

Libraries, Archives, and Museums Interwoven in the Digital South: From a Paucity of Resources to Data Richness for Understanding the South Asian Subcontinent



James Nye

Keywords Libraries · Archives · Museums · Metadata · Intellectual property rights · Digital cultural heritage coalitions

The cultural heritage of South Asia¹ is one of the wonders of the world. It deserves to be more fully understood, better preserved, and easily accessible.

This paper explores the current scarcity of digital resources for understanding the rich heritage of South Asia and some of the reasons why the situation exists. Second, the history of preservation and access programs focused on library, archive and museum resources of the South Asian subcontinent is considered along with recently inaugurated initiatives to improve access for scholars and for citizens. A concluding section identifies factors which would result in a richness of data resources for South Asia comparable to those currently available for other world areas.

1 A Dirth of Resources

The South Asian subcontinent might be succinctly described as located in the digital South. The phrase “digital South” was coined by the author to describe the paucity of resources and digital formats for understanding the cultural heritage of what some

¹South Asia is comprised of the sovereign nations Afghanistan, Bangladesh, Bhutan, India, the Maldives, Nepal, Pakistan, and Sri Lanka. The region is home to more than one fifth of the world’s population.

A version of this paper was originally presented on August 30, 2017 at DCH2017, a digital humanities conference held at Staatsbibliothek Berlin.

J. Nye (✉)

The University of Chicago Library, 1100 E. 57th Street, Chicago, IL 60637, USA
e-mail: jnye@uchicago.edu

call the Third World and others refer to as the global South. Comparisons with the enormous abundance of resources for understanding the cultural heritages of Europe, North America and East Asia reveal painfully stark disparities. There are numerous ironies, oddities, and obstructions that have given rise to this comparative dearth of resources from South Asia. Some of these are maladies dating to or are the consequence of British colonial domination of the region. But many, arguably most, should not be registered as colonial legacies. For example, in India we have the enormous irony of a country justly priding itself in educating some of the most adept practitioners in information technology and yet very few of those people address the needs of the region's cultural heritage. When we look at new initiatives such as the one announced last year by the Government of India for Massive Open Online Courses (MOOCs) it is shocking to find that only four or five out of more than 150 courses listed by the government address anything to do with the humanities and social sciences. Turning to Pakistan for another example, a country with Urdu as the national language, one finds the nation in the odd position of holding a modest percentage of Urdu language resources in the subcontinent. Attempts to improve access to the Urdu literary heritage require complicated interactions with the Government of India and with Indian institutions because the overwhelming majority of non-digital print assets in the Urdu language reside across the national border in India. Sri Lanka provides another example. Far too frequently one encounters the misplaced fear in the island nation that if resources from the country are digitized and made available for public access, no one will visit the archive collections and use the archival resources in that nation state. This is a pervading myth—a myth encountered virtually everywhere in the subcontinent—that runs absolutely counter to the needs of academics, both in country and worldwide, who are trying to understand the cultural richness of South Asia. A major museum in Nepal, severely damaged by the 2015 earthquakes, presents a deeply unfortunate example of a preservation irony. High-resolution images of all damaged objects in the museum along with detailed architectural drawings of the museum, part of a palace complex, would require terabytes of space for proper archiving. The funding agencies supporting restoration have not provided for repository space. Consequently, the images and drawings will almost certainly be lost for use by subsequent generations of scholars. The Maldives provides a final example of an anomaly. Beyond the digital South, one might reasonably define the Maldives as the *deep* digital South, a designation shared with Bangladesh. The Maldives was subordinated as a protectorate by the British under Ceylon (now named Sri Lanka) from the late nineteenth century. As a consequence of that complicated relationship with Ceylon, many of the most important assets, tangible and intangible assets for understanding the culture of that small but important island nation, were archived and held in Ceylon rather than in the country itself.

