

**University of California/Stanford
Government Information Librarians Group
Meeting March 20, 2003
9 a.m. - 4 p.m.
Agenda and Minutes
Location: UC Riverside Rivera Library, Room 403**

www.library.ucsb.edu/ucgils/Minutes_March20_2003.htm

- Directions: See the UCR campus webpages for campus map and directions to the campus. Directions: <http://www.campusmap.ucr.edu/directions.html> Campus map: <http://www.campusmap.ucr.edu/cgi-bin/mapit.cgi>
- Food arrangements: UCR will provide coffee and morning refreshments. We're planning on a working lunch and UCR will be hosting a buffet lunch.

Meeting Participants:

UC Berkeley: Jim Church, Harrison Dekker, Ron Heckart, Susana Hinajosa, Becky Lhermitte
UC Davis: Linda Kennedy, Marcia Meister
UC Irvine: Margaret Renton, Yvonne Wilson
UC Los Angeles: Jenifer Abramson, Jan Goldsmith, Kris Kasianovitz, Amos Lakos
UC Riverside: Margaret Mooney, Lynne Reasoner
UC San Diego: Elizabeth Cowell, Becky Culbertson, James Jacobs
UC Santa Barbara: Sherry DeDecker, Janet Martorana
UC Santa Cruz: Lucia Orlando
Stanford University: Joan Loftus
California Digital Library: Trisha Cruse

AGENDA

9:00AM - Coffee and morning refreshments
9:30-9:45 Introductions, agenda modifications, and Welcome by Dr. Ruth Jackson, University Librarian, UC Riverside
9:45 [Assessing our collections with statistics](#): GILS collections survey and the Foreign documents survey (J. Martorana and T. Cruse)
[Discussion of the SOPAG GP Task Force draft report](#) (L. Kennedy)
11:45 Working lunch: Continue morning discussion if needed.
12:45 [Mellon grant project](#) (T. Cruse)
2:15 Break
2:30 [Counting California](#) (T. Cruse)
[SCP for CalDocs](#) (B. Culbertson)
Federal updates: [ALA's Task Force on Restrictions on Access to Government Information](#) (S. DeDecker); PATRIOT Act.
Advocacy role for GILS?
Round robin (reporting via email) [Other?](#)
[List of Actions](#) generated from meeting

- Opening and Welcome by Dr. Ruth Jackson, UCR University Librarian - in her welcome she stressed important role of government information plays in our world, especially during these times.
- Janet Martorana - passed out final agenda and copies of round robins submitted by campuses via email. Kris Kasianovitz volunteered to take minutes; James Jacobs and Jan Goldsmith will also contribute notes.
- Introductions around the table: see agenda for attendees

ASSESSING OUR COLLECTIONS WITH STATISTICS (Janet Martorana, Trisha Cruse, Linda Kennedy)

Survey Statistics from December 2002 - distributed via email
Comparison is difficult because of the differing collections, focus, and way that statistics are kept on the campuses.

What do people see as strengths? Our print collections have been greatly reduced due to electronic distribution, especially Federal receipts. Trisha Cruse and Eric Forte did the Documents Data Miner (DDM) study by extracting data from DDM for all campuses. Trisha noted that this study was supposed to show what campuses were selecting uniquely. Overall, there were not a lot of unique titles (see Spreadsheet University of California Federal Depository Libraries: receipts by format: 2001/2002 vs. 1994/1995 UCB; UCSC 16 unique titles (see spreadsheet): Looked at resources from

the North and South campuses. This study shows that moving toward shared resources is a viable way to proceed. What data doesn't show-if titles are in various formats - distinction not reflected - will show up as a unique title.

Sherry Dedecker encourages us all to look at our collection and identify what is truly unique, as this will help guide us with weeding decisions. For example, UCSB is collecting water resources for all states; if other campuses are aware of this collection; this might help with collection policies and/or item-level agreements between campuses.

Suggestion to add this into the SOPAG report: what is out there that we can make use of for collaborative collection?

The question was raised as to how we can add, record, or find out about identifying the unique titles that campuses have? For example, UCB sends all of their documents to the Water Resources Archives so this didn't show up in their item selection for UCB. (<http://www.lib.berkeley.edu/WRCA>),

Discussion ensued about how we can better collect data to analyze collections across the UC's. Trisha noted that the group had hoped to do this, but there wasn't enough consistency across bibliographic records. Each campus catalogs differently (SuDoc, LC, some use Marcive, some OCLC, etc.), govpb code in the bib record's fixed fields (008/28 or 006/11, a, c, f, i, l, m, s, z) hasn't been applied consistently across campuses. Margaret Mooney suggested that using the SuDoc field should yield fairly reliable data. Perhaps with the new Melvyl-T we will be able to use it to better gather statistics. Again, the fact that every campus is very different in the way they collect statistics was emphasized. It is difficult to make exact comparisons across campuses.

Is it possible to devise some cataloging standards for government information bibliographic records so that in the future we will be able to better collect data on these collections?

Linda would like to work with UCLA to identify how many documents are in SRLF. What percentage of collection is there and could these then be identified as "shared documents"? UCLA at some point is hoping to do an inventory of SRLF holdings. There are numerous "brief records" for items that were sent to either SRLF or NRLF. Is it possible for the group to identify what would be minimum cataloging requirements?

