Purdue University

Purdue e-Pubs

Charleston Library Conference

Cooperative Collection Development Requires Access: SALToC—A Low-Tech, High-Value Distributed Online Project for Article-Level Discovery in Foreign-Language Print-Only Journals

Aruna P. Magier
New York University, aruna.magier@nyu.edu

Follow this and additional works at: https://docs.lib.purdue.edu/charleston

Part of the Library and Information Science Commons, and the South and Southeast Asian Languages and Societies Commons

An indexed, print copy of the Proceedings is also available for purchase at:
http://www.thepress.purdue.edu/series/charleston.

You may also be interested in the new series, Charleston Insights in Library, Archival, and Information Sciences. Find out more at: http://www.thepress.purdue.edu/series/charleston-insights-library-archival-and-information-sciences.

Aruna P. Magier, "Cooperative Collection Development Requires Access: SALToC—A Low-Tech, High-Value Distributed Online Project for Article-Level Discovery in Foreign-Language Print-Only Journals" (2014). *Proceedings of the Charleston Library Conference.*
http://dx.doi.org/10.5703/1288284315562

Cooperative Collection Development Requires Access:
SALToC—A Low-Tech, High-Value Distributed Online Project for
Article-Level Discovery in Foreign-Language Print-Only Journals

Aruna P. Magier, South Asia Librarian, New York University

Abstract

Foreign-language journals are an essential component of interdisciplinary area studies collections at research libraries but are, by definition, low-use materials. Librarians who select them seek to broaden these collections, reduce duplication, and enable shared access to them. The challenge is lack of article-level discoverability: these are print-only journals, not covered in online indexing/abstracting services. If users cannot discover these articles, then how can cooperating libraries share them, and distribute responsibility for collecting them, which is essential to coordinated collection development?

The SALToC project collaboratively address this issue by creating simple, centrally browsable tables of contents for target journals, through a low-tech, low-cost distributed process that benefits users at all participating libraries. For journals not available online nor included in article databases or indexes, this kind of discovery facilitates research by enabling scholars to use previously undiscoverable holdings of other libraries: they can now issue interlibrary loan, document delivery, and/or offsite retrieval requests, with full citations for desired articles. (Many libraries provide article document delivery, if the requester has a citation). Coordinated collection development (via planned reduction of duplication coupled with broader collective coverage) becomes supportable in the research library community only when shared access (and its prerequisite—discovery) is provided. The South Asian Language Journals Table of Contents (SALToC) project represents a proof-of-concept demonstration of the value of this approach. This paper shows how simple, "grass-roots" distributed efforts can contribute significantly to discoverability of hard to discover resources, thereby making coordinated collection development cost-effective, popular among users, and sustainable.

Introduction and Rationale

It is well understood that for any given unit of content that libraries make available to their patrons, the actual ability to use that content depends on discovery. If you can't discover that it exists, you can't locate it, can't request it, retrieve it, or use it. As many have observed, discovery is a prerequisite for access.

In the same way, if parts of a library's collections are not exposed for discovery, they also cannot effectively be shared with other libraries— because they cannot be requested by patrons at the partner institutions. In other words, discovery—or the lack of it—*delimits* what libraries can share with each other or borrow from each other. Stated from the patron's perspective, the key benefit of shared collections (access to more needed content) can only be derived for materials exposed for discovery.

In recent times, many libraries have responded to tight budgets by cutting back on their own collecting, relying more heavily on the collections of other libraries to help fill their users' needs. When groups of libraries that engage in collection sharing then coordinate their collection development going forward, they can reduce unnecessary duplication, redirect resources to broaden the scope of the community's aggregated shared collection, and increase the likelihood of being able to respond to their patron's needs on the whole. But the entire premise of coordinated collection development relies on access to one another's collections, which, as I have said, presupposes discovery.

But discoverability is not a yes-or-no characteristic. There are various kinds of exposure that libraries can provide, for different kinds of material, yielding different *degrees* of

discoverability, and requiring different amounts and kinds of investment to make the discovery possible. For example, at the low-investment end of the spectrum, simply having a book or journal issue present in the open stacks makes its content at least theoretically discoverable via browsing. At the other end of the spectrum, full-text digital presentation of collections of content, with full-text indexing, multiple descriptors and added-value access points, cross-referenced authority files and thesauri, etc., exposed through an intuitive interface for searching and browsing, with facets for narrowing or broadening one's search, relevance-ranking of results, presentation of related materials, and so on, all make the units of content much more discoverable. But, of course, creating the metadata and mechanisms that underlie that kind of discovery costs much more money!

