Creating Article-Level Discovery of Print-Only Foreign Language Journals

A Case Study of SALToC’s Distributed Approach

Aruna P. Magier

Foreign language journals are important components of interdisciplinary area studies collections at research libraries. In the US, although these are low-use materials almost by definition, they are indispensable for many types of research. Coordinated collection development among key libraries with shared interests in these materials is often the best way of broadening the collective collection, strategically reducing duplication to free up resources for broader acquisitions while relying on collection sharing infrastructures to implement shared access to the journals.

Collection sharing requires bibliographic access. With journals, sharing is usually via interlibrary loan and electronic document delivery, which require article-level bibliographic access for formulating the citation for a request. The challenge in providing article-level discoverability of print-only foreign language journals is that they are generally not included in online indexing/abstracting services. A frequent alternative for article discovery of unindexed journals is physical browsing, but many libraries do not circulate bound journals or will not lend them via interlibrary loan (ILL). Even for the institution’s own patrons, browsing is difficult because such low-use materials are typically sent to off-site storage. The ability to engage in cooperative collection development of low-use, print-only, foreign language journals is limited because of inadequate possibilities for article-level discovery, which are the prerequisites for collection sharing.

The South Asian Language Journals Cooperative Table of Contents (SALToC) project discussed in this paper addresses this issue by creating simple, online, centrally browsable, open access tables of contents for target journals via a low-cost, low-tech distributed process benefitting researchers at all libraries. SALToC also demonstrates the applicability of this approach for creating article-level discovery for any unindexed journals and is not limited to foreign language journals.

The SALToC model promotes research and leverages the value of otherwise undiscoverable library holdings. Many libraries provide scan-and-deliver services if the requester provides a citation. With the help of citations discovered
through SALToC, scholars can thereby access articles from these otherwise inaccessible print journals through document delivery, ILL, off-site retrieval, and other fulfillment services for which libraries already have the infrastructure. Beyond this immediate benefit for discovery and access, the SALToC model provides the secondary benefit of making cooperative collection development supportable. The 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 access resources, thereby making cooperative collection development cost-effective, popular among users, and sustainable.

**Background: the Context of Coordination in Collection Development**

Many academic libraries have responded to tight budgets by cutting back on the collecting of more specialized, lower-use materials, hoping to rely more heavily on other libraries’ collections to help fulfill their users’ research needs for materials they no longer collect. This is particularly true in the humanities and area studies, as observed by Jakubs: “It is not surprising that the cooperative arrangement has worked well with area studies and less well with the sciences, given that scholars whose work relies on difficult-to-find resources are pleased to have them at all and are more willing to be dependent on a partner institution, while scientists have a more urgent need for the materials and expect them to be in the campus library.”

When groups of research libraries coordinate collection development, they can thereby strategically reduce unnecessary duplication, redirect resources to broaden the scope and research power of the community’s aggregated shared collection, and increase the likelihood of being able to supply the materials that respond to their patron’s needs on the whole. But coordinated collection development by itself cannot produce these benefits: they are possible only when collection planning is implemented within an effective system of shared access to each other’s collections. Cooperative collection development can be achieved by the intersection of coordinated collection development, resource sharing, and bibliographic access.

**The Need for Discovery**

The ability of patrons to make effective use of library collections depends on discovery. In the absence of means of discovering the resources a library makes available, the patron cannot locate, request, retrieve, or use them. Discovery is a precondition for access. Conversely, lack of discovery translates into lack of use.

In the context of cooperative collection development, lack of discovery limits a community of libraries’ ability to share and borrow from each other. If parts of a library’s collections are not exposed to discovery, they will not be used by the library’s patrons and also cannot be shared with other libraries since patrons at partner institutions are unable to request them. From the patron’s perspective, the main advantage of shared collections and access to them is only achieved when these resources are exposed for discovery. Within a system of collaborating libraries, the value of discovery is a function of the breadth of potential use it enables across the populations of the institutions engaging in resource sharing. Having made the investment in acquiring and maintaining collections, it is in a library’s interest to leverage their value by providing discovery and access in the ecosystem of cooperative collection development.

**Discovery Mechanisms and Cost-benefit Analysis**

The value of discovery must be balanced with the costs. Discovery at various levels is provided through a variety of mechanisms, each with its own associated costs and benefits. Libraries provide various kinds of discovery tools for different kinds of resources that require differing amounts and kinds of investment to enable discovery. An instance of a low-investment option is having a book or journal issue in open stacks, making its content theoretically discoverable via browsing. Many have written about the value of serendipitous discovery through browsing. Conversely, full-text digital presentation of collections of content, with full-text indexing, multiple descriptors and added-value access points, thesauri, cross-referenced authority files, citation-linking, 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, etc. make the units of content much more discoverable. The monetary and staffing costs associated with creating the metadata and mechanisms underlying that kind of discovery are much higher.

What is the appropriate level of investment to deploy along this continuum of discovery for a given universe of content? Whether from the publisher’s perspective, aggregate’s business plan, or from that of the library community’s desire to leverage the value of its own holdings for its own patrons, such decisions are presumably made on the basis of anticipated use triggered by the discovery model. This means that potential use of any content is at least partly a function of the likelihood that patrons would use it if they could discover it. These kinds of cost-benefit calculations (better discovery means more usage, which means
either more profit [commercial business plan] or better service [library mission]) are clearly not infinitely scalable. For any content, usage will plateau at some level regardless of how much was invested in top-notch discovery. An idealized cost-benefit analysis for any content would also need to consider the theoretical level where usage would level off. Regarding the target content addressed in this paper (unindexed, print-only, South Asian language journals in US libraries), the most likely determinant of that plateau is language: the potential audience for the content cannot be larger than the number of people who can read the language, regardless of the quality of the discovery. Since that number in the US will be lower than the English-reading audience, it is safe to guess that the return on the discovery investment will plateau at a much lower point for these journals than for the much larger number of English-language journals these libraries hold. The “right” level of investment will likely be quite low for foreign language journals.

