

UC San Diego

Articles

Title

Bountiful Harvest: Batch Searching and Distribution of Electronic State Document MARC Records

Permalink

<https://escholarship.org/uc/item/79m754bd>

Journal

DttP: Documents to the People, 36(4)

Authors

Culbertson, Rebecca

Sklar, Annelise

O'Sullivan, Donal

Publication Date

2008

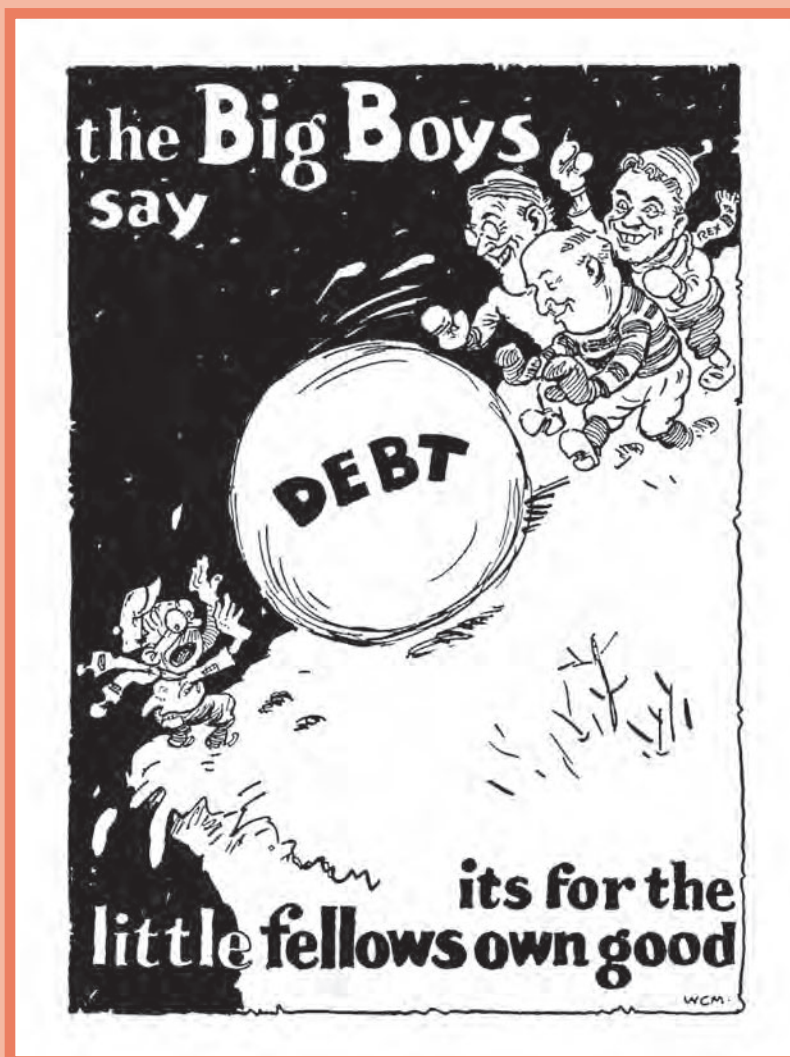
In This Issue

- “Official English”
Legislation
- Layers of Protection:
Beans, Biopiracy, and
Patents
- The Makah Whaling Saga
- The National Archives
and Private Partnerships

DttP

Documents to the People

Winter 2008 | Volume 36, No. 4 | ISSN 0091-2085



DttP: Documents to the People (ISSN: 0091-2085) is published quarterly in spring, summer, fall, and winter by the American Library Association (ALA), 50 East Huron Street, Chicago, IL 60611. It is the official publication of ALA's Government Documents Round Table (GODORT). The annual membership dues of \$20 include \$10 for an annual subscription to *DttP*; to nonmembers prepayment is required in the amount of \$35 in North America and \$45 elsewhere. POSTMASTER: Send address changes to *DttP*, 50 E. Huron St., Chicago, IL 60611. Periodicals postage is paid at Chicago, IL, and at additional mailing offices.

DttP features articles on local, state, national, and international government information and government activities and documents the professional activities of GODORT. The opinions expressed by its contributors are their own and do not necessarily represent those of GODORT.

