National Bibliography of U.S. Government Publications

Contents:

I. BRIEFING TOPIC: NATIONAL BIBLIOGRAPHY OF U.S. GOVERNMENT PUBLICATIONS

I.1 SETTING THE STAGE

GPO is developing the National Bibliography of U.S. Government Publications (National Bibliography) that fully meets the statutory requirements of 44 U.S.C. 1710-1711. By law GPO is charged with preparing and publishing a "comprehensive index of public documents," including "every document issued or published … not confidential in character."

GPO's National Bibliography is a comprehensive catalog containing descriptions and locations of U.S. Government unclassified publications in all formats. The National Bibliography describes any publication, regardless of form or format that any U.S. Government agency publishes, disseminates, or makes available to the public, as well as any publication produced for administrative or operational purposes, that is of public interest or educational value. Publications represented in the National Bibliography are acquired from official sources or sites, and are not subject to official use or security classification restrictions.

Initially the National Bibliography database will consist of GPO-produced bibliographic records representing titles cataloged from 1976 to the present. Once the ILS initial implementation is completed in spring 2005, GPO plans to expand the National Bibliography database by adding pre-1976 records, records created by cooperative cataloging partners and other institutions, as well as records converted from other bibliographic or metadata formats. Building the National Bibliography legacy database is a project that will span multiple fiscal years, while adding records for new publications is an ongoing effort.

Functionally, the National Bibliography will be a principal output of the Integrated Library System (ILS), Ex Libris' Aleph 500. GPO will develop one or more relational databases of library cataloging and other metadata records that describe and link to U.S. Government publications in a variety of formats and locations. The National Bibliography will utilize metadata resources created by GPO, by other agencies, and by partner institutions. The ILS' federated search function will provide the capability to search across a distributed set of metadata resources from other institutions and display a combined result set. The public face of the ILS will be GPO's Online Public Access Catalog.

Planning for the National Bibliography emphasized ways to improve the comprehensiveness of GPO's cataloging and metadata programs, to enhance the usability of GPO's bibliographic products and services, to operate a standards-based national library quality program, and to enhance the public's ability to identify and locate U.S. Government publications.

I.2 BIBLIOGRAPHIC CONTROL OF THE NATIONAL COLLECTION

Bibliographic access to all publications in GPO's portion of the National Collection will be provided through GPO's National Bibliography and potentially by other metadata services. Cataloging records for online publications will include a persistent link to the publication. GPO's goal is that digital publications will be accompanied by preservation metadata describing their content, file type, provenance, etc.

Bibliographic control will be provided to the individual publication level for all access copies of publications in the National Collection. Applying metadata at this level will enhance the performance of federated search tools and OpenURL linking technologies. GPO bibliographic records will conform to the practices and standards established for the National Bibliography. Digital publications intended for print-on-demand reproduction and sales will also have book industry standard metadata. The metadata for digital publications should indicate the permitted access to that item if any restrictions apply. Other or additional metadata systems or elements may be applied to other portions of the National Collection.

I.3 NEW INFORMATION

To improve communication about GPO's cataloging practices and developments, ID has created a National Bibliography News page on GPO Access at
http://www.access.gpo.gov/su_docs/fdlp/cip/index.html.

Beginning February 1, 2005, the OCLC holdings symbols of 30 regional depository libraries have been added to the bibliographic records GPO contributes to OCLC WorldCat. Working with GPO, OCLC has developed an automated batch loading process to set the holdings, which is run daily. This project will bring improved public and library community awareness of resources available in Federal depository library collections through an increased number of bibliographic access points and facilitated interlibrary loan of tangible U.S. Government publications. Adding OCLC symbols to GPO-produced cataloging originated as a suggestion from regional depository librarians. Most recently, the Oklahoma Department of Libraries and the University of Maine joined the project, bringing the total to 32 regionals participating.

At the Midwinter meeting of the American Library Association, GPO announced that the Office of Bibliographic Services would begin creating a unique bibliographic record for each manifestation, or version, of a title, at the same time that the ILS goes live. This issue was first raised in the *National Bibliography of U.S. Government Publications: Initial Planning Statement*, published in draft June 18, 2004, and was also discussed at the October 2004 meeting of the Depository Library Council. This issue engendered a large number of comments, the majority of which favored creating a unique record for each publication or version. Reasons supporting GPO's intent to create unique records include:

- More precise searching by format in the OPAC.
- Bibliographic record accuracy for libraries that receive only one format of a Government publication.
- Conformance with emerging national and international cataloging standards.
- Enables more precise version control and identification in content management systems such as GPO's Future Digital System.