As described below, a group of South Asian studies librarians attempted to address the challenges of producing a much-needed quantum of article-level discovery for scholars of South Asian studies—at an appropriately low level of investment—by creating SALToC (see https://archive.nyu.edu/handle/2451/33560). As an open access, low-cost table of contents discovery mechanism for print-only journals in South Asian languages, SALToC is also an effective tool for these librarians to coordinate their collection development investments. The hypothesis is that this model represents an appropriate level of investment to produce a much-needed quantum of discovery for this content.

**Foreign Language Periodicals in Research Libraries**

Many countries publish a broad range of periodical and journal literature of importance for research in the languages of their regions. While the best of these are considered as an indispensable component of interdisciplinary area studies collections at US academic research libraries, they are often, by definition, low-use, “obscure” research materials. They frequently share the fate of other low-use materials, which is to be omitted from prospective collection development policies. With libraries increasingly focusing on their “core” needs, usage data, patron-driven decision-making, etc., librarians with the mandate to ensure access to these kinds of resources for their area studies constituencies coordinate with their counterparts at other libraries. Through careful coordination, they can strategically reduce duplication regionally or nationally, and redeploy greatly needed resources. Depending on collection development philosophies and missions, these resources can be counted as bottom-line savings, redirected locally for other collecting priorities or used to target additional needed journals to round out the collective collection, relying on collection sharing infrastructures to implement shared access to the journals among the group.³

The challenge to this coordination is the general lack of discoverability of papers in the journals collected through such cooperative arrangements. If users cannot discover papers in these unindexed off-line journals (except via physical browsing), then how can cooperating libraries share them and distribute responsibility for collecting and retaining them? Answers to these questions are essential to coordinated collection development.

The rise of e-journals in general has facilitated vast increases in discovery at the article level. Before e-journals existed, libraries facilitated title-level discovery of journals through creation of a single bibliographic record with subject headings and access points. A library patron could discover a journal on a particular topic, but the only way to find individual articles was to physically browse the holdings or citations from other articles. Occasional printed author and subject indexes produced by a journal’s publisher enhanced article-level discovery, later followed by third-party indexing and abstracting services that enabled article-level discovery across multiple journals. Usually, such bibliographic-only databases have been more recently 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 discovery spectrum as previously described.³

Each improvement in article-level discovery along that continuum omitted some journals as discovery-providers deemed that content not worth the investment necessary to raise it to the current high-end. For the reasons outlined above, foreign-language journals in general are often the most overlooked in this process, particularly those that are not available as e-journals. Because of language, their potential for increased use seemed too low to warrant attention. When libraries have continued to collect and retain them for the inherent research value of their content (and to serve local research and teaching priorities), bibliographic access at the article level has generally continued to be possible only by physical browsing. While countries differ greatly in their level of technological development, ubiquity of internet access, and prevalence of online publishing, area studies librarians report that much of this kind of content is not available online and not indexed.¹⁰

This is largely the case with journals and periodicals in the languages of Africa, South Asia, Southeast Asia, much of Eastern Europe and Central Asia, the Caribbean, and even parts of Latin America. Whereas Troost reports that there is broad online accessibility of Chinese and Korean
Asian studies scholars struggle to discover relevant information and to keep up with the deluge of publications. Differing publication conventions and categorization systems in Asia can make it difficult for scholars to employ the same discovery methods as they do with Western publications…. One of [the] challenges is that rates of digitization and the ability to access digital materials varies widely within and among regions of Asia. For instance, scholars researching in East Asia generally found that Korean institutes “scan a lot and they put up a lot [of] stuff online, so it’s just really accessible,” whereas in Japan and Taiwan, remote access to databases and other digitized government documents is only available to citizens of those countries, if not only in-person. In other regions, such as in South Asia, “there’s not a lot of stuff online” despite increased digitization.12

With regard to South Asia, Rader observes, “A number of factors might have determined the dearth of discoverability of periodical literature from South Asia—a lack of broad commercial profitability, a complicated colonial/post-colonial relationship with English, lagging technology to effectively display and search materials in non-roman scripts. This glaring gap in easy discoverability however should not be equated with an absence of content of interest for writers and researchers.”13

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

How do Libraries Treat Low-usage Print Journals?

Not surprisingly, due to the high costs of maintaining print collections, libraries respond to low usage by moving these print journal runs to compact, off-site locations for retrieval on request.14 In some libraries, if the constituency for the content is complacent, the library may simply cancel their subscriptions to these journals or deaccession the existing backfiles to save space. But for an unindexed print-only journal, moving the holdings to remote storage enacts a self-fulfilling prophecy of zero usage.15 Held off-site where they cannot be browsed, no one will discover articles in these journals, no one will use them, and a history of non-usage will inevitably lead to deaccessioning to save on the storage costs.

Approaches to the Problem

Clearly, since research libraries continue to collect them for their research value, omitting foreign language print-only journals from the trend towards better article-level discovery is a problem. Many approaches to address this problem have been attempted. For example, working with the professional organizations of area studies librarians such as CONSALD’s (Committee on South Asia Libraries & Documentation) Journals Subcommittee, JSTOR has sought 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 enterprises has thus been growing gradually in recent years. However, the growth has only been in English-language journals (the largest language of publishing in South Asia and many other parts of the world).

Approaches from Scholarly Societies

One might anticipate that specialized scholarly societies, representing constituencies with a great stake in discovery and access to these materials, would make the necessary investment to start filling that gap. An example is the Association for Asian Studies (AAS)—an international scholarly professional association of over 7,000 members—to serve the interdisciplinary needs of broad Asian Studies research.16 Among its activities in support of this field, AAS has a broad publishing agenda, including the premier bibliographic reference tool for Asian Studies—the Bibliography of Asian Studies (BAS).17 The BAS was produced as an annual print index from 1936 through 1990. With a grant from the Mellon Foundation, the BAS was transitioned to its current incarnation as a cumulative online database with nearly a million records, originally hosted at the University of Michigan (UM).18 To support production and administrative costs and the salaries of paid indexers, AAS charged institutional and individual subscription rates. Nonetheless, the costs and complexities of its production, and the challenges of updating the database’s infrastructure and capabilities, could not be indefinitely sustained within the original online framework of informal academic institutional agreements at UM.19 In 2016, AAS decided to transfer the BAS’s infrastructure and hosting to EBSCO

11 A 2018 study commissioned by Ithaka S+R entitled “Supporting the Changing Research Practices of Asian Studies Scholars” notes that subscriptions to these journals or deaccession the existing content is complacent, the library may simply cancel their request.