Editorial Staff

Please see the website for full contact information: www.ala.org/ala/godort/dtpp/aboutdttpp.htm.

Lead Editor: Andrea Severson, 18 Weatherstone, Lake Oswego, OR 97035; (301) 951-4749; dttpp.editor@live.com

Editors:

Jackie Druery, Queen's University; (613) 533-6401; drueryj@post.queensu.ca

Chuck Eckman, UC Berkeley; (510) 642-7999; fax (510) 643-8179; ceckman@library.berkeley.edu

Julie Linden, Yale University, (203) 432-3310; fax: (203) 432-3214; julie.linden@yale.edu

Helen M. Sheehy, Pennsylvania State University Libraries; (814) 863-1347; fax: (814) 863-1403 hms2@psu.edu

Contributions Editor: Ben Amata, CSU Sacramento; (916) 278-5672; fax: (916) 278-7089; bamata@csus.edu

Reviews Editor: Susan Tulis, Southern Illinois University, Morris Library, Mailcode 6632, 605 Agriculture Dr., Carbondale, IL 62901, (618) 453-2522, fax: (618) 453-3440; stulis@lib.siu.edu

Advertising Editor: Sherry DeDecker, Davidson Library, University of California, Santa Barbara, CA 93106-9010, 805-893-3713, fax: 805-893-4676, dedecker@library.ucsb.edu

Advertising: Inquiries about advertising may be addressed to the Advertising Editor. Please see www.ala.org/ala/godort/dtpp/adrates for rates and publication schedule. *DttP* accepts advertising to inform readers of products and services. *DttP* will adhere to all ethical and commonly accepted advertising practices and will make every effort to ensure that its practices in relation to advertising are consistent with those of the other Association publications. *DttP* reserves the right to reject any advertisement deemed not relevant or consistent to the above or to the aims and policies of ALA.

Distribution Manager: ALA Subscription Department, 50 E. Huron St., Chicago, IL 60611. 1-800-545-2433, press 5; fax: (312) 280-1538; subscriptions@ala.org

Subscriptions/Claims: For subscriptions, prepayment is required in the amount of \$35 in North America, \$45 elsewhere. Checks or money orders must be made payable to "ALA/GODORT" and sent to the Distribution Manager. Changes of address and claims six months of the date of the issue must be sent to the Distribution Manager. To purchase back issues, write to: UMI, 300 North Zeeb Rd., Ann Arbor, MI 48106.

Contributions: Articles, news items, letters, and other information intended for publication in *DttP* should be submitted to the Lead Editor. All submitted material is subject to editorial review. Please see the website for additional information, above.

Indexing: Indexed in Library Literature 19, no. 1 (1991), CSA Worldwide Political Science Abstracts 33, no. 1 (2005), and selectively in PAIS 33, no 1 (2005).

Editorial Production: ALA Production Services—Troy D. Linker, Chris Keech, Tim Clifford, and Justine Wells.

Columnists:

By the Numbers
Stephen Woods
Pennsylvania State Univ.
swoods@psu.edu

Geospatial News
Marcy M. Allen
Pennsylvania State Univ.
mallen@psu.edu

International Documents Roundup
James Church
Univ. of California, Berkeley
jchurch@library.berkeley.edu

News from the North
Michael McCafferty
Univ. of Toronto
michael.mccafferty@utoronto.ca

State and Local Documents Roundup
Kris Kasianovitz
Univ. of California, Los Angeles
krisk@library.ucla.edu

Tech Watch
Valerie Glenn
Univ. of Alabama
vglenn@ua.edu
Amy West
Univ. of Minnesota
westx045@umn.edu

Tips from Tim
Tim Byrne
Info. Int'l Assoc.
byrnet@osti.gov

Washington Report
Kevin McClure
Chicago-Kent College of Law
kmclure@kentlaw.edu

DttP

Documents to the People

Winter 2008 | Volume 36, No. 4 | ISSN 0091-2085

Columns

- 4 Editor's Corner
- 5 From the Chair
- 7 Washington Report
- 8 News from the North
- 9 State and Local Documents Roundup
- 11 Tech Watch

