

SAMP Digitization Projects Guidelines

South Asia Materials Project

SAMP supports the digital reformatting of material that adheres to its mission and scope (preservation of scholarly material from or relating to South Asia, not readily available to North American scholars). Traditionally, SAMP has concentrated on preservation of newspapers, journals, monographs, and archival materials, but to address emerging research trends will consider projects that seek to digitally preserve important video and audio materials, as well as material born digitally.

SAMP will support reasonable costs related to organization, description, preparatory conservation work, scanning, and presentation of materials. SAMP encourages the use of existing digitization capacity of partner institutions, but can consider the purchase of equipment and provision of digitization training under certain circumstances. It is broadly consistent with the goals of SAMP to include capacity building components in collaborations with South Asian memory institutions, and SAMP will consider funding for pilot digitization projects as a way of exploring possibilities for broader partnerships/resource sharing arrangements. Wherever possible, SAMP would also like to include a microfilming component within collaborative digitization projects.

Proposals to SAMP should adhere to appropriate current best practices in terms of selection, conversion, metadata, and access. Projects should be well-planned and where possible benefit from lessons learned in digital project manuals or comparable existing projects. If copyright permission is necessary, explain the nature of the partnership between the participating institutions. See list of resources for digitization standards and digitization project management attached at the end of this document.

The SAMP project proposal should include an explanation of any deviations from the principles listed below:

- The original project material is in a format that be easily digitized.
- The proposed digitization process will capture the full intellectual content of the source materials.
- The material is either in the public domain, permission to digitize has been secured from the copyright holder, or the project falls under ARL's Fair Use Guidelines.
- The proposed technical specifications will reflect the recommendations described in the standards resources cited in this document.

- Metadata will be created in accordance with best practices, so that it will provide access to specific images or text files, and may be interoperable with the OAI-PMH (Open Archives Initiative Protocol for Metadata Harvesting) standards in order to allow for harvesting of the data whenever possible.

The digitized materials and metadata will be deposited at CRL, which will be granted a non-exclusive license to ensure the long-term retention and maintenance of the digital files, as well as rights to display these files on its own servers or through partner institutions. CRL will neither engage in nor permit for-profit uses of these files.

Any successfully funded proposal will be subject to additional terms and conditions in the Center for Research Libraries Terms and Conditions of Award document.

Proposal Guidelines

Please include the following sections in your proposal.

I. Narrative

Title & Abstract

Please provide a brief description of your proposal. This should include information about the resources to be digitized, their provenance, scholarly value, and reasons for digitization. Please include basic bibliographic information, such as publisher, date of publication, author, etc. and whether the contents are unique to your institution or held by other institutions, including North American holdings. Describe other preservation/digitization efforts related to this material.

Content

Further describe the content proposed for digitization.

- What kinds of materials will be digitized (e.g. print, photographs, audio, video, realia, etc.)? Provide data on numbers of pages, images, etc.
- What are the physical characteristics of the original material (size, color, volume)?
- What is the condition of the material? Is the content brittle or otherwise endangered?
- What provisions are in place for conservation work in anticipation of digital reformatting, or for the potential loss of content during the digitization process?

Copyright / Permissions

Describe any issues regarding the intellectual property of the content.

- Is the material in the public domain?
- If not, who owns the copyright?
- Have you secured permission from the copyright holder to make these freely accessible? Please provide documentation of appropriate authorization by the copyright holder.

- Will you assign rights to SAMP to store, access, and reproduce the material for other not-for-profit purposes, in accordance with the Terms and Conditions?

Conversion Procedure

Describe the means by which you propose to digitize the material.

- Describe the facilities that will be used for digitization, and the appropriateness of the equipment to the format and condition of the material.
- What specifications do you propose using to convert the material? Include details such as resolution, file format, file sizes, etc.
- What quality control will be used to ensure best practices are adhered to throughout the conversion process?
- If OCR is generated, will it be edited or uncorrected?
- Will your project include a microfilming component?