In this paper, I explore the relationships between discovery and cooperative collection development. I focus on a particular category of content for which libraries have traditionally not invested much in its discovery—if they collect it at all. I am speaking here of foreign language journals, and especially print-only journals.

Many countries of the world publish a broad range of journal literature of importance for research, in the languages of their regions. While the best of these are considered to be an essential component of interdisciplinary area studies collections at US academic research libraries, they are often, by definition, low-use, "obscure" material, and it makes sense for the librarians who collect them to seek to broaden these collections in a coordinated way, reducing duplication nationally, and enabling shared access to them. But the problem has been lack of discoverability of the articles in these journals: they are print-only, not covered in online or printed indexes or indexing/abstracting services. If users cannot discover the articles in these journals (except via physical browsing), then how can cooperating libraries share them, and distribute responsibility for collecting them, which is essential to coordinated collection development?

The rise of e-journals in general has enabled vast increases in discovery at the article level. Before

e-journals existed, libraries created title-level discovery of the journal itself through creation of a single bibliographic record, with subject headings, etc. A patron could discover that a journal on a certain topic existed, but the only way to find individual *articles* was to physically browse the holdings. Occasional printed author- and subject-indexes brought out by a journal's publishers enhanced article-level discovery, later followed by third-party indexing and abstracting services that enabled article-level discovery across multiple journals. By and large, such bibliographic-only databases have been supplanted or made redundant by the rise of full-text e-journals with publisher-level cross-journal article discovery through the publishers' or aggregators' interface, and integrated discovery systems across all those systems at the higher end of the spectrum as described above.

But each improvement in article-level discovery along that continuum left some journals behind, as discovery-providers deemed that content not worth the investment needed to bring it up to the current high-end. The journals generally left furthest behind in the process are foreign-language journals in general, and those that are not publishing online in particular. Articles in print-only foreign journals are all too often stuck back in the dark ages of discovery.

Even where libraries have continued to collect them for the inherent research value of their content (and to suit the local research and teaching priorities), bibliographic access at the article level has only been possible by physical browsing. While there are large differences from country to country, due to differences in the penetration of online publishing, area studies librarians report that vast swaths of this kind of content is not only not online, it is not even indexed anywhere. For example, my colleagues among Middle East Studies librarians report almost none of the periodical literature from that region is online or indexed (in print or online fashion). The same is largely the case with journals in the languages of Africa, South Asia, Southeast Asia, much of Eastern Europe and Central Asia, the Caribbean, and even parts of Latin America.

Because this content is in foreign languages, of course, its usage is much lower than that of journals in English or other commonly taught languages. But even those students and scholars who can read those languages, and whose research would benefit from access to these articles, tend not to use them because discovery is so difficult compared to our growing expectations of keyword searching or "click-and-read."

Not surprisingly, libraries respond to low usage in several ways: if the constituency for the content is complacent, the library may simply stop collecting these journals, and withdraw the backfiles to save space. Or, perhaps less radically, they will move these print journal runs to compact, off-site storage, for retrieval on request. But for an unindexed print-only journal, this solution enacts a self-fulfilling prophecy of zero usage. Held offsite where they cannot be browsed, no one will discover any articles in these journals, no one will use them at all, and a history of nonusage will inevitably lead to withdrawal, to save on the costs of storage.

To the extent that foreign print-only journal content has value for research, one could say that leaving these journals behind—so to speak—is a shame. If one wanted to address the problem, a number of approaches might be possible. For example, working with the professional organizations of area studies librarians—such as the Journals Subcommittee of CONSALD (Committee on South Asia Libraries & Documentation), JSTOR has been seeking to identify print journals from abroad that are priorities for negotiation with publishers to get their runs digitized and included in JSTOR. JSTOR's coverage of foreign journals from places like South Asia with vast print-only publishing industries has thus been growing gradually in recent years. But even there, the growth has been in English-language journals (the largest language of publishing in South Asia, as in many parts of the world). So the foreign language journals yet again get left behind.

The question then arises of whether or how to create article-level access to these print-only orphaned journals. Attempts have been made in this direction, sometimes through cooperative efforts of librarians who try to create some of the same kind of indexing and abstracting value seen in an earlier era for English-language journals. But, creating article-by-article metadata for such databases is quite expensive, especially when you consider the need for language-expertise to go along with the indexing or cataloging skills needed to do it.

In an era when even large research libraries working together are having trouble securing the human resources to do original cataloging of their backlogs of books and journals, how likely is it they could prioritize the indexing or abstracting of the vast number of individual *articles* in all those demonstrably low-use, specialized foreign language journals? With some notable exceptions (like HAPI, Hispanic American Periodicals Index published out of UCLA), these attempts have been spotty, limited in scope and/or hard to sustain. The cost-benefit analysis tends to make these efforts hard to justify.