Articles

- 13 Student Papers: The Assignment
Cass Hartnett and Amy Stewart-Mailhot
- 15 "Official English" Legislation and Its Effects on Limited English Proficient People
Catherine McMullen
- 21 Layers of Protection
Beans, Biopiracy, and Patents
Rebecca Z. Kuglitsch
- 26 The Makah Whaling Saga
Indigenous Rights, Endangered Animals, and the U.S. Government
J. M. Rohan
- 29 The National Archives and Private Partnerships
A New Trend for the Digital Age?
Esther E. Benson
- 34 George W. Bordner and Pennsylvania Documents
A Remembrance
Mary Spila
- 36 Tips from Tim: Real Life Search String
Tim Byrne

Review

- 39 *Managing Electronic Government Information in Libraries: Issues and Practices*
Reviewed by Kirsten Clark

'Round the Table

- 41 Report from the World Library and Information Congress
74th IFLA General Conference and Council in Québec
Jackie Druery and Marcy Allen
- 43 Councilors Report
- 44 Index to Advertisers

Copyright © 2008 ALA/GODORT

About the Cover: The cover photo is from the UCLA Online Campaign Literature Archive, a century of Los Angeles Elections, digital.library.ucla.edu/campaign. The image is from the 1936 Landon/Knox campaign. (http://digital.library.ucla.edu/campaign/librarian?VIEWPDF=1936_007_003_a). Thanks to Kris Kasianovitz and Gabriella Gray for their suggestions and assistance with the cover selection.

Kit. The *Kit* is being developed by the Critical Thinking Consortium (www.tc2.ca), a non-governmental organization comprised of boards of education, professional associations, and universities. Each module contains lesson materials, classroom activities, links to articles, photographs, graphs, and data and is aimed at the intermediate and secondary (grades seven to twelve) level. The first five modules focus on population change, aging, the family, immigration, and the native population. The remaining modules will cover work, employment and education, ethnic origins and visible minorities, and earnings and income. Each module is available in HTML or PDF and all may be found from the main page (www12.statcan.ca/english/census06/teacherskit/index.cfm). A number of other lesson kits and lesson plans are also available covering such topics as the environment, social trends, and agriculture (www.statcan.ca/english/kits/kits.htm). They are targeted toward the elementary, intermediate, and secondary school levels.

Over the summer, the final versions of a number of 2006 Census *Reference Guides* were released. The guides contain definitions and explanations on census concepts, data quality, and importantly, historical comparability. Eleven such reports have been released to date, starting with the *Families Reference Guide* released in October 2007. The most recent, the *Education Reference Guide*, came out in August 2008.

On a related note, mention should be made of one highly useful resource. Constructing time series of truly comparable statistics can, at times, be a difficult task. Apart from methodology and concepts, there remains the question of whether or not the numbers actually exist. Simply put, if the question was not asked, the numbers will not be there. Laine Ruus, data librarian at the University of Toronto, maintains a highly useful guide to Canadian census questions asked since Confederation, which contains, in tabular format, a list of all questions asked since 1951 (www.chass.utoronto.ca/datalib/censusq.htm). The table indicates the censuses in which they were asked and when they were *first* asked going back to 1871, thus providing users with quick and easy access to the contents of censuses over a period of time.

In March, Statistics Canada launched the *Canada Yearbook Historical Collection* making available, free of charge, yearbooks published between 1867 and 1967. Historical texts, tables, charts, and maps are included and the collection is supplemented by learning resources for students and teachers. The individual yearbooks are available in PDF and the collection includes the eighty-seven English and seventy-two French editions issued over the one hundred years since the Confederation. Separate gateways exist for the English and French collections (www65.statcan.gc.ca/acyb_r000-eng.htm

and www65.statcan.gc.ca/acyb_r000-fra.htm, respectively).

Web Guides

There is at present no comprehensive guide to Canadian government information similar to those provided for Americans at institutions such as the University of Michigan or by professional associations such as GODORT's Federal Documents Task Force. However, many specialized guides do exist that, if taken together, might serve the same purpose as their more comprehensive counterparts south of the border. Sherry Smugler of the University of Toronto has produced an excellent resource entitled *Canadian Parliamentary Publications: A Research Guide* (link.library.utoronto.ca/MyUTL/guides/index.cfm?guide=parl). Though not a guide to the workings of Parliament as such, it does enumerate Parliamentary publications and places them in a context that makes it easy to understand the legislative process. To my mind, it is at present the best such guide and would be useful to bookmark. A more process-oriented guide, taking the form of a tutorial, is being revised for the government information students at the University of Toronto's Faculty of Information. It will be featured and discussed in the next column.

