Cataloging U.S. Depository Materials: A Reevaluation

Alice Harrison Bahr

Automation is reshaping the rhetoric and content of an old debate: to catalog or not to catalog federal publications. The availability of quality records online, keyword search capabilities of online catalogs, and other developments argue persuasively against patent acceptance of old premises. This article reviews previous arguments, discusses conditions redefining them, and proposes partial guidelines for reevaluating current U.S. depository cataloging policies and practices.

Cataloging United States government depository materials is an old debate. Beginning in the 1930s and continuing into the 40s, it was sparked by an increase in the number of depository materials. From 400 in 1900, for example, the number rose to 4,300 by 1930. Typically, numbers dictated organization. Libraries that selected small numbers cataloged them, a practice endorsed for the smaller library by Mary Hartwell, cataloger for the Office of the Superintendent of Documents. Those that received larger numbers maintained separate collections arranged alphabetically by agency, by type of material, or by Superintendent of Documents (SuDocs) classification number, thus organizing materials by agency and series rather than by subject. The latter system was particularly easy, since materials were sent to depository libraries with shipping lists supplying SuDocs numbers. Other libraries used both approaches. They cataloged some materials and placed others in separate collections.

In time, arguments for the two basic arrangements became set. As Waldo points out, their bases "were merely unsupported opinions and assumptions." Separatists pointed to the shortcomings of the card catalog: its inability to index serial publications, its paucity of subject headings, and its difficulty of use, especially for the user confronting the unwieldy U.S. author drawer. There were shortcomings with cataloging itself. It was not suitable for all materials, particularly pamphlets and posters. It imposed a classification system, either Dewey or Library of Congress, that precluded arrangement by agency, useful to some researchers. It increased delays between receipt and availability of materials. It was expensive. Quality records were sparse, necessitating original cataloging, and frequent title and agency name changes required multiple record handling. Relying on higher-quality indexes, with better quality indexing, provided better access to materials and information. Promoting staff familiarity with government materials, separate collections improved the quality of reference service.

On the other side, advocates of cataloging acknowledged its expense, but justified cost on the basis of improved, simplified access. Cataloging spared users the aggravation of consulting several indexes and the annoyance of learning a second classification system. It also provided immediate feedback on library holdings. In

Alice Harrison Bahr is project librarian for online systems at the libraries at Cedar Crest and Muhlenberg Colleges, Allentown, Pennsylvania 18104.
short, it was expensive and time-consuming, but resulted in better service. The old arguments are changing. Technology is bypassing them and raising new questions. The availability of Government Printing Office (GPO) cataloging on networks, record retrieval on some networks by SuDocs number, and online catalogs and their related use studies both beg the assumptions of the past and highlight new needs. Among these are the need to keep abreast of technological developments, to reevaluate present policy in light of those developments, and to formulate policies based on both demonstrated user needs and technological developments.

TECHNOLOGICAL CONSIDERATIONS

Technology is reshaping old arguments by creating access other than through the card catalog, by streamlining cataloging procedures, and by creating a new form of the card catalog. Bypassing traditional card catalog access altogether, for example, Bruce Morton made use of available software and Carleton College’s DEC VAX-11 to produce a SuDocs shelflist accessible not only by SuDocs number but by title, keyword, and Boolean operations. This is an exception, however; most libraries have relied on online vendors, national agencies, and bibliographic utilities to create new options.

The basis for most new possibilities is cooperation among the Government Printing Office (GPO), Library of Congress and Online Computer Library Center (OCLC). Deciding to speed up production of its Monthly Catalog, the official catalog of U.S. government congressional and departmental publications, GPO began cataloging on the OCLC shared cataloging network in 1976. To do so, it had to abandon its local thesaurus and conform to Anglo-American cataloging rules and Library of Congress subject headings. Conformance ushered in GPO’s new role as cataloging authority for U.S. government materials, which became official on January 2, 1981. GPO’s use of OCLC produced the major benefit to libraries: availability of a substantial number of timely, high-quality records.