12 Print journal runs to compact, off-site locations for retrieval on request.

13 Not surprisingly, due to the high costs of maintaining print collections, libraries respond to low usage by moving these print journal runs to compact, off-site locations for retrieval on request.

14 In some libraries, if the constituency for the content is complacent, the library may simply cancel their subscriptions to these journals or deaccession the existing backfiles to save space. But for an unindexed print-only journal, moving the holdings to remote storage enacts a self-fulfilling prophecy of zero usage.

15 Held off-site where they cannot be browsed, no one will discover articles in these journals, no one will use them, and a history of non-usage will inevitably lead to deaccessioning to save on the storage costs.

16 Approaches to the Problem

Clearly, since research libraries continue to collect them for their research value, omitting foreign language print-only journals from the trend towards better article-level discovery is a problem. Many approaches to address this problem have been attempted. For example, working with the professional organizations of area studies librarians such as CONSALD’s (Committee on South Asia Libraries & Documentation) Journals Subcommittee, JSTOR has sought 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 enterprises has thus been growing gradually in recent years. However, the growth has only been in English-language journals (the largest language of publishing in South Asia and many other parts of the world).

17 Approaches from Scholarly Societies

One might anticipate that specialized scholarly societies, representing constituencies with a great stake in discovery and access to these materials, would make the necessary investment to start filling that gap. An example is the Association for Asian Studies (AAS)—an international scholarly professional association of over 7,000 members—to serve the interdisciplinary needs of broad Asian Studies research. Among its activities in support of this field, AAS has a broad publishing agenda, including the premier bibliographic reference tool for Asian Studies—the Bibliography of Asian Studies (BAS). The BAS was produced as an annual print index from 1936 through 1990. With a grant from the Mellon Foundation, the BAS was transitioned to its current incarnation as a cumulative online database with nearly a million records, originally hosted at the University of Michigan (UM). To support production and administrative costs and the salaries of paid indexers, AAS charged institutional and individual subscription rates. Nonetheless, the costs and complexities of its production, and the challenges of updating the database’s infrastructure and capabilities, could not be indefinitely sustained within the original online framework of informal academic institutional agreements at UM. In 2016, AAS decided to transfer the BAS’s infrastructure and hosting to EBSCO.
while retaining in-house editorial control. The BAS is now distributed commercially among EBSCO’s other database offerings. The BAS is a vast index of journal articles and book chapters from and about Asia, and is most noteworthy for its coverage of material from Asia that is not discoverable through other sources. For article discovery, nothing compares with its scope, but, as with JSTOR, the foreign language journals are again omitted: the BAS describes itself as “the single most important record of research and scholarly literature on East, Southeast, and South Asia written in Western languages.” Journals in the languages of Asia are not included.

The American Bibliography of Slavic and East European Studies (ABSEES) is another example of bibliographic production by a scholarly association: it was originally published as an annual printed volume from 1956 through 1994 and later switched to electronic only. It was produced under the auspices of the former American Association for the Advancement of Slavic Studies (AAASS) and supported by grants from various funders. Like the BAS, ABSEES started as a production of a scholarly society, hosted within a variety of academic settings, including the Graduate School of Library and Information Science at the University of Illinois at Urbana-Champaign. The cost of production, development, and hosting this resource became too high for local investment at the library (though most of its more than 135,000 records are from English sources, mostly published in the US and Canada), and the product was therefore converted to a commercially distributed database provided by EBSCO.

What can be seen from the BAS and ABSEES’ trajectories is that the limited resources available to scholarly society publishers dictate that even within the scope of specialized area studies, they must focus their efforts on activities that will benefit the largest potential population of their constituents: academic readers of English and other Western languages. Though not themselves geared for profit in the manner of commercial publishers and aggregators, these scholarly societies make choices about coverage (and eventually about moving hosting to a commercial provider) by the same kinds of logic of return on investment, which will tend to yield the same result for foreign language journals: the latter do not seem to warrant much (if any) investment.

**Approaches from the Library Community**

If commercial and non-commercial providers and scholarly societies are not likely to invest in creating article-level discovery of foreign language print journals, who will? While many will agree that these may be important for research, and that certain kinds of research simply could not take place without them, who has the motivation and resources to prioritize serving this potential constituency by creating a framework to enable their discovery? The answer would seem to point to the research library community itself, where the mission of connecting scholars to the content they need has always entailed making investments in enabling discovery (e.g., through original cataloging) of more obscure, low-use materials in their research collections.

Over the last few decades, libraries have experimented with numerous approaches to the problem with regard to foreign language journals. These have included local projects where individual libraries provided scanned tables of contents for some off-site materials on their websites and retrieval-request interfaces. There are also single institution, open access projects like the South Asian Periodicals Index (SAPI) from University of Wisconsin Madison Libraries. A variety of collaborative ventures have been attempted, including:

- the Latin American Periodicals Tables of Contents (LAPTOC) project hosted at Vanderbilt University;
- the former Southeast Asian Serials Index (SASI) hosted at Australian National University; and
- the Thai Journal Index project hosted at University of Washington.

What these efforts have in common is that, for a universe of foreign language journals, they seek to recreate the same kind of indexing and abstracting functionality towards the high-end of the discovery continuum described above. They create individual records at the article level, deploy metadata fields to the extent that their production model and human resources allow, and present the records with some level of indexing and a search or browse interface for discovery of the articles.

But beyond the infrastructure costs of hosting and maintaining such databases, the work of creating article-by-article metadata is expensive, especially when considering the need for the language expertise and skills to accomplish indexing or cataloging. Even within distributed production models like LAPTOC where the work of contributing records from nearly a thousand journals published in twenty-nine countries is distributed between dozens of participating libraries, the size of the investment has created severe limits on the extent of coverage or on the sustainability of these projects.

