Classification at Another Crossroads: Exploring E-book Call Numbers

by Daniel Saulean

Digital formats represent the materials on which today's libraries spend most of their budgets. As a result, a great deal of cataloging flows into these formats. Catalogers have been cataloging online content for some time. They adapted their general rules to the specificity of the Internet as soon as it became evident that end users appreciate the convenience of electronic resources. A key element of the cataloging process—classification—had a controversial status in the new electronic environment, with many questioning its benefit in searching for e-resources.

CLASSIFICATION AND CALL NUMBERS

In the early 1990s, new user tasks formally emerged and soon became a new direction in metadata and cataloging. Innovative catalogers embarked on the mission of elaborating appropriate rules to support the FRBR model, a process that expanded beyond traditional MARC21 cataloging. The essential tasks of finding, identifying, selecting and obtaining materials helped shape the structure of the new generation of faceted online catalogs (Tillet 2005). These faceted catalogs remain in service in many libraries even today. They bring precision and order to a collection of multiple formats, enabling discovery by means of filters. However, even when all resources benefit from the same metadata quality standards, there is a format-related dissimilarity when it comes to access. While the print book sits in an orderly fashion on the library shelf, an electronic one often resides on a vendor or publisher platform, where its full text is reachable from the catalog or discovery layer via a string of clicks, link resolvers, and filters that do not particularly entice readers.

Bibliographic classification is a valuable library service that supports user tasks. An alphanumeric class identifier becomes an effective element that guides the user directly to the item. It supports shelf browsing—a good old-fashioned method to discover similar literature based on the concept of collocation. At the same time, a class identifier embedded in the descriptive metadata facilitates collection development or user experience studies, turning into a tool for collection management and planning for better library service. Many collection analysis and stack maintenance techniques rely upon classification as a qualitative instrument of measurement.

Classification intertwines closely with subject analysis. The reason for which catalogers assign subject headings to resources is the same reason for performing classification. Subject analysis connects conceptual and descriptive cataloging. The call number is a byproduct of a classification based on one of the schemes: either LC, Dewey, Universal Decimal, NLM, SuDoc—to name the universal ones. It clusters together works, expressions, and manifestations, ensuring all of them are collocated in the physical as well as virtual space. Moreover, the call number is a unique access point that is complex enough to distinguish representations from one another. It is the core element of the item description and epitomizes granularity in organizing resources in the library.

Subject headings offer access points that are not unique resource identifiers. By contrast, when it comes to resource description, the classification numbers can be as unique as ISBNs; this is why, in practice, they serve multiple collection management and assessment purposes.

Daniel Saulean is Continuing Resources Metadata Librarian at the University of Vermont.

Classification is a standard procedure in creating the surrogate for books in a public catalog, which dates back to the beginning of libraries and cataloging. Call numbers, which grow out of classification through cuttering, represent traditional access points to which the patrons are generally accustomed when they carry out either a search or retrieval.

VIRTUAL SHELF BROWSING

Empirical studies maintain that the lack of visibility of e-resources makes researchers unaware of their existence, creating the assumption that collections are much smaller (Tingle and Teeter 2018, 165). Because e-books are not stored in the physical stacks, an integrated shelf of print with electronic books seems out of the question. Instead, the online catalogs and their hyperlinking capabilities make possible virtual shelf browsing, a simple and effective enabler of serendipitous discovery. Hence, the importance of including the call number information in the e-books metadata added to the catalogs regardless of type of resources, in order to ensure the application of uniform standards for discovery and access.

A number of functionalities within traditional integrated systems revolve around classification, at both their front and back end. Almost taken for granted in the print world, call numbers were a controversial topic from the very beginning of e-books cataloging. Many catalogers continued to create call numbers for monographs in electronic format, while others were quick to drop them off from the MARC records, arguing that these materials do not require physical shelving, even though practitioners demonstrated the benefits of online retrieval by call numbers from the very onset of the Internet (Chan 1986).

Searching the catalog on a given topic may be a challenge without librarians' help, especially when the subject terms belong to a controlled vocabulary. Librarians always underlined the value of a subject search tool, while end users and developers did not fully make sense of its purpose (Markey 2006). This type of search requires a librarian's expertise, which is customarily part of bibliographic instruction.

Subject analysis—the process leading to an individualized classification number for a given resource—may be deemed a cost-intensive activity with low expected return on investment. Additionally, it demands the maintenance of a database subset of authorized names and subject headings that are continuously added, updated, or cancelled. By contrast, a free full-text search by keyword or phrase across an e-books platform seems a much more cost-efficient alternative. However, would the mere existence of the full-text search be enough to render an entire science of assigning subject headings pointless? As long as the OPAC and e-book platform are two different databases, with separate indexes and limited interoperability, the answer is evidently no.

The precision of the results of a subject search based on controlled vocabulary outperforms the fuzzy outcome of a keyword search. The truth is that a subject search will always be a secondary option to a keyword one, particularly when subject indexes are made of pre-coordinated strings.