II.   REVISED ASSUMPTIONS

II.1   The Integrated Library System assists GPO in building the National Bibliography. A 21st century National Bibliography cannot be operated on GPO's legacy information technology (IT) platforms.

II. 2   GPO will continue to contribute records to OCLC.

II.3   Initially, GPO's core National Bibliography database format will be AACR2 cataloging records in MARC21 format. As cataloging standards and metadata schema evolve, GPO's use of metadata schema and cataloging formats will also evolve.

II.4   Other bibliographic formats may be ILS/National Bibliography system inputs or outputs. Possible system input or output formats include, but are not limited to:

1. MARC21
2. ONIX
3. GILS
4. Preservation metadata
5. Dublin Core
6. COSATI

II.5   When GPO performs original cataloging, the default level of cataloging applied to U.S. Government publications, regardless of format, will be equivalent to OCLC "Full" level. Following applicable standards, GPO will apply different levels of cataloging in some cases, such as when converting records for older materials without the described publication in hand; e.g., records derived from the GPO shelf list.

II.6   National Bibliography cataloging records will come from multiple sources, including:

a) GPO original cataloging
b) Record imports
c) Conversion of pre-1976 GPO bibliographic records
d) COSATI-MARC conversions
e) Harvested metadata
f) Metadata from content originators
g) Records contributed by depository libraries and other partners
h) Preservation metadata from legacy digitization projects
i) Conversion of the GPO historical shelf list

II.7   The National Bibliography database will provide public access and location information for all of GPO's holdings in the National Collection. Bibliographic access to NC items held by partners may be provided through federated searching.

II.8   GPO cataloging policies and procedures will be documented and communicated proactively.

II.9   GPO will continue its participation in the LC Program for Cooperative Cataloging (PCC) Programs. These include:

        a) NACO (name authorities)
        b) SACO (subject authorities)
        c) CONSER (serial cataloging)

II.10   GPO will continue to assign Superintendent of Documents classification numbers to publications made available to Federal depository libraries regardless of format.

II.11   No fees will be charged to search the National Bibliography or to link to publicly available data described therein.

II.12   No fees will be charged to depository libraries to obtain bibliographic records from the National Bibliography.

II.13   Records obtained from the National Bibliography database are free from copyright or redistribution restrictions.

II.14   Location information in the National Bibliography includes persistent links to online resources and a mechanism to identify the depository libraries that hold tangible publications; e.g. the "Locate Libraries" function.

II.15   GPO will identify and participate with other metadata development and standards-setting organizations.

III.   QUESTIONS TO COUNCIL, WITH COUNCIL DISCUSSION

III.1   QUESTION: Are the revised assumptions on the National Bibliography still appropriate, or should something be changed or added?

DISCUSSION BY COUNCIL:
Council generally agreed with Assumptions 1-4. Following the Council discussion, a reference to converting records from the GPO shelf list was added to Assumption 5. Council discussion regarding OCLC possibly asserting a copyright restriction over GPO's contributed bibliographic records led to a new Assumption 13.

This led to a lengthy discussion of Assumption 7, and how the National Bibliography will search or retrieve content from different parts of the National Collection. Assumption 7 was amended to refer to GPO's holdings, and the possibilities of searching content held by others using federated searching.

Assumption 12 was amended to state that libraries can acquire GPO bibliographic records from the ILS free of charge, regardless of whether push or pull technology is employed.

Assumption 14 was amended by adding the "Locate Libraries" function as a clarifying example.

III.2   QUESTION: Now that the plan for the National Bibliography has been revised in light of discussions with Council and other public comments, are there elements of the plan that need further refinement or need to be added?

DISCUSSION BY COUNCIL:
Council encouraged GPO to investigate pushing National Bibliography and cataloging information out, in addition to the National Bibliography News web page on *GPO Access*. GPO is investigating using RSS technology for this and similar applications.

Council asked GPO to clarify the intent to apply the Superintendent of Documents classification number to both digital and tangible content. GPO agreed, and pointed out that the abbreviated class number used briefly for digital content has been replaced with the full application of the classification system. Further discussion ensued about Library of Congress and Dewey classification numbers. GPO described a pilot initiative with the Library of Congress on assigning LC classification for Congressional materials. GPO has received few requests to assign Dewey class numbers, although that has been under discussion with LC.

IV.   QUESTIONS FROM COUNCIL ADDRESSED AT THE MEETING

Council questioned GPO concerning a discussion at the April 18 regional meeting concerning ways to help identify and control the inventories of Government information around the country and in regional and selective libraries. One of the ways that came up was the proposal by Paul Gherman of SOLINET about working with OCLC to post holdings to OCLC of primarily things in storage, in a controlled storage environment, like a light archive. GPO has briefly reviewed the Gherman document and, without in-depth analysis, believes the plan has some possibilities for light archive management.