Further complicating the situation in the digital South and specifically in South Asia, the administrators and staff responsible for cultural institutions are often viewed with far less respect than they deserve. It is often said that librarians and archivists in the subcontinent are viewed as little more than peons. A consequence of this is that very few of the brightest people are attracted into the field of library science

and archives management. This may partly explain why so little of the information technology energy is going towards libraries. If one takes the case of archives, it is common joke, but a substantial truth, that to be appointed as the head of an archive in South Asia is considered a punishment posting. If one has misbehaved, if one has done something to attract the ire of the political elite, one will end up appointed to what is seen as the lowest ranking opportunity in civil service, that is, to head an archive. Finally, South Asian museum have, until very recently, lacked adequate support for educational programs and found themselves well out of sync with best practices museology.

Further compounding the noted complications, highly variable climate combined with often inadequate housing facilities have resulted in gigantic loss over the centuries. Manuscripts fall prey to insect and rodent damage or the vicissitudes of monsoon climate. Books and periodicals, in addition to the maladies faced by manuscripts, were frequently printed on paper with high acid content and consequently, many cultural resources have disappeared without a trace. Museums have been underfunded, especially following mid-twentieth century independence from colonial rule. In sum, the material and human infrastructure in South Asian libraries, archives and museums is severely wanting.

2 Previous and Current Programs

On a more positive note, libraries, archives, and museums have benefited over nearly a century and a half from legislation and programs designed to register and distribute publications from the South Asian subcontinent. The 1867 Press and Registration of Books Act was enacted to regulate printing presses and newspapers printed in India. It provided for the registration of those publications and deposit of copies with archives in India. Additionally, copies of the publications could be selected by five collections in England. A consequence of that act is that a significant percentage of pre-Independence publications were documented in quarterly lists of publications produced and some of the publications were added to collections in the United Kingdom.² Provision for transmission of printed publications to the five collections in the United Kingdom terminated, for the most part, with independence in 1947 and 1948 for the former colonies in the subcontinent.

Beginning in the late 1950s, the Library of Congress began an important program for collecting and disseminating to selected U.S. academic libraries important publications through cooperative programs run from offices currently based in New

²While compliance with the 1867 act were uneven across time and administrative units, the best estimates are that no more than 85% of imprints were registered and deposited. Further, practices at the British Museum and the India Office Library, two of the libraries in England entitled to request copies strongly favored publications in English. In the early twentieth century, less than 15% of books in the regional languages of colonial India were requested in many years and, most of the newspapers and other periodicals in regional languages were acquired as specimen issues rather than as continuous runs.

Delhi and Islamabad.³ The South Asia Cooperative Acquisitions Program (SACAP), as it is known, has provided to academic institutions in the United States copies of resources that are the envy of many of our colleagues in South Asia itself. The collection of publications delivered to the University of Chicago through the program, for example, consists of more than 700,000 volumes published since 1960 in the regional languages of the region with an estimated value of USD 34 million. A distinctive and important feature of SACAP is that where national boundaries have divided sub-regions with a shared language—Urdu for example, between India and Pakistan—SACAP has allowed U.S. institutions to build collections in a way that has supported scholarly needs and will eventually support collaborative digitization, to the benefit of our colleagues in the subcontinent.

The South Asia Materials Project (SAMP, known until 2014 as the South Asia Microform Project)⁴ based in the United States, is another example of an important collaborative program for preservation and access to the cultural heritage of South Asia. Over more than a half century, SAMP has prepared high quality microfilm surrogates of resources necessary for understanding South Asia as a region and has made those resources available to researchers at member institutions. The program, recently renamed the South Asia Materials Project, has used member contributions in excess of USD 1.1 million over the past 50 years for the preparation of resources that address the needs of researchers in the humanities, social sciences, and the sciences. Since SAMP microfilms contain high-resolution images, they are a vital asset for future digitization that will benefit our colleagues in the subcontinent. One specific SAMP initiative is noteworthy both because of the methods deployed in the execution and the resulting resources preserved. The Microfilming of Indian Publications Project⁵ was undertaken with colleagues in South Asia between 1990 and 2001. Using the *National Bibliography of Indian Literature, 1901–1953* as a source of well-selected historical publications, SAMP managed to create high-quality preservation microfilm copies of approximately 38,000 volumes. Those volumes are assets that are currently being digitized for open access.

