

Staging Plan for the Digitization of Federal Depository Library Collections held by CIC Libraries

October 2009

Background

The library directors of the Committee on Institutional Cooperation (CIC) charged this steering committee with responsibility for developing and overseeing a comprehensive collection management strategy for U.S. government publications across the CIC libraries. The first phase of this project is to digitize a significant proportion of CIC holdings of U.S. Federal Documents, and make them publicly accessible through the Hathi Trust Digital Repository.

Using their high-quality, sheet-fed page scanning operations, Google will partner with the CIC to digitize these collections. In the pilot phase of this project, the University of Minnesota delivered nearly 85,000 volumes of duplicate federal government publications to Google.

The staff at the Center for Library Initiatives has identified Penn State as the next library in the queue to send federal government publications to Google. This plan concerns the staging of library shipments to Google once Penn State's initial commitment has been fulfilled and will serve as a guide for subsequent libraries.

Scope

The scope of materials to be included in this project is limited to print publications distributed to CIC libraries through the Federal Depository Library Program (FDLP). While developing a truly comprehensive digital collection of all federal government publications is a laudatory goal, the bibliographic tools needed to identify this comprehensive collection do not currently exist. However, we can determine with some accuracy the number of print publications there were distributed through the FDLP. In addition, we are not recommending the digitization of microform distributions. Moreover, depository distributions represent the bulk of the federal government publications held by CIC libraries. Reducing the duplication of these materials within the CIC will afford significant space savings.

From a project management perspective, we think that the narrower scope of this project is justified, but this should not prevent libraries from contributing non-depository materials, especially since local practices may make distinguishing between depository and non-depository holdings impractical.

Guiding Principles

- In compliance with Title 44 of the *United States Code*, only publications held by selective federal depository libraries, or duplicate copies of publications held by regional federal depository libraries are eligible for this project.
- Libraries must comply with the disposal guidelines of the appropriate regional federal depository library for all materials targeted for digitization.

- The project implementation is designed to minimize the overall costs associated with preparing materials for digitization, so libraries that already have large portions of their collections cataloged and/or bar-coded may be asked to contribute a larger percentage of their materials to the project.
- In selecting materials for digitization, libraries will adhere to all format, size and condition standards adopted Google, or any other digitization partners.
- Because this project has been endorsed by all CIC library directors, all CIC libraries will participate, either by contributing actual items or by resourcing the project in other ways, such as creating or purchasing cataloging records.

Staging

This plan is staged so as to maximize the volume of materials digitized early on, and to minimize the amount of local processing required of each participating library, while still providing for a comprehensive digital collection of materials distributed to the CIC libraries through the FDLF. There are three stages in the plan, as well as a process for materials in need of special handling.

In the first stage, selected libraries will contribute large collections (approximately 60,000 or more volumes) of cataloged federal publications that meet the format, size and condition standards adopted by Google. Priority in the queue has been given to libraries that have significant holdings of bar-coded material. With the assistance of Google, each library participating in this stage will develop a pick-list of unique titles to be digitized. Libraries should plan for the digitization of as many of titles as possible. This stage will proceed sequentially through the queue until completed.

The second stage applies to institutions contributing smaller collections of cataloged federal publications that meet the format, size and condition standards adopted by Google. With the assistance of Google, each library participating in this stage will develop a pick-list of unique titles to be digitized. This stage may run concurrently with the first stage, with materials being shipped to Google for digitization as intake capacity becomes available. A library's place in the queue will depend upon available capacity.

The third stage is reserved for federal publications held by CIC libraries that meet the format, size and condition standards adopted by Google, but that are not cataloged by any CIC libraries. In this stage each library within the CIC will take responsibility for the publications of a group of individual federal agencies, and using standard bibliographic tools will identify, locate and catalog any remaining materials. This stage may run concurrently with the first two stages, with materials being shipped to Google for digitization once cataloging is completed, and as intake

capacity becomes available. A library's place in the queue will depend upon available capacity.

For classes of materials that do not meet the format, size and condition standards defined by Google (e.g. oversize materials, posters, maps, items with inserts) the steering committee will identify both donor libraries and possible digitization partners. Agreements with potential digitization partners are subject to approval by the CIC directors. Once digitized, these publications will be deposited directly with the Hathi Trust.

Provisional sequencing

Based upon a preliminary survey, the steering committee has provisionally ranked participating libraries into three tiers, with Tier 1 generally representing those libraries with the largest number of ready-to-go (cataloged and bar-coded) materials and Tier 3 the least. Libraries are listed in alphabetical order, not in the order they might occupy in the digitization queue.

Tier 1) Illinois-Urbana Champaign, Indiana, Michigan State

Tier 2) Michigan, Northwestern, Purdue, Wisconsin

Tier 3) Chicago, Illinois-Chicago, Iowa, Ohio State

In the coming weeks, the steering committee will refine this ranking and begin to identify libraries for inclusion in each of the stages outlined above.

Ongoing project management

Once candidate libraries have been queued for participation in this project, the steering committee will be responsible for tracking commitments and serving as a resource group. Concurrently, an implementation team should be formed to develop routines and standards for the quality control of digitized material and to track contributions and completed scans.