The type of records available has expanded, and reported hit rates are excellent. In October 1981 GPO assumed cataloging responsibility for materials distributed by the National Audiovisual Center. In 1982 the University of Washington reported finding 75 percent of materials on OCLC. Its hit rate for U.S. government materials was 95 percent. A 1985 Muhlenberg College study reported a 96 percent hit rate for selected categories of U.S. materials (see table 1).

Backlog at GPO is not significant. As soon as GPO catalogs materials, records are available on OCLC. In 1984 a one-month study of “priority one” cataloging items (congressional materials) indicated that 72 percent were input in 15 days and average input time was 24.4 days. The SUNY/Potsdam Library (State University College at Potsdam, New York), which catalogs U.S. materials on OCLC, begins searching one to two months after receipt of documents.

Cooperation between GPO and bibliographic utilities has increased the number of quality records available to libraries relatively soon after publication. It has also created new means of access to those records and new cataloging services. In 1980, OCLC surveyed libraries to determine their interest in an even quicker cataloging process, DARP (automatic distribution of cataloging). Targeted initially for regional depositories (those libraries receiving all available depository materials), the service would have created individual catalog records for libraries receiving depository series. Cataloging would have been automatic, completed as soon as GPO cataloged an item in a series but due to the cost of the service, DARP never got off the ground.

There is a difference between having records available and having quick, convenient access to them. For cataloging purposes, access by SuDocs number is desirable. It allows catalogers to work directly from shipping lists to search for records. OCLC made SuDocs number access (OCLC’s gn: government number key) available in 1980. In 1985 it became possible to search for records by any SuDocs number attached to a title, as of the point of cataloging. SuDocs numbers change when an issuing agency changes, which
TABLE 1
OCLC Record Availability

<table>
<thead>
<tr>
<th>Shipping List Dates</th>
<th>Cataloging Candidates (#)</th>
<th>Available Series Cataloging (#)</th>
<th>Hearings</th>
<th>Geological Survey Professional Papers (#)</th>
<th>Geological Survey Bulletins (#)</th>
<th>Non-GPO Record (#)</th>
<th>No OCLC Record (#)</th>
<th>OCLC Record Availability (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>P840601-1 to P840831-5*</td>
<td>97</td>
<td>4</td>
<td>27</td>
<td>12</td>
<td>6</td>
<td>15 (14/DLC)†</td>
<td>4</td>
<td>96%</td>
</tr>
<tr>
<td>M840601-1 to P840831-8‡</td>
<td>109</td>
<td>10</td>
<td>37</td>
<td>12</td>
<td>6</td>
<td>12 (15/DLC)‡</td>
<td>3</td>
<td>96%</td>
</tr>
<tr>
<td>Total</td>
<td>189</td>
<td>9</td>
<td>61</td>
<td>18</td>
<td>14</td>
<td>45 (23/DLC)</td>
<td>7</td>
<td>96%</td>
</tr>
</tbody>
</table>

*P = paper shipping lists.
†DLC = the number of non-GPO records that are Library of Congress records.
‡M = microfiche shipping lists.

means one title may have several SuDocs numbers. Both searching capabilities enhance access to records.

Another benefit of GPO’s use of OCLC is the generation of machine-readable records on tape. Used to produce Monthly Catalog, the tapes have other uses. They are loaded regularly onto other shared cataloging networks like the Research Libraries Information Network (RLIN) and Western Library Network (WLN) to provide users of those systems with access to the same GPO records as OCLC users. Availability is not as immediate as in OCLC, since GPO tapes are loaded monthly, and these networks do not yet provide SuDocs number access.⁴

Tapes have another potential use for libraries using or planning to implement an online catalog. They provide an alternative means of including GPO records in a local online catalog. One possibility is to download records from a shared cataloging network by means of an interface. The other is to load GPO tapes into the online catalog. This requires tape manipulation both to load the tapes and to eliminate records neither received nor cataloged, i.e. maps, serials, etc.