Over the past twenty years several experiments have been undertaken to acquire, preserve, and disseminate private collections in India, Pakistan, and Nepal through collaborative engagements between North American libraries and bodies in the subcontinent. Specifically, institutional and external funds from North America have been raised to purchase and process private collections; new bodies or existing institutions in South Asia have provided homes for the collections; and as the collections have been preserved, cataloged, and made accessible, ownership of those collections is being ceded by the North American institutions or consortia to the South Asian

³The Library of Congress websites for the New Delhi and Islamabad Field Offices are respectively at <https://www.loc.gov/acq/ovop/delhi/> and <https://www.loc.gov/acq/ovop/islamabad/>.

⁴SAMP, celebrating its 50th anniversary in 2017, maintains a website at <https://www.crl.edu/programs/samp>. It is noteworthy that the change of SAMP's name, from microform to materials, signaled a change in emphasis from preservation microforms to digital copies as a preferred form of conservation.

⁵Please see the Microfilming of Indian Publications Project website at <http://dsal.uchicago.edu/bibliographic/nbil/nbil.php>.

institutions. Those South Asian libraries and archives have provided an excellent base for ongoing collaborations. They have been led and staffed by some of the finest librarians and archivists in the region and brought imaginative new approaches to the preservation and dissemination of resources, most of which are extremely rare and endangered. For example, the Roja Muthiah Research Library in Chennai, the Madan Puraskar Pustakalaya in Kathmandu, the Mushfiq Khwaja Library and Research Centre in Karachi, and the Centre for Studies in Social Sciences, Calcutta have all received multiple grants from the Endangered Archives Programme⁶ for preservation of collections at risk.

The British Library's project titled Two Centuries of Indian Print⁷ is, in its first phase, digitizing more than 1,000 early printed Bengali books published between 1713 and 1914. Work is being undertaken with the School of Cultural Texts and Records at Jadavpur University and several other Indian institutions. If additional funding is obtained, the project may be extended to include materials in other South Asian languages and expand the participating institutions to include those specializing in other regional languages.

A newly formed organization called the South Asia Open Archives (SAOA)⁸ is building upon antecedent programs to remedy some of the circumstances that have resulted in the digital South and to facilitate better understanding of South Asia's cultural riches. A founding group of nearly thirty institutions in South Asia, North America, and Europe joined in 2016 to establish and maintain, for non-commercial open access, a collection of materials for the study of South Asia and to federate with other collections of resources. All academic disciplines are encompassed, from the humanities through the sciences. Activities addressing major objectives during the first five years are funded by the USD 750,000 so far contributed by SAOA members and through external funds from foundations and other sources. Those objectives include:

- mass digitization of textual, visual, and sonic resources for open access,
- creation of large sets of numerical data through conversion of printed sources such as the decennial censuses of the subcontinent,
- resolving questions of intellectual property rights in South Asia through work with the responsible government bodies,
- training library and archive staff for engagement following international standards,
- collaboration for improvement of existing metadata and implementation of linked open data for resource description, and

⁶The Endangered Archives Programme, based at the British Library, is described on the website at <https://eap.bl.uk/>. Specific projects for the South Asian institutions mentioned are also described on that site.

⁷A project website is located at <https://www.bl.uk/projects/two-centuries-of-indian-print>.

⁸Information on the South Asia Open Archive is available at <https://www.crl.edu/samp-open-archives-initiative>. The charter for SAOA is found at <https://www.crl.edu/sites/default/files/attachments/pages/SAMP%20Open%20Archives%20Plan%202016-01.pdf>. Please note Appendix I of that document enumerating digitization priorities during the first five years.

- provision of interfaces for exploring SAOA resources and interlinking them with those of other trusted institutions—for example, interfaces for geo-navigation and analysis of large data sets, in support of collaborative, cross-disciplinary work.