State and Local Documents Roundup

Bountiful Harvest: Batch Searching and Distribution of Electronic State Document MARC Records

Rebecca Culbertson, Annelise Sklar, and Donal O'Sullivan

Time wasted is time lost. What better way to make use of a document cataloger's time than to automate the searching process in OCLC? This is what we are doing at the University of California (UC) San Diego. The key component of this strategy is the *up*: search command in OCLC. This command only became available to OCLC Connexion users in April 2008 and has remained largely undiscovered.

Until July 2008, the UC Shared Cataloging Program (SCP) cataloged online California state documents for all ten UC campuses. While most of the cataloging was value-added copy cataloging, it still required a significant amount of time to find new documents on each website on the government information librarians's list of priority agencies. Some of this time

has been put to better use by subscribing to agency RSS feeds and e-mail lists.

Another solution developed by SCP staff involves batch searching OCLC Connexion for MARC records, downloading to a local save file, and then exporting to the library catalog. While the process does not retrieve every online state document cataloged in OCLC in a given time period, it does allow a library to capture a good portion of very current online state documents cataloging with minimal effort—assuming that someone out there is doing the original cataloging.

In this algorithm, two searches are performed: one for records cataloged by the California State Library, and one for records cataloged by all other libraries. Each search looks for records with specific attributes (for example, material type online and material type state government publication) that have been updated in a particular month. Once the records are downloaded to local Connexion files libraries might choose to add permanent URLs, harvest the documents for an online archive, or send the list to the appropriate bibliographer for review. The records can then be batch-exported to the integrated library system using a combination of macros. OCLC has a number of such macros; optionally, it is very helpful to use a repeat-loop macro program.

The following procedure was tailor-made to search for California state document MARC records but it is easily adaptable to other states by replacing state-specific information. Smaller states could possibly do one search that would get everything at once. Public services librarians take note; it is time to pass this article to your cataloger.

1. Preparation

- a. In OCLC Connexion, create two local save files—one for newly created records and the other for updated records. The updated records may be handled a little differently than the new.
- b. Create Connexion macros that will not only apply the constant data used by your institution but will also allow for certain cataloging inconsistencies. For example, some records have a 007 field and some do not. It is possible to do this work using constant data records alone, but this does have the potential of adding unnecessary duplicate fields (and creating more work for the cataloger). For UC, SCP developed macros to add the appropriate 793 0 local title fields, 949 fields for item and bibliographic records, and a 936 field for tracking purposes (e.g. 936 CalDocs 200807).

- c. Optionally, create separate repeat-loop macros (UC San Diego uses MacroExpress) to enable the cataloger to input how many times they want the Connexion macros to repeat.

2. Searching

- a. There are two searches to be done in Connexion. The first gathers all records that have either been newly input or updated by the California State Library; the second gathers records from all other cataloging institutions during the same time period.

i. **mt:url and cs=cax and mt:sgp and up:200807??**

This search string retrieves records that contain a URL, were cataloged by the California State Library (040 CAX), have a GPub code “s” and were entered or updated during July 2008.

ii. **mt:url not cs=cax and mt:sgp and pl:Sacramento and ll:eng and up:200807??**

This search string is a bit more complex but retrieves records of online state documents published in Sacramento and not cataloged by the California State Library. Limiting to documents published in Sacramento eliminates documents published by other states. Additional searches can be done on other common cities of publication within a state. The ll:eng code limits the catalog record language to English (i.e., excludes parallel records), though the document itself could be in any language.

- b. The resulting records from both of these searches are initially saved to the **new** save file.