Cooperation between national agencies and shared cataloging networks weakens the argument that quality records are unavailable, a main pillar in the argument against cataloging U.S. materials. Similarly, the existence of the online catalog erodes another premise, that excessive record handling contributes to the high cost of cataloging. In manual catalogs, record changes require pulling, editing, and refiling cards, usually seven to a set. An online catalog eliminates these steps. While this is a benefit regardless of what materials are being cataloged, it is a particularly strong one for depository materials, which undergo frequent main entry and title changes.

Those frequent changes are inherent in the nature of both serials and government organizations. While little can be done to eliminate the former, technology can help simplify the latter. Some online catalogs incorporate authority control for personal authors, and a few offer it for corporate authors and series. If the catalog also offers online global change, one change to an authority record will automatically alter the same way all affected headings throughout the database.

A common argument against cataloging U.S. materials is the paucity and inade-
quacy of subject headings. Online catalogs have the capacity to increase the traditional author, title, and subject heading access points. To search the manual catalog successfully, users must know precise authors and titles, guess the correct subject term, or know how to use *LC Subject Headings* volumes. Most online catalogs are more forgiving. They provide keyword access to some or all fields as well as SuDocs number access. The latter is increasingly useful as more and more sources include these numbers. Public Affairs Information Service announced recently that 65 percent of indexed U.S. government materials include SuDocs numbers.

By simplifying record changes, eliminating filing, and increasing access points, online catalogs can facilitate a variation of the old either/or proposition to catalog U.S. materials or to house them separately. These catalogs make it easy to combine approaches, namely, to catalog materials, classify them by SuDocs, and yet keep them separate. The same arrangement is possible in a manual catalog, but filing alone argues against it. The library at SUNY/Potsdam, which has cataloged on OCLC and classified by SuDocs since the 70s, reduces filing by cataloging selectively.

Technology, of course, is not a panacea. An automated SuDocs shelflist has the advantage of creating more access points than a traditional card catalog, but it also forces users to consult more than one catalog to locate information.

The availability of records on shared cataloging networks has been an important technological advance. However, not every library uses one, and not every network offers SuDocs access. For libraries using networks, whether or not SuDocs access is available, there are other possible constraints. GPO is slow to catalog certain materials. Possible downscaling of cataloging might also present problems. There has been no report about GPO's 1982 discussions on the possibility of downscaling AACR2 cataloging to augmented level II and no longer providing corporate and personal name authority work.

Specific GPO practices may be of greater concern. For example, GPO provides no collective series entries for titles in monographic series, only individual analytics. This means, for example, that if a library staff decide that a few subject headings will suffice to lead users to the content of U.S. Geological Survey Professional Papers series, they must, if using GPO copy, provide individual analytics for titles in those series. The Depository Library Council, an advisory body comprised of librarians, has repeatedly passed resolutions requesting series cataloging, but official GPO response has been that current practice conforms to depository law. This problem is not insurmountable, however; it requires only that a library create a few collective series entries.

Technological changes are catalysts. They provide new possibilities and challenges, not perfect answers—microfiche records provide an example. GPO catalogs only paper copies of publications that are converted to microfiche for depository distribution. A note indicates that distribution to some depositories was on microfiche, and the word *microfiche* appears after the item number. However, libraries with substantial fiche collections might not welcome the editing required. Exchanges between the Depository Library Council and GPO have thus far produced only the following resolution: GPO's policy is in accord with Title 44 of the U.S. Code, and librarians are free to modify records. Actually, a number of libraries do so, creating suitable records in most cases. This is especially true for government hearings.

Shipping lists from GPO are not always correct, creating another cause for concern depending on the delay between receipt of lists and actual cataloging. Revisions to shipping lists, called "corrections lists," are issued periodically to clear up incorrect SuDocs numbers, typographical errors, and wrong item numbers (the numbers used to order depository series). GPO tape users face greater problems. For example, tapes include all GPO cataloging, and unless a library has received all these materials (some of which are non-depository, that is, not offered to depository libraries), records must be eliminated either in-house or by a vendor. Regardless of who processes the tapes, records are
usually extracted by item number. Since it is not uncommon for depository libraries to fail to receive some items ordered, records may be created for titles not in fact held. The result is like cataloging on the basis of what is ordered instead of what is received. In addition, tape loading on an online system requires manipulation of data.