Online catalogs expanded the range and reliability of search methods, and call-number or shelflist browsing is among them. Similar to subject search, faceted browsing by class number is less popular. Nonetheless, librarians continue to use it and to promote it to patrons, as it brings validity to any literature search. The direct benefit would be locating books by consulting the LC schedules in advance and finding out what a library has in a specific class or subclass area. This is an effective help for someone interested in a topic already defined in the LC classification scheme,

but less effective when venturing into uncharted areas. After all, the effectiveness and universality of the search capability depend on the extent to which specific formats of library resources are consistently classified and represented in the catalog.

Call numbers for e-books provide an instrument for a virtual browsing facet. In reality, a call number is unlikely to be the first step in a catalog search, but, once a relevant one is found, it serves to expand the search by adding the surrounding numbers, which hyperlink to the titles classified in the same area. Users interested in enlarging their literature searches could take into consideration the titles within the same virtual location, which may represent resources in various formats. Therefore, a call number browsing facet can be precise, comprehensive, and serendipitous at the same time. This facet is more efficient and practical than a subject facet, on the condition that all the format types are represented in the catalog (Rose 2012). While call number functionality is still an option in classic catalogs, it is pervasively missing from publishers' platforms. An ad hoc survey of several online e-books platforms did not identify a call number browsing option for these portals, which is very much indicative of the disinterest of publishers in library-type classification.

ENHANCING E-BOOKS' USABILITY

To ensure the universal application of the rules, cataloging of electronic and print books should be equally consistent. In theory, the type of format should not affect the value added that the call numbers can bring to the end user. In practice, formats do not benefit from equal treatment, and there are differences in approaching classification for electronic as compared to print formats (Frederick 2016, 134).

Ever since e-books first made their way into public catalogs, end users are grappling with the difficulty of finding them. Some years ago, a study found only 1.3% of patrons consider e-books easier to locate than print (Jamali, Nicholas, and Rowlands 2009, 39). Researchers who looked at the underutilization of e-books by undergraduate students emphasized the importance of making e-books as findable as other materials. Findability of e-books starts with their inclusion in the catalog (Lamothe 2013, 40) and depends on a spectrum of practical usability features as well as the user-friendliness of the discovery tool (Potnis, Deosthali, Zhu, and McCusker 2018, 115). The completeness of the surrogate metadata creates conditions for better visibility and higher positioning in the ranking algorithms, enhancing overall discovery and usability.

Studies argue that end users favor electronic over print if provided with choice options. As suspected, when e-book metadata is not present in the library catalog, the usage of e-books is low (Goodwin 2014; Yuan, Van Ballegooie and Robertson 2018). The surrogate metadata for the e-books that libraries receive from various publishers goes, most of the time, directly into the public catalogs, sometimes with or without local editing. Most likely, an e-book record with a missing classification number will be loaded as-is. Due to diminishing cataloging expertise in libraries at a time of contracting budgets, libraries accept substandard records with the hope of enhancing findability, ignoring the fact that deficient metadata has an underwhelming impact on discovery. When libraries acquire vendors' sets of e-books, many of them prefer activating the titles directly in their discovery layers, bypassing the OPAC and relying entirely on non-edited publisher metadata that regularly lacks class number info. The likelihood of discovery layers to accommodate call numbers is also very low, hindering once more the call number browsing facet.

UNCERTAINTY OVER THE FUTURE OF CLASSIFICATION

History suggests that the theory and practice of library classification does not necessarily intersect. Controversies over bias and hierarchies of concepts representing classes and subclasses on one hand and the impracticality of classifying e-resources by traditional means on the other are evidence of the challenging times professional standards undergo nowadays. Conserving a library classification system that flourished within the MARC environment and was designed to serve the print world seems rather out of place in the context of the semantic Web. Already, in many libraries, newer generations of discovery layers do not display the call number data for e-resources. Even more, some libraries adopted a simplified, bookstore type of collections organization that leaves behind the classic schemes catalogers developed over time, which became esoteric to the public.

A check of the newly issued cataloging rules reveals a diminishing interest in classification within the professional community as well. The new RDA Toolkit index has no entry for classification or call numbers. The RDA Registry, which maps RDA elements and relationships to linked data, does not comprise classification as such. None of the hundreds of newly created vocabulary entries and properties related to classification or call number is among the elements or values listed in the registry. Yet BIBFRAME 2.0 developed a number of properties describing subject and classification numbers, which may stem from a universal scheme or controlled vocabulary. Conversion of bibliographic records created according to previous cataloging standards may conserve the pre-existing data, including classification, although LD4P does not express particular commitment to continuing the creation of class numbers.

As long as the classic catalogs or next-gen ones are still around, the call number will largely persist. If the libraries abandon the catalog and move away from its traditional information organization principles, established in an exclusively print environment, call numbers would likely fade away. It is unclear which model the catalog will transition to, if any, but change is in the air.