The question of harvesting Government publications from the Web was raised, with some concern that the volume of the harvested content could result in backlogs in cataloging and other processes. Building on this was a discussion of the necessity to rely on traditional cataloging or metadata-based retrieval in the emerging information technology environment. GPO responded that in the Future Digital System GPO has developed a requirement to harvest any associated metadata when we harvest content. That would provide GPO with metadata that was created by the content originator, and then GPO would be able to enhance it. Council also stated that today it is possible to make documents searchable and retrievable without any metadata at all, and if GPO should get to the point at which cataloging is up to capacity, the recommendation of that Council member is to keep on getting materials that are retrievable and searchable, even if they cannot be cataloged as completely as called for here.

V.   AUDIENCE QUESTIONS ADDRESSED AT THE MEETING

The facilitator of the Council sessions accepted questions from the audience written on GPO-supplied cards. Nine of the nine questions were answered during the Council session. Those questions and their answers are summarized below.

V.1   QUESTION: Given that the FDLP is moving to electronic, will GPO look at using a metadata standard other than MARC21, one that is more appropriate to electronic formats? Will the metadata be linked directly to the publication so it's easily available to anyone using it, especially important in versioning?

RESPONSE: That is what GPO is planning in the Future Digital System metadata requirements analysis. There is really no single scheme that meets all of the anticipated metadata requirements, so we are looking at a combined metadata scheme for the Future Digital System, which will incorporate a multitude of elements, will be linked to the content, and will facilitate the unambiguous identification of versions.

V.2 **QUESTION:** Will GPO continue to post its monthly cataloging output to the FBB? Will all record levels be included? What about records created by partners?

RESPONSE: GPO plans to make its cataloging records available from the Integrated Library System. GPO will be migrating away from using the Federal Bulletin Board as the dissemination tool for distributing records, but the records will still be available. This will include all the record levels we create. We intend to distribute records originally created by partners, subject of course to any potential copyright issues that may turn up in our various arrangements. GPO's intent and desire is to make as many records as possible freely available, absent any unforeseen restrictions.

V.3 **QUESTION:** Isn't all of the National Collection GPO's?

RESPONSE: The National Collection is a distributed collection, and partners may hold different pieces of it at different locations. Conceptually, we consider depository library collections to be part of the National Collection.

V.4 **QUESTION:** Many pre-1976 records are in OCLC. Can GPO upgrade those records and then own those records for the National Bibliography?

RESPONSE: It has been our assumption that we could, but we're going to validate that and make sure.

V.5 **QUESTION:** Will you classify tangible items that are acquired for the National Collection but not distributed to depositories?

RESPONSE: When GPO finds a fugitive document and digitizes it as part of the legacy collection of things we are digitizing, GPO will apply SuDocs classification to the digital version, as long as the publication in question is in scope for distribution to depository libraries.

V.6 **QUESTION:** Restated the individual's desire for classification numbers even for publications not distributed, to support occasions when libraries choose to make local copies of digital publications.

RESPONSE: That is why GPO changed the practice on the classification of the digital versions to the full SuDocs class, as the abbreviated classification didn't lend itself to printing out a copy and integrating it into a physical collection.

V.7 **QUESTION:** Thank you for providing a no-fee environment for downloading records. Will batching of records be possible?

RESPONSE: It is GPO's intent to provide batching capability, because nobody wants to download those records one by one.

V.8 **QUESTION:** The concept of identifying depository library holdings might be started as a pilot project to have regional libraries report holdings for documents they have and intend to keep,

preferably, and accomplished through technology. Holdings should also be set for digital copies created by libraries with the library's approval. Any comment on that?

RESPONSE: GPO will investigate what is involved in doing this. Our experience with the OCLC project suggests that starting with regional libraries for any kind of a holding project makes sense, because there's not as much de-accessioning from the collections to keep track of.

FOLLOWUP DISCUSSION: The ensuing dialog emphasized the desirability of applying holdings information to bibliographic records for various versions of Government publications. These might include printed copies that libraries had acquired from sources other than the FDLP, for copies printed locally from online resources, or digital files converted from print materials.

V.9 **QUESTION:** This a statement supporting the continued use of the Superintendent of Documents classification system, and the utility of those numbers in the digital environment. For example, the institution's catalog users were browsing the catalog by call number for not just tangible but also electronic versions.

RESPONSE: GPO affirmed the value of the comment.