SAOA has drawn inspiration and guidance from other projects, especially those associated with the digital South. The guiding principles of the International Dunhuang Project have resonated particularly well with SAOA. Those principles include: (a) free access to resources provided to all users, (b) comprehensive coverage of the resources distributed across repositories around the world, (c) requirement that images and data created be of very high quality, (d) assurance that data gathered is as rich as possible, (e) collaborative engagement by participating repositories, (f) maintenance of autonomy of partner institutions, and (g) provision for the long-term sustainability of the program and access to the scholarly resources. The Endangered Archives Programme's (EAP) commitment to avoid duplication of effort with other non-profit programs is another bedrock principle for SAOA. Rather than duplicating effort, SAOA seeks to complement and complete resources for presentation to scholars and citizens. Additionally, EAP's focus upon resources at risk has captured SAOA's attention as fundamentally important. Those risks may be from environmental, socio-political, or other factors that would imperil culturally significant resources.

3 Future Prospects and Recommendations

Moving forward, building upon earlier and current programs and projects, there are promising ways in which the digital South can be further addressed and the current dearth of digital resources can be replaced by abundance, by a richness of resources comparable to that currently available for many other world areas.

The DCH2017 conference on Digital Cultural Heritage held in Berlin offered a remarkable opportunity for learning, sharing approaches to the technical challenges faced by participants, and thinking strategically about what the participants might accomplish through collaboration. The papers included in this collection of essays address many of issues and approaches presented at the conference. One of the most important themes presented in papers and explored in discussions during the conference is development of partnerships across disciplines and across geographical sectors. This theme was also prominent in the Google summit on culture held in Florence in mid-February 2017.⁹

⁹The author participated in the two-day Google summit titled Digital Meets Culture. Participants from twenty-three countries included leaders of cultural institutions, academics, representatives of UNESCO, the European Commission, and national governments. Sessions at the summit centered on the preservation and accessibility of cultural heritage in the digital age. The summit produced "Recommendation for the future of digital and culture" focused "on the importance of digital technologies in fostering accessibility and audience engagement, in preserving and promoting cultural diversity and in enhancing the digital footprint of cultural institutions through digital skills." Please see the recommendations at https://www.lib.uchicago.edu/e/su/southasia/Recommendation-Digital_meets_Culture_2017.pdf.

Those involved in digitization of cultural resources related to South Asia will want to give special attention to a few considerations as they develop new programs for improved access. These suggestions are based on lessons learned through prior engagements by the author and his colleagues while working on preservation and access projects.

1. Documentation of library, archive, and museum holdings is fundamentally important. Because very few resources are currently described with appropriate metadata, the hard work of cataloging physical objects is necessary to support decisions about priorities for digitization activities. In other words, surveys of public and private collections in South Asia followed by cataloging of unique resources is an essential prerequisite for informed digitization programs.
2. Collaborations between actors in the subcontinent and those elsewhere in the world improve the likelihood of success in preservation, digitization, and presentation projects. The development of collaborative linkages takes time, it requires diplomatic skills, and the understandings informing collaborations must be renewed and refreshed on a regular basis if partnerships are to produce the expected results.
3. Funding is necessary and yet often difficult to secure for projects. Funds for sustaining digital resources are even more difficult to obtain and yet without corpus funding the future of the digitized cultural resources will not be secure.
4. Human infrastructure is essential. Problems noted earlier in this paper regarding perceptions of librarians, archivists, and museum staff often make it difficult to locate well-trained and committed staff in South Asia. Programs for training in best practices are often necessary before projects can be undertaken.
5. Private repositories rather than public institutions in South Asia are often the best targets for preservation and digitization. Private collections are often superior to those in institutions. Further, bureaucracy associated with public collections can be a major impediment in accomplishing preservation and access objectives.
6. Copyright and intellectual property regimens must be honored scrupulously. This is important for both legal and ethical reasons. Having declared that principle, it is important to work with government officials responsible for intellectual property in sovereign nations of South Asia to secure the broadest interpretation of what copyright law does and does not permit.

Recognizing the impediments present when working with cultural resources related to South Asia, building upon previous and current programs for digitization and access, and with attention to the cautionary notes included in this paper, one can reasonably expect that new projects will enhance the understanding, improve the preservation, and ease access to the rich cultural heritage of South Asia.

James Nye James Nye has served as Bibliographer for Southern Asia at the University of Chicago Library since 1984. He is responsible for collections on South Asia and services to readers. Nearly one million volumes comprise the University's collection about South Asia. All languages of the subcontinent are encompassed. His academic training is in Sanskrit and Sanskrit literature.