Thus, for example, the SAPI project, supported by University of Wisconsin, indexed six South Asian journals, only two of which are non-English. After creating nearly a thousand detailed indexing and abstracting records for these two South Asian language journals in a distinct, tailored open access infrastructure, the project was decommissioned in the summer of 2018. The LAPTOC project, now collaboratively supported by LARRP (the Latin Americanist
Research Resources Project consortium based at CRL, produced a more impressive, searchable, open access database of more than 340,000 article-level records. It unfortunately could not garner resources sufficient to sustain it after fifteen years of coverage (1994 to 2009).

The Thai Journal Indexing Project was organized by the Committee on Research Materials on South East Asia (CORMOSEA) member libraries and others and supported for two years by grants from the Association of Research Libraries (ARL) under the AAU/ARL Global Resources Project. This Thai project deploys a unique combination of hand-crafted indexing and abstracting metadata with page images of the first page of each article to enable snippet-type browsing of the full text. It covers six Thai language journals but seems to have gone dormant following the conclusion of its grant funding in 1998. For now, the resources provided by most of these projects are still available on the web, each hosted in its particular structure and institutional setting or archived in the Internet Archive, but without any new investment in content or updating of the database or its functionalities.

The SASI was an open access index of about 77,000 records for articles from or about South East Asia derived from very selective coverage of about 140 journals. The records were created or contributed by the Australian National University (ANU) Library, Monash University’s Asia Institute, and the Royal Netherlands Institute of Southeast Asian and Caribbean Studies (KITLV). The records consisted of article titles, sometimes with keywords, presented in a searchable database hosted by ANU. The Southeast Asian language content consisted of records created in the 1990s under a now defunct Indonesian Serials Database plus records from a small number of journals indexed by the KITLV.

The challenges faced by all these models include the unsustainably high cost of creating the records and the informal or potentially unstable commitment to maintain the infrastructure for one-off, “niche” products like these on the part of the hosting institutions or consortia. Local stakeholders move on, institutional memories fade, and local IT resources to support or update those dormant or “legacy” projects get deployed for other priorities. These challenges have necessitated trade-offs such as aiming for slender coverage, minimal metadata, or scaling back or terminating projects once the initial funding is gone.

The Hispanic American Periodical Index (HAPI) has taken a different approach. HAPI, originally a print bibliography from the 1970s, is now an online database with more than 335,000 citations (about a fifty-fifty mix of online and print-only content) to articles in more than 700 journals (with about 400 journals being indexed currently, of which about 80 percent have links to full text). The citations are full-fledged bibliographic entries, with subjects, added descriptors, and links to online full text where available (e.g., for the open access journals). The search interface is sophisticated, with many advanced features for searching, sorting, limiting and exporting result sets. The work of creating the records is essentially crowd-sourced among about thirty volunteer librarians and scholars at US universities and a few Latin American countries. The database has been hosted and maintained by the University of California Los Angeles’ Latin American Institute. The trade-off that has enabled this project to remain operational is that it is not an open access production; while institutions in Latin America and the Caribbean receive free subscriptions, other institutions pay for subscriptions at tiered rates tied to student FTE levels. The potential academic audience for Spanish-language journals is undoubtedly much higher than that for Thai, Punjabi, Hindi, Vietnamese, etc., and this cost-recovery model of paid subscriptions to cover hosting costs, combined with crowd-sourced production of metadata, might therefore be sustainable. However, there is some indication that even under this kind of cost-recovery model there is concern regarding HAPI’s long-term sustainability.

In an era when even large research libraries that are collaborating are challenged with securing or retaining the necessary human resources to provide original cataloging of their own backlogs of books and journals, how likely is it to sustainably prioritize the indexing or abstracting of the vast number of individual articles in all those demonstrably low-use, specialized foreign language journals? With notable exceptions like HAPI, such collaborative attempts generally have been spotty, limited in scope, and/or hard to sustain. The cost-benefit analysis for true indexing and abstracting of these inherently low-use materials tends to make such efforts hard to justify.

**SALToC as a Discovery Mechanism: Case Study**

What follows is a case study of a very different, experimental, library-based collaborative approach undertaken by a group of South Asian Studies librarians targeting a select subset of print-only journals in languages of South Asia. The South Asian Languages Cooperative Tables of Contents (SALToC) project was specifically designed to be sustainable by avoiding the pitfalls of the kinds of projects outlined above, by not aiming for true indexing or abstracting, deploying instead a very low-tech, low-investment, distributed online project to enable open access article-level discovery towards the low end of the discovery continuum with no required new back-end or interface programming. SALToC enables discovery by online browsing of tables of contents, which is certainly incrementally better than no
discovery, and provides 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 enabling scholars to locate previously undiscoverable journal holdings. Exposing article citations for discovery in this way makes it possible for scholars to place ILL requests, document delivery requests, and off-site retrieval requests with full citations for the desired articles. While many libraries do not lend print journals, many do offer article-level document delivery on request, but only if the requesting institution provides a full citation. 35

SALToC’s “grass-roots” distributed type of approach offers a proof-of-concept demonstration that the goals of coordinated collection development (jointly planned reduction of unnecessary duplication and enrichment of the collective collection) can be enabled even for specialized, low-use foreign language print journals through cost-effective shared access via a low-lying discovery layer. How it works, the values it provides, and its potential weaknesses follow.

SALToC’s History and Goals

Collecting from South Asia has a long history at many US research libraries. Although efforts at inter-institutional collaboration started earlier, the 1962 advent of the cooperative acquisitions program of the Library of Congress (LC) in India (initially subsidized with rupees from the US PL480 program) enabled a broader range of US libraries both to expand their collection coverage from the region and to begin exploring methods for coordinating their profiles. 37 This was possible because most of these libraries were acquiring books and journals from the LC program via a single profile structure. National efforts at South Asia collection coordination—with varying degrees of success—became more visible after a 1974 Boston Conference on South Asian Library Resources in North America, organized under the South Asia Council of the Association for Asian Studies, which highlighted gaps in the national collection from the region. 38

CONSALD has been operating and expanding continuously since the 1960s. It functions as the professional library organization of South Asian Studies specialist librarians in collection development and technical services roles from all the North American research libraries supporting interdisciplinary South Asian Studies. CONSALD membership currently numbers about fifty. 39 Meetings, projects, and collaborations address a wide range of issues particularly relating to collection development, access, preservation, and retention of materials from South Asia.