I am now going to provide a case study of another new, experimental approach being undertaken by a collaborating group of South Asia librarians, targeted at a selected group of print-only journals in languages of South Asia. It is called SALToC: the South Asian Languages Cooperative Tables of Contents project. It was designed to be a low-tech, low-investment, distributed online project to enable article-level discovery towards the low end of the discovery continuum: discovery by online *browsing*—which is certainly incrementally better than no discovery, and which provides, as I will show, quite valuable benefits for access and cost-efficiencies for coordinated collection development.

A distributed low-cost system of creating simple, centrally browsable tables of contents in a sustainably accessible infrastructure with low-institutional barriers can facilitate research by finally enabling scholars to make use of these previously undiscoverable journal holdings. Discovering articles this way makes it possible for them to issue interlibrary loan requests, document delivery requests, and offsite retrieval requests, with full citations for the desired articles. While many libraries do not send lend print journals on interlibrary loan, many do offer

article-level document delivery on request, but only if the requesting institution provides a *full citation*.

The "grass-roots" distributed nature of the SALToC approach I will describe here offers a proof-of-concept demonstration that coordinated collection development and reduction of unnecessary duplication can be sustainable, popular among users and cost-effective in the research library community through the shared access made possible by this low-lying discovery layer.

Let us look at the SALToC model in some detail to see how it works, and what its values and weaknesses might be.

Goals of the Project

The group of South Asia librarians, representing the needs of their respective South Asian Studies constituencies, worked with me to develop a set of operating criteria and goals for a joint project. The main goal of SALToC was to enable article-level discovery of vernacular language journals from their collections that were identified as not otherwise discoverable, because they are not online and not included in existing full-text or bibliographic databases. Key objectives included:

- Enabling patrons to identify and access articles in their own collections or request them from other libraries through standard ILL.

- Enabling patrons to use citations to request journal articles from offsite storage.

- Enabling cooperating libraries to provide digital document delivery ("scan and deliver" service) for articles in these vernacular journals, just as they already are doing for print journals in English.

- Providing an online substitute for physical browsing.

- Enabling offsite storage decisions for these low-use journals.

In contrast to other, more expensive discovery systems for mainstream materials, we also decided to clarify what SALToC is *not* meant to be. It is

- Not a journal article indexing project.

- Not an indexing and abstracting service.

- Not a searchable, structured database of citations.

- Not a table of contents alert service.

Process

The concept was to find a way to keep this as low-tech and low-cost as possible, while still providing real discovery value. The steps are simple, and do not require any special skills or highly trained staff at each stage of the process. Once the infrastructure was set up, the actual ongoing production is carried out almost entirely by student assistants who do not need to know the relevant languages.

1. Each participating library makes simple page-image PDF files by scanning the table of contents of each issue of the target journals they are contributing.

2. They annotate each issue's PDF file with a couple of basic fields pulled from their existing catalog record and volume holdings: title, imprint, volume, issue number, date.

3. These annotated PDFs are sent to me at the central SALToC repository at NYU.

4. At NYU, these SALToC tables of contents files are stored as part of the University's institutional repository in the DSpace platform, with a separate permanent URI for the separate landing page for each journal, with its accumulating tables of contents, and acknowledgement of the contributing library.

5. *A link to permanent URI for each journal's landing page is then added to the regular catalog record for that journal in the contributing library's own OPAC (to*

enable discovery of the ToCs by their own local users).

6. NYU catalogers then update the WorldCat record for each SALToC journal to add the same link to the ToCs there (to enable discovery of the ToCs by everyone else).

7. A simple DSpace collection page for whole SALToC project, listing and linking all the journals included, is also maintained in NYU's repository, to allow the participating librarians to promote and highlight SALToC to their patrons (through LibGuides, bib instruction, etc.), beyond the linkage for each title provided through OPACs and WorldCat.

Participants

Current participating libraries contributing ToCs to SALToC include University of Hawai'i, University of Pennsylvania, Princeton, Columbia, Yale, University of Chicago, University of Washington, Center for Research Libraries, and Library of Congress overseas field office in Islamabad. Several other libraries are now preparing to start contríbuting ToCs as well. The home of the SALToC project is New York University, whose contribution to the collaboration is hosting SALToC on their existing DSpace platform as part of their institutional repository, and having students add each new ToC to the collection as it comes in.

Workload

The distribution of production work among the participants is simple. Contributing libraries scan the ToC of each issue of their journal as received (about 10 minutes), and submit it to NYU. They also—just once for each journal—have their catalogers update the local OPAC record to add the 856 link to the permanent URI for that journal (less than 5 minutes).