Both anecdotal and Web traffic data shows that OPAC searches are down from year to year in academic libraries. This is valid at my home institution—the University of Vermont—but there are signs that the trend goes across the board. Several years ago, some began openly advocating for the disappearance of the classic catalog and discontinuing investment in discovery systems by replacing them with Google Scholar, other similar search engines, or straight-up research databases (Kortekaas and Kramer 2014). A library-without-a-catalog concept emerged, although without gaining much popular traction thereafter. The local catalog without users, highly expensive to maintain, represents the driving force behind the bold move, a new reality that libraries and librarians seem to face by looking at sinking traffic statistics and failure to attract patrons from new generations.

Cataloging without a catalog will no longer be business as usual. A longtime sacred cow of cataloging, classification is already at a crossroads and in a vulnerable position. It is hard to justify investing in such a complex and politically controversial activity with diminishing popularity among users. The next-gen catalogers show no appetite for learning the old rules. As a likely result, the emerging models of library services platforms will gradually phase MARC21 out, driving classification into a corner, unless machines will start learning subject analysis, traditionally done by humans, and save call numbers from extinction, too. There are hopes that information technologists will rescue library classification through an automated taxonomic analysis of digital content—a protracted change that startups such as Yewno recently put in practice. Artificial intelligence, rath-

er than human beings, will likely lead the revolution about to unfold in the field of library classification (Schreur 2020).

LOOKING ANXIOUSLY AHEAD

While libraries have long striven to bring the Web into their catalogs, the Web always wanted to bring the libraries and their collections into its realm. The tension between the libraries, which serve a group of local patrons, and the Web, which brings together the whole world, will continue in the years to come. When the time is ripe to complete the migration from MARC21 to linked data, libraries will have to decide the fate of their complex legacy metadata, and an important part of the transformation will be what to do with their classification systems.

WORKS CITED

Chan, Lois Mai. 1986. "Library of Congress Classification as an Online Retrieval Tool: Potentials and Limitations." *Information Technology and Libraries* 5, no. 3: 181–92.

Frederick, Donna E. 2016. *Managing E-book Metadata in Academic Libraries: Taming the Tiger*. Chandos Publishing.

Goodwin, Cathy. 2014. "The e-Duke Scholarly Collection: E-book v. Print Use." *Collection Building*. doi.org/10.1108/CB-05-2014-0024.

Jamali, H. R., D. Nicholas, and I. Rowlands. 2009. "Scholarly E-books: The Views of 16,000 Academics: Results from the JISC National E-Book Observatory." *Aslib Proceedings* 61, no. 1: 33-47. doi.org/10.1108/00012530910932276.

Kortekaas, Simone, and Bianca Kramer. 2014. "Thinking the Unthinkable: Doing Away with the Library Catalogue." *Insights* 27, no. 3: 244–8. doi.org/10.1629/2048-7754.174.

Lamothe, Alain R. 2013. "Factors Influencing the Usage of an Electronic Book Collection: Size of the E-book Collection, the Student Population, and the Faculty Population." *College & Research Libraries* 74, no. 1: 39–59. doi.org/10.5860/crl-301.

Markey, Karen. 2006. "Forty Years of Classification Online: Final Chapter or Future Unlimited?" *Cataloging & Classification Quarterly* 42, nos. 3–4: 1–63.

Potnis, Devendra, Kanchan Deosthali, Xiaohua Zhu, and Rebecca McCusker. 2018. "Factors Influencing Undergraduate Use of E-books: A Mixed Methods Study." *Library & Information Science Research* 40, no. 2: 106–17. doi.org/10.1016/j.lisr.2018.06.001.

Rose, Mary Z. 2012. "The Ship Has Sailed and We Aren't on It: How Catalogers Could Support User Tasks and Why We Won't." *Journal of Library Metadata* 12, nos. 2–3: 127–39. doi.org/10.1080/19386389.2012.699828.

Schreur, Philip E. 2020. "The Use of Linked Data and Artificial Intelligence as Key Elements in the Transformation of Technical Services." *Cataloging & Classification Quarterly* 58, no. 5: 473–85. doi.org/10.1080/01639374.2020.1772434.

Tillett, Barbara B. 2005. "FRBR and Cataloging for the Future." *Cataloging & Classification Quarterly* 39, nos. 3–4: 197–205.

Tingle, Natalia, and Keith Teeter. 2018. "Browsing the Intangible: Does Visibility Lead to Increased Use?" *Technical Services Quarterly* 35, no. 2: 164–74. doi.org/10.1080/07317131.2018.1422884.

Yuan, Weijing, Marlene van Ballegooie, and Jennifer L. Robertson. 2018. "E-books versus Print Books: Format Preferences in an Academic Library." *Collection Management* 43, no. 1: 28–48. doi.org/10.1080/01462679.2017.1365264.