In March 2013, in recognition of the particular problems related to coverage of South Asian journals in US library holdings, and the article-level discovery and access issues outlined above, CONSALD created a Journals Subcommittee. To improve full-text access to journals from the region, this subcommittee worked with JSTOR to successfully advocate for inclusion of a wider selection of English-language South Asian journals in the JSTOR database. The Subcommittee also receives regular reports on South Asian coverage in the BAS. Although JSTOR and the BAS provide extremely valuable, broad access to articles on the region, their coverage remains limited to English and other western-language sources. A recent study of Asian Studies scholars’ research practices concluded that while “many Asian studies scholars expressed that they are not particularly challenged in their ability to access information published in the U.S.,” scholars “experience difficulty discovering materials published outside of the West, often having to travel and spend significant amounts of time browsing through libraries, archives, and bookstores to discover information relevant to their research.” 40 The issues of technology shortfall for providing discovery and access to non-Roman-script materials (such as lack of optical character recognition capabilities for these scripts) are also highlighted in this report.

To address some of this discovery shortfall, CONSALD’s Journals Subcommittee began in 2013 to explore the possibility of creating their own non-commercial project to enable browsing of tables of contents of South Asian language journals. A formal proposal was presented at the October 17, 2013 CONSALD meeting. 41 This quickly led to a refined set of goals, operating criteria, and methodologies for the joint project, thereafter dubbed SALToC. SALToC’s main goal that emerged from these deliberations was enabling article-level discovery of vernacular language journals from CONSALD collections identified as not otherwise discoverable because they are not online and not included in existing full-text or bibliographic databases. Key objectives included:

- freezing patrons to identify and access articles in their own collections or request them from other libraries through standard ILL;
- allowing patrons to cite articles in their research and to use the citations to request journal articles from offsite storage;
- making it possible for cooperating libraries to provide digital document delivery (“scan and deliver” or “photocopy and deliver” services) for articles in these vernacular journals, just as they already are providing for print journals in English;
- providing an online substitute for physical browsing; and
- allowing libraries to select runs of journals for removal to offsite storage without sacrificing some level of discovery and bibliographic access.

While many libraries do not lend print journals, many do offer article-level document delivery on request, but only if the requesting institution provides a full citation. 35

SALToC’s “grass-roots” distributed type of approach offers a proof-of-concept demonstration that the goals of coordinated collection development (jointly planned reduction of unnecessary duplication and enrichment of the collective collection) can be enabled even for specialized, low-use foreign language print journals through cost-effective shared access via a low-lying discovery layer. How it works, the values it provides, and its potential weaknesses follow.

SALToC’s History and Goals

Collecting from South Asia has a long history at many US research libraries. Although efforts at inter-institutional collaboration started earlier, the 1962 advent of the cooperative acquisitions program of the Library of Congress (LC) in India (initially subsidized with rupees from the US PL480 program) enabled a broader range of US libraries both to expand their collection coverage from the region and to begin exploring methods for coordinating their profiles. 37 This was possible because most of these libraries were acquiring books and journals from the LC program via a single profile structure. National efforts at South Asia collection coordination—with varying degrees of success—became more visible after a 1974 Boston Conference on South Asian Library Resources in North America, organized under the South Asia Council of the Association for Asian Studies, which highlighted gaps in the national collection from the region. 38

CONSALD has been operating and expanding continuously since the 1960s. It functions as the professional library organization of South Asian Studies specialist librarians in collection development and technical services roles from all the North American research libraries supporting interdisciplinary South Asian Studies. CONSALD membership currently numbers about fifty. 39 Meetings, projects, and collaborations address a wide range of issues particularly relating to collection development, access, preservation, and retention of materials from South Asia.

In March 2013, in recognition of the particular problems related to coverage of South Asian journals in US
To highlight the contrast with other, more expensive discovery systems developed for more mainstream materials, this group clarified what it did not want SALToC to be. This was necessary to control expectations and to keep the project “low to the ground,” to reduce the costs of creating discovery, and thereby improve the return on investment. SALToC is:

- not a journal article indexing project;
- not an indexing and abstracting service;
- not a searchable, structured database of citations; and
- not a table of contents alert service.

Because these South Asian studies librarians lack significant project resources from their libraries at their disposal, they narrowed SALToC’s planning parameters to keep it as low-tech and low-cost as possible while still providing real discovery value.

**SALToC Project Methodology**

In keeping with these principles, SALToC was given a simple workflow requiring only modest investment of human resources from each participating library. The steps are simple, and do not require any special skills or highly trained staff, at each stage of the process. The central infrastructure for accumulating the distributed TOC content and placing it online in a sustainable open access archive was established by the South Asia Librarian at New York University (NYU). SALToC went live in early 2015. The workflow steps follow:

- Student assistants do not need to know the relevant languages and are quickly trained at each participating institution. They make simple, page-image PDF files by scanning the TOCs of each issue of the target journals contributed by their institution. A separate TOC file is created for each issue of the journal.
- The PDF files are annotated with two basic Roman script bibliographic fields transcribed from the journal’s local existing bibliographic record (title and imprint), and a field derived from the local volume holdings (or bound-volume spine labels): volume, issue number, date. The annotation also includes the journal’s OCLC accession number.
- The annotated PDFs are transmitted to the central SALToC repository at NYU where the files are ingested as part of the university’s institutional repository (“Faculty Digital Archive”) maintained on the university’s DSpace platform with a separate permanent URI for each journal’s “landing page,” which also displays links to its accumulating run of TOCs (PDF bitstreams), and acknowledgement of the contributing library. See http://hdl.handle.net/2451/33893 for an example of a typical SALToC journal landing page.
- An access link to the permanent URI for each journal’s landing page is added to the bibliographic record for that journal in the contributing library’s OPAC or integrated discovery layer to enable discovery of the TOCs by local users.
- Under the auspices of the Cooperative Serials Cataloging Program (CONSER), authorized NYU catalogers update the OCLC record for each SALToC journal, adding the same link to the journal’s SALToC landing page (to enable discovery of the TOCs by others). See http://www.worldcat.org/oclc/214908417 for an example of a WorldCat record with the SALToC access link for TOCs.
- A simple DSpace “collection” page for the entire SALToC project, continually updated to list and link all the journals included, is also maintained in NYU’s repository to enable participating librarians to promote and highlight SALToC to their patrons through LibGuides, bibliographic instruction, etc. beyond the linkage for each title provided through OPACs and WorldCat. The SALToC collection page is available at https://archive.nyu.edu/handle/2451/33560. Since the SALToC site is crawled and indexed by Google, the project site and each of its journal titles are also discoverable through general internet searching.