At NYU, the ongoing workload is just adding each PDF file to the DSpace repository as it is received from the contributors (less than 5 minutes). The one-time work for setting up each journal title in SALToC is just creating its permanent landing page in DSpace (containing the brief bib info from the

annotation, a link to the WorldCat record, and acknowledgment of the contributor), which takes about 10 minutes, and updating the master record for the journal in WorldCat (less than 5 minutes).

Sustainability

Our goal was sustainability. Among other reasons, we wanted to make sure that the catalogers' one-time-per-journal work of adding links into OPAC and WorldCat records (simple as it is) would not need to be updated or revised later. While catalog links to web content are notoriously ephemeral, and tend to go out of date and become dead-end links very quickly, we wanted to reassure the catalogers that they were linking to permanent content at a permanent address. We addressed overall sustainability in several ways:

- SALToC uses a light-weight, low-tech maintainable infrastructure.

- It requires minimal resource investment in terms of human workload and system resources.

- It provides demonstrable value for our patrons.

We also managed to avoid reinventing the wheel. We created what is undeniably a niche product (that fills a specific need not otherwise filled for South Asia vernacular journals), but *not* based on a separate niche infrastructure. Basing SALToC in the University's DSpace required no specialized programming, workflow, database, or server maintenance. It uses a system already put in place and maintained to do what the University is already committed to doing anyway: a permanent repository for faculty, with long-term institutional commitment for permanence and permanent URLs. So SALToC gets to leverage the value of that existing infrastructure without adding any extra cost.

Scalability

How scalable is this model? What are the potential limits on its growth? SALToC is scalable because contributor institutions add as much or as little as they want. The barriers to entry are exceedingly low, and so far, the decisions about

making these minor investments have been kept very close to the "grass roots" level (i.e., the front-line area studies librarians who work most closely with the scholars who benefit from SALToC). But each contribution (table of contents) adds *incremental* value to SALToC, creating permanent discoverability for the corresponding articles in the libraries' print holdings, via the links in the OPACs and in WorldCat records. Now that this infrastructure is in place, the work to insert each successive contribution is negligible: it is completed in mere minutes by nonspecialist staff and students.

So SALToC seems to be both sustainable and scalable. In contrast, projects that have attempted to create discovery at higher points on the continuum, for example by creating searchable, structured databases of full article citations (like LAPTOC, the Latin American Periodical Table of Contents project) require actual data entry for each article at each participating institution. This dramatically increases the cost of production, making the project less sustainable or scalable. Similarly, the South Asian Periodicals Index at University of Wisconsin is also a manually constructed article-by-article citation database. It ceased production after six journals (all but two of them in English).

Learning from projects such as LAPTOC and SAPI, participants in SALToC chose to use meet the needs of discovery through *browsing*: mere page images, with no data-entry and no language skills required. After only six months in production, SALToC already contains about 300 tables of contents files in six languages of the Subcontinent.

Going forward, the South Asia Librarians have been conferring together to develop consensus criteria for selecting and prioritizing additional journal titles from their collections: periods of coverage and completeness of holdings, target subjects, target languages, and soliciting faculty input to identify specific high-value titles for inclusion.

Conclusion

In designing the SALToC model, we have tried to find an appropriate value point between the ideal and the real. Recognizing that article-level discovery is a matter of degree, and that the high-end of the continuum (the ideal) would require high levels of investment that would skew the cost-benefit ratio making the whole enterprise unsustainable in institutional contexts of constricted resources, we sought instead to create something much more modest, of incremental value that could be sustained because it is unlikely to fall to the cost-cutting axe.

It is too soon to conduct a full-scale evaluation and accounting of fully loaded costs and values, costs-per-use, impact factors, etc. And we have already heard from some that comparing SALToC to other resources they are familiar with (for example, JSTOR), it clearly seems home-grown and improvised, and lacking in features they love in those high-end productions. On the other hand, we are also receiving enthusiastic reports of use cases of how SALToC is enabling scholars to delve into the journal content we are collecting for them across our cooperating institutions, in ways that seem to validate the premise of coordinated collection development. So we are encouraged to stick to our SALToC slogan: *Don't let the excellent be the enemy of the good!*

References

SALToC: South Asian Language Journals Cooperative Table of Contents Project. https://archive.nyu.edu/handle/2451/33560 (accessed 12/5/14)

Magier, Aruna P. 2014. Cooperative collection development requires access—SALToC (Charleston 2014 presentation SlideShare). Retrieved from http://www.slideshare.net/arunamagier/aruna-magier -charlestonpowerpoint