3. New records

- a. In the **new** save file, sort the records by OCLC number so that the records are listed in ascending order, i.e., the lowest OCLC number is first.
- b. Starting from the top of the list and moving downwards, identify the first OCLC record where the ENTERED date matches the month that you are searching. Every record prior to this on the list should then be moved to the **updated** save file.
- c. Sort results by **title**, giving a cursory glance to make sure there are no records that should not be there. If you find differing format records with the same 856 field, delete one of the records (for UC, priority order is: online, print, CD-ROM).

- d. Sort results by **format** and highlight all books.
- e. Apply the repeat-loop macro which in turn applies the Connexion macro adding whatever local information is required.
- f. Batch export the records to the local ILS using another MacroExpress macro.

It should be noted that after a detailed examination of the updated records from July 2008, we judged the changes in the monographic records to be reasonably innocuous and as such could be included in the batch process. Records that have appeared before will show up with the “held” symbol if you have attached your library holdings to the records as you go along. Serials and databases, on the other hand, as is their nature, remain problematic and will have to be dealt with on a one-by-one basis.

As mentioned previously, before loading records into the catalog, some libraries might want to add persistent URLs (BibPURLs) or deposit copies into an online archive and use those links. In some cases, government information bibliographers might want to take a look at the records and select only those that meet their collection criteria. Once items have been exported into the local ILS, post-cataloging processing can be completed as usual.

Rebecca Culbertson, Shared Cataloging Program, California Digital Library, rculbertson@ucsd.edu.

Annelise Sklar, Librarian for Political Science, Law & Society, and State, Local, & International Documents, University of California, San Diego, asklar@ucsd.edu.

Donal O’Sullivan, Electronic Resources Cataloger, Shared Cataloging Program, California Digital Library, dosullivan@ucsd.edu.

Tech Watch

Twittering Away the Day

Amy West

Twitter is an account-based microblogging service. You publish updates of 140 characters or fewer and find others to follow as well as having them follow you. Updates are public unless you use a specific command, although a public update can be directed at another user.

You send an update that will show up in the Twitter view

of anyone who follows you and any public searches of Twitter. As you begin to follow other Twitter users, you will see their general updates and their updates directed at specific users. To direct an update to specific user, type “@username” in your post (figure 1).

These screenshots were taken from Twitter on the web, but you can interact with Twitter from a variety of other tools. Twirl is a very powerful and popular desktop client for Twitter and Twitterfox is a popular Firefox browser plugin for viewing Twitter activity. For an extensive list of links to the Twitter applications mentioned in this column, see mashable.com/2008/05/24/14-more-twitter-tools.

Twitter can also be used to broadcast RSS feeds, making it an effective mechanism for keeping up-to-date. Free Government Information (FGI) uses it to share feeds (see figure 2).

What’s nice about Twitter is that even though I haven’t directly participated in any of these instances, I can still benefit from my colleagues’s conversations, as well the feeds they share. That one can get something out of Twitter without having



Figure 1. Directing an Update

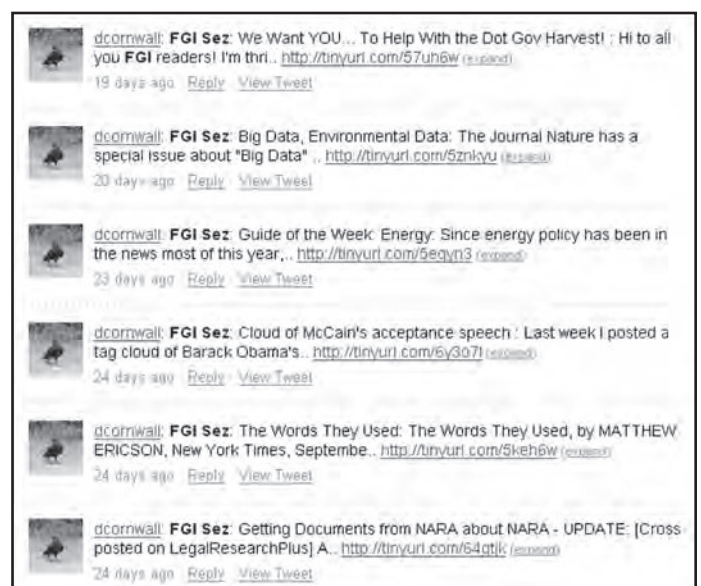


Figure 2. Sharing Feeds