Not every library has, or plans to have, an online catalog. Even for those who do, GPO tapes present additional difficulties: they come without documentation. Control numbers like OCLC, SuDocs, technical reports, and contract numbers are either omitted or appear in incorrect fields. Correction tapes are not issued. Consequently, errors must first be detected and then corrected manually. There are multiple records for errata slips and periodical issues, and not all series, subjects, and names conform to LC form.

Fortunately, many of the GPO tape errors and limitations have been corrected. The Depository Library Council has been requesting since spring 1979 that GPO convert traced names, series, and subjects to LC form; add control numbers such as item, SuDocs, and stock numbers (used for direct purchase of materials) to appropriate fields; and correct typographical and other errors. The Public Printer announced at the fall 1984 Council meeting that the project would be undertaken and supervised by Judy Myers at the University of Houston. Funds were never allocated. One year later Brodart announced its GPO file, which includes LC subject and name headings and control number corrections. Correction pages from annual Monthly Catalog volumes are being keyed in. While Brodart is still investigating procedures to handle separate serial records, its GPO file offers dramatic evidence of how quickly technology changes situations.

COST CONSIDERATIONS

The availability of quality records in various forms, a major obstacle to cataloging, is part of a larger consideration: cost. Even if quality records are available, are they affordable? The assumption has been no, despite the absence of comparative cost data for processing, maintaining, and supervising separate collections. Fortunately, new possibilities have been explored with an eye to reducing expenses. Available cost data highlight the weakness of previous assumptions and also undercut them.

The cost of Carleton College's automated SuDocs shelflist was determined by counting the bytes per record and then calculating the required disk space for a given number of records. At 310 bytes per record, the storage cost for 431,600 records was $8,500, the same price as a 260,000-block disk and drive. Disregarding input time and use and storage costs, all of which were absorbed by the Computer Center, the unit cost for the shelflist was $1.98, slightly higher than OCLC first-time use charges.

In 1984 Bowerman and Cady, suspecting libraries might include records in their online catalogs if they were available at "a small fraction of the cost incurred in a traditional cataloging environment," compared costs for various methods of obtaining records. The researchers took a sample from a GPO test tape to develop a cost-effectiveness model on which to base a comparison of four ways of obtaining machine-readable records. Two of the four ways involved tapes: the first manipulated in-house, the second by Marcive, a commercial firm that strips records from tapes for libraries. The other two options involved automated cataloging services, either online through OCLC's network or offline, inputting records on a microcomputer and sending them to Library Systems Services, Inc. (LSSI) for tape or card production.

The results were as follows: a library cataloging approximately twelve thousand publications (about 35 percent of available items exclusive of maps, serials, and microfiche) would spend about $1,500, $1,800, $17,000, and $2,600 to obtain records from GPO, Marcive, OCLC, and LSSI, respectively. Record extraction by Brodart, unavailable at that time, reduces costs further: $600 for twelve thousand records ($0.05 per record), excluding tape and profiling charges.

While the study notes the need for data processing personnel for in-house manipulation and raises appropriate questions
about quality control in MARCIVE and LSSI processing, it confines costs to the acquisition, not the storage, of records. Regardless, it dispels the idea that under any and all circumstances cataloging U.S. materials is prohibitively expensive.

USER CONSIDERATIONS

Under some circumstances, then, cataloging of GPO materials is affordable. But is it desirable from the users' perspective? Early studies indicate catalogs are under used and misused. Recent studies reveal little enthusiasm for the enhanced search capabilities of online catalogs. One use study indicates certain faculty locate government publications outside of the library. All studies indicate that much more needs to be known about the use of libraries, government materials, and catalogs before a decision is made about cataloging U.S. depository materials.