The contributing institutions’ workflow for scanning the issues is thus light, and the brief metadata for the annotations is a matter of transcribing available bibliographic and enumeration/chronology information. Unlike indexes and other searchable databases that require subjects and descriptors, SALToC does not require description, cataloging, or metadata creation because that work has been done in local catalogs and OCLC at the journal-title level. The selection of titles for SALToC is determined by consensus among the CONSALD Journals Subcommittee and the contributing libraries, and no library assumes more of this workflow than its staffing can easily absorb.45

The workflow at NYU’s central file repository is likewise light. After the initial 2015 development of the SALToC template by NYU Libraries Digital Library Technical Services staff within the DSpace repository platform, the ingest of the contributed PDF files is straightforward, with an interface that supports manual ingest by students (who do not need to know the relevant languages) or batch-ingest of files organized by simple file-naming conventions (e.g., Astha_01.01.pdf, Astha_01.02.pdf, etc.). The task of adding the enumeration/chronology labels for the link to each bitstream file involves only copying the enumeration/
chronology field from each PDF file’s internal annotation and pasting it into the bitstream description label. With that infrastructure and standardized workflow in place, adding TOCs for the existing titles is very straightforward. When new titles are added to the SALToC project page, NYU’s SALToC Project Editor edits the DSpace collection page to add a link to the new landing page and informs the CONSER-authorized cataloger to update the OCLC record to include the new URI link to the title’s landing page into the OCLC record’s 856 field.\(^43\)

Beyond local decisions regarding staff capacity for scanning and contributing TOCs, SALToC group considerations for coverage and prioritization include issues of language (for example, focusing on widely collected versus “overlooked” South Asian languages), discovery correlated to existing access and delivery mechanisms, and subject scope as related to coordinated collection development among the group. Two years of detailed discussions have resulted in flexible parameters. Individual participating institutions could contribute journals according to local priorities determined by their South Asia librarian and input from relevant scholars. The general consensus was to focus on annuals, bi-annuals, or quarterly titles with more complete holdings, though monthlies are occasionally included. Many SALToC titles have their runs completed by coordinated contributions from multiple participating libraries. The subject selection of serials for SALToC coverage focuses on history and humanities, literature, and social sciences. As of this writing, SALToC consists of twenty-six titles, plus 1,545 files in eleven languages, supplied by twelve partners.\(^44\) A thirteenth partner, NYU Library, contributes the hosting services and institutional repository infrastructure described above.

**SALToC Within the Context of Cooperative Collection Development and Distributed Print Archiving Commitments**

Beginning in 2010, South Asian Studies librarians, including many CONSALD members, organized a collective referred to as SACOOP (South Asian Cooperation).\(^45\) SACOOP held annual workshops to offer consensus methodologies to address specific aspects of cooperative collection development.\(^46\) The 2010 workshop focused on rationalizing these libraries’ South Asia monograph profiles in their participation in the South Asia Cooperative Acquisitions Program (SACAP) of LC’s New Delhi field office to begin to orchestrate a well-rounded national collection.\(^47\)

In 2011, this group’s attention turned to considering collaborative approaches to South Asian serials. The Fall 2011 SACOOP Workshop focused on highly subscribed and least-subscribed journal titles acquired through LC’s New Delhi-based SACAP with the goal of shifting resources to broaden the collective collection and fill in gaps in the national coverage of the journals using detailed subscription and holdings analyses contained in a 2010 working paper by Wright.\(^48\) Simultaneously, the group recognized that the coordination of collecting required appropriate discovery and access, as reported in its 2011 workshop report:

> It was agreed that institutions should better coordinate serials subscriptions on a national level. That said, if cooperation determines that access is not to be local, proximate and reliable access to the literature is critical. Libraries must lobby for better indexing and discovery tools—or to create them ourselves—so that discovery leading to successful interlibrary loan is actually feasible…. Participants expressed strong interest in extending the availability and discovery of indexing for top journals to allow institutions to more comfortably relinquish physical access to local copies.\(^49\)

These deliberations eventually led to the concerted effort to create at a minimum browsing discoverability for articles in the least accessible South Asian journals—those in South Asian languages produced only in print. From this endeavor, SALToC was born.

With the gradual coordinated shifts of subscription resources for these journals, and with the successful launch of SALToC to begin to provide the discovery necessary for ILL and physical access, in 2006, the SACOOP group heeded the need for explicit print-retention commitments for assured future access to materials collected under cooperative agreements.\(^50\) A working group was created, charged with exploring methods and terms for print retention agreements, and for an envisioned SACOOP distributed print archive in particular. With the goal of identifying and prioritizing the South Asian materials to be targeted for print retention commitments, a SACOOP Print Retention Content Group was also created. This working group’s ideas and general recommendations were discussed and endorsed at the fall 2017 SACOOP Workshop.\(^51\)

The purpose of commitments to print retention include simply ensuring continued access to the targeted material for the holding institution and providing some level of assurance to other libraries that the material will continue to exist in the holding library. The other libraries may have a “stake” in the target material, either in terms of: (1) an expanded universe of content to which the patrons of non-holding libraries could have theoretical access or (2) another holding library’s ability to deaccession their holdings of the target material (e.g., to save money or space) without completely losing access to the targeted content.
This is explained in the Content Working Group’s core principles.\textsuperscript{52}