Metadata

Describe method of creating and providing access to metadata for the digitized files.

- What metadata standard will be used?
- At which level will the metadata be applied?
- Who will prepare the metadata?
- Will MARC bibliographic records be created?
- Will the metadata be harvestable using OAI-PMH standards?

Added-Value Features

Describe any proposed products beyond digital image files. For instance:

- Will text files be made searchable via the application of Optical Character Recognition software or double-keying?
- Will searchable text files be marked up in accordance with specific schema?

- Will numerical files be rendered in forms suitable for statistical manipulation?
- Will cartographic and related materials include geospatial referencing?

Access

Describe how the users will access the data.

- What delivery system do you propose to use?
- In what format will the files be delivered?
- Will the data be freely available on the internet? If not, what limitations to access will be in place for this data (and why)?
- What search and browse capabilities will be used to access the data?
- Will the metadata allow for easy harvesting of data?

Archiving

Describe terms for the preservation and ongoing maintenance of content.

- What is your process for sustained preservation of the files?
- Will the data be archived at any location(s) other than CRL?
- How will you deliver the files to CRL?
- What will you do with the original source material?

II Plan of Work

A detailed workplan should include an estimated schedule for the digitization project, broken down by the phases of the project (selection, permissions, preparation, conversion and quality control, metadata creation, delivery, preservation, etc). The workplan should also include information about the staffing needed to complete all aspects of the project.

III. Budget

A detailed budget should include estimated costs for the digitization project, broken down by the phases of the project. The budget should include any project support requested of SAMP, as well as expected from sources other than SAMP.

Digitization Standards Resources

California Digital Library. *CDL Guidelines for Digital Images*. v. 2.0 2005.

<http://www.cdlib.org/inside/diglib/guidelines/bpgimages/>

CDL Guidelines for Digital Objects. v. 2.0 2007. [includes metadata standards]

<http://www.cdlib.org/inside/diglib/guidelines/>

CDP Metadata Working Group. [was Western States Group] *Dublin Core Metadata Best Practices*. v. 2.1.1 2006 <http://www.bcr.org/cdp/best/dublin-core-bp.pdf>

Copyright Term and the Public Domain in the United States. Peter Hirtle, Cornell Copyright Information Center. 2008

http://www.copyright.cornell.edu/training/Hirtle_Public_Domain.htm

Digital Library Federation. *Benchmark for Faithful Digital Reproductions of Monographs and Serials*. 2002. <http://www.diglib.org/standards/bmarkfin.htm>

Western States Digital Imaging Best Practices v. 1.0 2003

<http://www.bcr.org/cdp/best/digital-imaging-bp.pdf>

Digitization Project Management Resources

BCR CDP Digital Toolbox. <http://www.bcr.org/cdp/digitaltb/index.html> See especially *Getting Started with Digital*

Projects http://www.bcr.org/cdp/digitaltb/getting_started/faq.html

Institute of Museum and Library Services. *NLG Project Planning: a Tutorial* http://www.imls.gov/project_planning/

NISO Framework Advisory Group, National Information Standards Organization. *A Framework of Guidance for Building Good Digital Collections*. 2nd edition. 2004.

<http://www.niso.org/framework/framework2.html>

North Carolina ECHO (Exploring Cultural Heritage Online). *Guidelines for Digitization*. 2007. <http://www.ncecho.org/guide/toc.html>

Northeast Document Conservation Center. *Handbook for Digital Projects*, edited by Maxine K. Sitts. 2000. <http://www.nedcc.org/oldnedccsite/digital/dighome.htm>

Association of Research Libraries, Center for Social Media, and Program on Information Justice and Intellectual Property. *Code of Best Practices in Fair Use for Academic and Research Libraries*. 2012. <http://www.arl.org/bm~doc/code-of-best-practices-fair-use.pdf>

rev. 03/13