According to previous surveys, only 59 percent of library patrons use the catalog. Most catalog users look up one entry and stop. They locate correct subject headings only 50 percent of the time. In 1958 ALA surveyed 5,494 catalog users in thirty-nine libraries. Results showed a failure rate of 20 percent for known-item searches and 13 percent for subject searches.

The capabilities of the online catalog are not as important to user satisfaction as suspected. The Council on Library Resources (CLR) supported a 1982 study of twelve thousand online catalog users in twenty-nine libraries. Eighty-five percent reported finding some or all of what they were searching for, and subject searching was of greater interest than the ALA study concluded. Online catalog capabilities such as keyword access and search qualifiers by date, language, and Boolean operators, however, were not perceived universally as benefits. The CLR study found that language and call-number search limits had a slightly negative effect on satisfaction and that keyword and Boolean operators were unrelated to satisfaction.

Some users bypass the catalog altogether. In his 1984 study of government publications use, Peter Hernon indicated that "academic social scientists rely upon their subject literature and interpersonal sources (e.g., colleagues) for awareness of source material. They do not make extensive use of indexing and abstracting services and bibliographies housed in libraries." Frequently, they obtain materials from outside the library.

Rather than making definitive statements, these studies point out how little is known and how much remains to be seen. After examining two hundred catalog use studies in the most extensive book on the subject, Redesign of Catalogs and Indexes for Improved Subject Access, author Pauline A. Cochrane concludes, "we come away from catalog use studies quite discouraged about the present state of use of our catalogs, with no sure guidance about improvements and the impact of changes."

No evidence supports the assumption that cataloging is always desirable. No substantive evidence supports the opposite view. Technological changes make that clear. They encourage investigation of present situations, their corresponding possibilities, their costs, and their effectiveness in meeting user needs independent of old assumptions. For single answers, they substitute questions.

PARTIAL GUIDELINES FOR REEVALUATING CATALOGING POLICIES

The old question of whether or not to catalog U.S. depository materials raises no single query. It was made into one by technologies that limited options, elevated and fixed cost considerations, and made assumptions about user needs. Newer technologies suggest several questions: they focus on current and future levels of automation, available technologies, and the need to know more about collections and their use.

1. Does the library catalog on a shared cataloging network? If so, what access does the network provide to GPO records?
2. Does the library have or plan to implement an online catalog? Will it permit keyword, SuDocs number searching? Does it include authority control and global change capability?
3. How does the library's clientele learn...
of government publications (e.g., through journal articles, colleagues, indexes, GPO sales brochures, news broadcasts, etc.)?

4. How does the library's clientele look for and obtain government publications? Do patrons look in the library?

5. What are the major strengths of the government publications collection (e.g., congressional materials, geological materials)?

6. Is there a clientele for those collection strengths?

7. Is cataloging the best means of addressing the needs of that clientele?

8. Is manpower available to define the means and related costs of cataloging?

9. Will cataloging be selective? If so, what categories or types of materials will be excluded?

No single library can address all these questions, but, for some, GPO shipping lists may provide one option for testing assumptions, exploring possibilities, and determining costs.

What can shipping lists do to answer these questions? For RLIN and WLN users, they can test the speed with which records can be searched without SuDocs access. For all shared cataloging network users, they can help determine network costs for cataloging U.S. materials. (See table 2.) They can help identify the type of entries available for specific depository series (analytics or series), their timeliness, and their sources. The 040 field contains the cataloging agency. If GPO appears first, it was the first cataloging agent. If it appears elsewhere, GPO modified an existing record. If the Library of Congress has modified a GPO record, DGPO/GLC will be found in this field. Finally, for both network and nonnetwork users, shipping lists can assist in identifying collection strengths, providing crucial direction for determining what will be cataloged, and forming policy that reflects those decisions.