Priorities for retention commitment could be keyed to judgment of the relative size of the “stake” that the collective community of libraries and scholars has in that material. That is, the larger the stake in a given set of material (for SACOOP or the whole community of libraries and scholars), the more important it will be to target that material for retention commitment. Of the two stake elements above, (1) should be given a higher weight because all will have a stake in material that uniquely enriches the total universe of valuable research content in libraries, and its value depends upon its committed retention. But stake element (2), while theoretically enabling some set of specific holding libraries to save money and space by deaccessioning their copies of the target material (on the basis of the committed retention by one library), will produce a much narrower benefit (only for those libraries, not for the universe of scholars from all libraries). Element (2) should be given a lower weight than (1) in targeting materials for retention. Everyone benefits from a library committing its holdings of important research materials, but when a library withdraws duplicated copies it only benefits itself.

Another set of considerations that impact the evaluation of the stake of a given set of material are the conditions of access. For example, the community as whole has a much smaller stake in the retention commitments of material retained in a library that will not lend it than the same material in a library that circulates it. The general principle is that the terms of the stake vary according to the degree of access (including discovery) provided by the holding library.\textsuperscript{53}

With these principles in mind, consensus emerged in 2017 among the SACOOP members with the idea of interlinking the creation of article-level discovery of South Asian language periodicals through SALToC with the targeting of specific titles for retention commitment. Implementation of explicit institutional commitments along these lines is not at the sole discretion of the South Asian Studies librarians themselves, and so it progresses incrementally according to the general frameworks and infrastructures for retention commitment available at each library. Commenting on the need for libraries to commit to new ways of sharing, Rader wrote, “Access to content is critical for success in all research areas. As we respond to changes in the North American research library environment—reduced budgets, for sure, but also increased opportunities to rely upon each other through deep collaboration—we are called to build and deploy new inter-institutional structures to ensure the ongoing discovery, access, and use of materials. The interconnected relationships of SACOOP and SALToC epitomize what is possible when we work together for a common good.”\textsuperscript{54}

Conclusion

Nothing about SALToC is radical or entirely new. Considered in the aggregate, SALToC model’s features represent a more intentional effort to locate an appropriate value point between the ideal and the real. Recognizing that article-level discovery is a matter of degree, and that for specialized, low-use research materials like these, discovery mechanisms at the high-end of the continuum (the ideal) would require high levels of investment that would skew the cost-benefit ratio and make the whole enterprise unsustainable in institutional contexts of constrained resources, SALToC aimed instead for something much more modest, and of incremental discovery value, that could be sustained and scaled because it is unlikely to be cut.

How sustainable is SALToC? The SALToC model avoids the pitfall of relying upon purpose-built or one-off structures with potentially impermanent locations on the web. It was designed to ensure that the catalogers’ one-time-per-journal investment of work to add TOC links into OPAC and WorldCat records would not need to be updated or later revised.\textsuperscript{55} Catalog links to open web content are ephemeral and tend to become outdated and quickly become dead-end links.\textsuperscript{56} SALToC’s approach reassured catalogers that they were linking to permanent content at a permanent address. SALToC achieves overall sustainability in several ways:

- it uses a light-weight, low-tech, maintainable infrastructure;
- it requires minimal resource investment of human workload and system resources; and
- it provides demonstrable discovery value for researchers, through targeted browsing of tables of contents.

The collaborating South Asian Studies librarians who devised SALToC 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. Housing SALToC in NYU’s DSpace institutional repository required no specialized programming, workflow, database, or server maintenance. It uses an existing system and maintained to do what the University is already committed to doing: providing a repository built for faculty, with long-term institutional commitment for permanence and permanent URIs. SALToC thereby leverages the value of that existing infrastructure without additional cost.

How scalable is the SALToC model? What are the potential limits on its growth? SALToC is scalable because contributor institutions add as much or as little as they want. As of this writing, SALToC grows at an average rate of about 500 TOC files per year. The barriers to entry are
exceedingly low, and so far, three years into the experiment, 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). Each contribution (TOC) 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. With this infrastructure in place, the work to insert each successive contribution is negligible: it is completed in minutes by non-specialist staff and students.

SALToC therefore 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 and the University of Wisconsin’s South Asian Periodicals Index) require actual data-entry for each article at participating institutions. This dramatically increases the cost of production, making the project less sustainable or scalable. Learning from projects such as LAPTOC and SAPI, SALToC participants chose to meet the needs of discovery through browsing: page images, with no data-entry and no language skills required. Expanding the channels of discovery by embedding title-level SALToC TOC links into the individual full catalog records in WorldCat provides the added value of enabling users to find wanted articles and to simultaneously see which libraries have the relevant holdings.

For low-use material like these South Asian language print journals, it is still too soon to conduct a full-scale evaluation and accounting of fully loaded costs and derived values, costs-per-use, research impact factors, user experience, etc. Some have compared SALToC to other resources with which they are familiar (for example JSTOR). It clearly seems home-grown and improvised, it lacks features available with high-end productions, and it lacks a search box. However, SALToC is also generating enthusiastic reports of use-cases on how it is enabling scholars to delve into the journal content these libraries are collecting for them across the cooperating institutions in ways that seem to validate the premise of coordinated collection development.

References and Notes


10. James Simon, Vice President for Collections and Services at the Center for Research Libraries (CRL) in Chicago, notes that “The supply chain of scholarly resources has changed radically in the past twenty years, yet there are vast tranches of information not yet integrated into the global information environment. A recent survey of Africana journals at Michigan State University suggests nearly 80 percent of Africa-based journal titles are inaccessible or, at best, partially indexed in bibliographic tools. This paucity of discovery is even more acute for journals published in vernacular script that defy easy encoding.” (personal communication, July 6, 2018). Likewise, Ramzi Nasser and Kamal Abouchedid, in their overview of journals in the Arab world, report that there is “a lack of systematic archiving as well as indexing and bibliographies that leaves many researchers in the dark about work of others.” “Problems and the Epistemology of Electronic Publishing in the Arab World: The Case of Lebanon,” First Monday 6, no. 9 (2001), accessed July 8, 2018, http://firstmonday.org/ojs/index.php/fm/article/view/886/795.


15. “The library community’s increasing reliance on remote storage has fragmented the one remaining recourse for discovery: browsing the shelves. Libraries have been slow to exploit the possibilities of digital technologies to integrate discovery of content held five, fifty, or five hundred miles away from the patron,” James Simon, personal communication, July 6, 2018. The access problems specifically associated with off-site storage of serials—especially unindexed serials—were well anticipated by Dan C. Hazen in 2000: “Moving long runs of unindexed serials can also be particularly grave in terms of diminished user access. . . . Serials, for instance, are at once attractive and problematic candidates for storage. Moving a serial can save lots of space, but without complete and ready bibliographic access via indexes or citation databases, effective intellectual access is almost impossible. . . . We can by now represent detailed serial holdings in our online records. With appropriate initial processing, users should thus be able to verify a library’s precise holdings of an off-site serial. Knowing what is inside these volumes, however, can be far more difficult. Printed indexes are an obvious resource, and many serials regularly produce their own cumulative indexes. External indexing services may also cover a specific journal, though it is important to confirm both time frame and completeness. Where indexing does not exist, or even in addition to indexing, digital technologies may assist in creating information on the contents for inclusion in an online catalog. One approach is to scan page images of tables of contents for users to consult online. . . . Creating searchable text files of tables of contents, which could support queries based on author name, keywords, and the like, might be a (more expensive) next step.” Daniel C. Hazen, “Selecting for Storage: Local Problems, Local Responses, and an Emerging Common Challenge,” Library Resources & Technical Services 44, no. 4 (2000): 176–83. See also David Block, “Remote Storage in Research Libraries: A Microhistory,” Library Resources & Technical Services 44, no. 4 (2000): 184–89.


23. University Library, University of Illinois at Urbana-Champaign,
An interesting potential counterexample is the Byara database, a searchable compendium of nearly 13,000 bibliographic records of articles in Tibetan language journals. Latse Library (New York), Byara Database, accessed July 11, 2018, http://www.latse.org/byara. In Byara, most titles are translated into English, and keywords have been added to enhance the search capacity of the database. Many of the abstracts have also been translated into English (see the description at http://www.latse.org/catalogs). Unlike the other article discovery endeavors described above, the Byara Database is the product of a private library (the Latse Library) that is part of a supporting foundation (the Trace Foundation) that is devoted to “the continuity, development, and vitality of Tibetan communities.” The Trace Foundation, About Trace, accessed July 11, 2018, http://www.trace.org/about. The long-term sustainability of the Byara Database under this private foundation sponsorship is unknown.

According to Peter Bae, Assistant University Librarian for Scholarly Content Services, Princeton University Library, personal communication, June 29, 2018; Reference and User Services Association of the American Library Association (RUSA), Interlibrary Loan Code for the United States, approved by RUSA Board January 11, 2016, accessed July 14, 2018, http://www.ala.org/rusa/guidelines/interlibrary. Section 4.3 of the Code indicates that the requesting library is required to “describe completely and accurately the requested material following accepted bibliographic practice.”


At NYU, the ongoing workload consists only of adding each PDF file to the DSpace repository as it is received from the contributors (less than five minutes). The one-time work of setting up each new journal title in SALToC consists only of 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 ten minutes, and updating the master record for the journal in WorldCat to add a link back to the SALToC landing page for the title.

43. At NYU, the ongoing workload consists only of adding each PDF file to the DSpace repository as it is received from the contributors (less than five minutes). The one-time work of setting up each new journal title in SALToC consists only of 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 ten minutes, and updating the master record for the journal in WorldCat to add a link back to the SALToC landing page for the title.

44. Center for Research Libraries, Columbia University, Cornell University, Harvard University, Library of Congress Field Office Islamabad, Princeton University, University of California Berkeley, University of Chicago, University of Pennsylvania, University of Texas, University of Washington and Yale University.


53. The effectiveness of formal print retention commitments in empowering cooperative collection development among library communities, as well as in local decision-making, critically depends upon registering and disclosing those commitments to all concerned, and even making those disclosures available for macro analysis. This is an explicit principle and requirement in the HathiTrust Shared Print Program for monographs. See, for example, the Disclosure section and the discussion of the HathiTrust Shared Print Registry in HathiTrust Shared Print Program Operating Policies and Guidelines, June 2017 [revised May 2018], accessed July 8, 2018, http://www.hathitrust.org/sites/www .hathitrust.org/files/HathiTrust%20Shared%20Print%20 Policies%202018%2005%20rev.pdf. It is also central to the June 2018 announcement of the Mellon Foundation’s grant to OCLC and the Center for Research Libraries “to enhance the underlying infrastructure of the OCLC WorldCat database and CRL’s Print Archives Preservation Registry (PAPR) to accommodate and make accessible actionable data for shared print serials management.” OCLC Awarded Mellon Foundation Grant to Register Library Retention Commitments for Print Serials in WorldCat, accessed July 8, 2018, http://www.oclc.org/en/news/releases/2018/20180624-oclc -awarded-mellon-foundation-grant.html. While SACOOP discloses its serial retention plans to its own members in its workshop and reports outcomes documentation, it will probably have to work with each member institution to ensure that any formal, MoU-based commitments do get centrally registered in PAPR to be of maximal value to non-SACOOP members as well.

54. Mary Rader, personal communication, July 6, 2018.

55. This operation takes less than five minutes per title, according to Joyce Bell (Cataloging and Metadata Services Director, Princeton University Library), personal communication, November 2015.

2003 and found that they had a half-life of only two years (i.e., on average 50 percent of webpages disappear within two years). Faith Oguz and Wallace Koehler, “URL Decay at Year 20: A Research Note,” *Journal of the Association for Information Science & Technology*, 67, no. 2 (2016): 477–79. With the reduction in barriers to self-publishing on the web, and the ubiquity of internet access world-wide, the average half-life of web content has presumably grown even shorter since then.