Realizing that the card catalog may not be the single most appropriate vehicle for locating government materials, that not all materials (e.g., posters, pamphlets) merit cataloging, and that cataloging all materials increases cataloging costs, a number of libraries have either established or recommended policies for cataloging select categories of U.S. materials. Peter Graham recommends that large research libraries concentrate cataloging efforts on collections that are not indexed. At the University of Houston, priority is given to "publications of agencies that the average user does not know are government agencies." Whether policies are set or explored, careful arranging and coding of shipping lists can verify the feasibility and costs of policies. Shipping lists must be representative of type and extent of receipts. Monthly statistics can establish the latter. Then lists can be separated into microfiche and paper piles to examine each format separately and into chronological piles to test the timeliness of available records. Coding should reflect specific interests. For instance, items received on deposit can be checked, those considered for cataloging circled, and candidates for series cataloging marked with an S. Special types of materials that could dramatically increase cost if cataloged may be studied separately (e.g., hearings might be preceded by an H), and cataloging agent can be indicated as GPO, LC, or Other.

Such an examination begins to provide a picture of what a library can afford to catalog on the basis of type of catalog records.

**TABLE 2**

<table>
<thead>
<tr>
<th>OCLC Cataloging Costs</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Estimated Number of Cataloged Titles (annual)</strong></td>
</tr>
<tr>
<td>-----------------------</td>
</tr>
<tr>
<td>400</td>
</tr>
</tbody>
</table>

Notes: An online catalog would eliminate the need for cards and reduce costs to $588. No reliable estimate of staff costs exists. With minor record changes, most titles could be processed in five minutes or less. At $10 an hour, 300 hours of professional cataloging time would cost $3,000.
rather than on number of titles; whether records are available; whether they are timely; whether sufficient numbers of records are available for microfiche; and, consequently, whether or not fiche should be included in cataloging programs. For most libraries, whether or not they are able to determine these considerations by searching items on a shared cataloging network, shipping lists may suggest criteria for including or excluding certain types of materials from cataloging. In short, they can be used to qualify general assumptions about the feasibility and affordability of cataloging U.S. materials in a specific environment.

**SUMMARY**

Arguments against cataloging U.S. materials have assumed that cataloging was difficult in the absence of available records, time-consuming because of numerous title and agency name changes, and expensive as a result. Other negative reasons have been that cataloging limits shelf arrangements to Dewey or LC and causes delays between receipt and availability of materials to patrons. On the positive side, despite inadequate and difficult entries, cataloging the materials provides the most convenient, direct, and useful means of access.

In an automated cataloging environment, all of these assumptions can be disproved. Shared-cataloging network users have online access to records; online catalogs minimize the record-editing process and facilitate the simultaneous, separate SuDocs shelf arrangement and cataloging of materials, thus eliminating the usual delays associated with cataloging: the availability of tape records, vendors who can manipulate tapes, and selective cataloging reduce traditional expenses. Finally, studies have shown that the card catalog is not as pivotal to all users' needs as was once thought.

The new cataloging environment mandates a new look at an old debate, encouraging exploration of alternatives and dismissal of some earlier assumptions. Most important, the online alternatives highlight the central importance of knowing more about the need for, use of, and methods of obtaining government information.

---

**REFERENCES AND NOTES**

2. Ibid., p.320.
11. While RLIN and WLN networks provide subject access, they do not provide SuDocs access to GPO records. Subject access to OCLC GPO records is available for the last few years of GPO rec-
ords through BRS. According to the February 1986 OCLC Newsletter, subject access should be available to libraries in early 1987.


17. Ibid., p.10.


19. Ibid.


22. A 354MB disk drive holds approximately 82,000 fully indexed records of 4,000 characters each and costs roughly $20,000. For this configuration, per-record storage costs would be $.24.


24. Ibid., p.29.

25. Ibid.


31. Government publications staff discarded a departmental card catalog after a one-month study indicated that 30 percent of questions could be answered on the basis of prior knowledge and 50 percent through indexes.


33. Ibid., p.124.