Suggestion: We should examine these brief records that were created and sent to a RLF to look at what is or isn't in the bibliographic record. For example, in order to pull up records for materials that are in NRLF, you have to use a call number to do the search; in many cases you cannot search for UN materials, international and federal pamphlets, and foreign government documents by using an author search. (Noted that getting the staff time and funding to go back and review/cleanup such records was probably not feasible on many campuses)

Due to space issues across campuses, more and more will be sent to storage. The question was raised, why select documents, give them minimum bibliographic records and send them to storage, especially if someone else is collecting and giving access to them? Can we get more flexibility with FDLP and the item number selection issue (i.e. if you select an item number to get a particular publication(s), then you also have to take all of the other bits and pieces). Regarding the item number issue, Linda said she was going to work with FDLP to try to come to an agreement that if 1-2 of campuses retain, then depositories can discard these items or they might be able to separate out as a separate item number.

Suggestion: the SOPAG Task Force report should emphasize that materials that are sent to storage areas HAVE adequate cataloging records before they can be sent. Group appeared to be in agreement with this.

The Foreign Documents survey again demonstrated that it is difficult to show cross campus comparison; but individual information is very rich. Chuck Eckman at the last UCGILS steering committee put up information for sharing and used it in conjunction with other surveys/studies, so that we can look at more of the detail

Action: Post Chuck's files from the Foreign Documents Survey to the UCGILS website (this closes a chapter on the 10 year Foreign Documents Survey)
([TOP](#))

DISCUSSION OF THE SOPAG TASK FORCE ON GOVERNMENT INFORMATION (Linda Kennedy)

University Librarians across all campuses are actively discussing this topic. LAUC is also interested in this issue. Task Force membership consists of several UL's and AUL's, see roster:
http://www.slp.ucop.edu/sopag/govinfo_roster.html

Charge is to develop a plan or model for a unified depository library structure at UC's-- See charge:
<http://www.slp.ucop.edu/sopag/govpubsTF7.pdf>

Current environment of government information - we are receiving 70% of Federal documents in electronic format; and getting much less paper. Depository Library stated that materials are distributed in electronic format; the essential titles list is not growing.

State agencies are also conserving by distributing materials electronically. The idea behind being a complete California depository (i.e. "save everything forever") is too difficult to maintain (and too difficult to change the depository rules). California State library is amenable to one campus saving a copy for us to meet the depository requirements.

Linda wants to really stress the unique nature of a government document in Task Force report, i.e. the lack of standardized indexing, need for special tools to access information, etc. How do we deal with these unique service and access issues?

Discussion of a model for single and distributed collections:

Several people expressed deep concern for having only one or two copies of an item.

Most were in agreement that at a minimum we should have copies of materials at 2 locations and not at a single location. Also, there needs to be a good system in place to facilitate ease of retrieval and distribution (see Recommendation #1 "Delivery" on Draft report, p. 5)

Key Issue: Barriers to access of information: Rapid delivery of materials from RLF's is necessary. Several campuses noted that many users will opt not use material if they have to wait or if it is in a format (i.e. microform) that is perceived as difficult to use. Possibility for RLF to receive request, digitize document, and deliver electronically. Are the RLF's equipped (and willing) to do this? Service Points other than documents reference desks need to aware of the shared collections and the ability to request the materials.

Local Options need to be supported! Please emphasize this and make sure it is well integrated into the report. We need to "empower selectors" so they will be assured that local decisions will be protected and respected. Many were concerned that local decisions for retention might not be recognized by UL's if a decision is reached to have shared copies, distributed collections, or if they see electronic copy is being retained, etc. Campuses need support so that they can make local decisions to protect their collections. Also, the group felt it necessary to emphasize collective memory of decisions and local needs. Due to the coming staff changes (administrators, new documents librarians, etc.) have to have a structure in place to make sure that collections and local decisions are protected.

??Big question - How do we make sure we preserve local collection needs; then share collection strengths with others who have less of a demand for other's strengths? Can we set up "distributed responsibility" of subject areas, e.g. UCB Latin America, UCLA Africa???

An example of this already being put into practice are the Elsevier titles are being put into RLF's; then campuses can discard print; materials that are sent to storage become designated shared copy;

Shared copies can still be problematic; because of classes on certain campuses have high demand for materials. Then, a local decision would be to retain these materials not as a shared copy, so that these needs can be met.

We need to work together to make sure that high use materials have good bibliographic access to successfully share; make sure at least one campus has the collection that can be effectively shared.

Most people supported looking at major titles as a first step.

Sharing of information is important, either through various lists or bibliographic records. This will help us all make better collection development decisions.

Last Copy Model

If something is identified as a last copy, it will be retained and saved. The bibliographic record would need to include notes that state something is shared copy and/or last copy. Again, "big ticket" items will be the first project area.

Electronic materials and archiving

We can look at government information in terms of "everything is going electronic" - but we also know that if it's on web it will go away. Most felt that we need to actively do something to archive, preserve what is on the web; then this would give us an opportunity to withdraw what is in print.

The Goals of the Collection:

- Retain service.
- Improve access (through digitization or other electronic means).
- Save space.
- Preserve government information. (Crucial to capture what is on the web - enable us to discard print and retain a single shared copy)

Trisha pointed out that digital is the big ticket item, but space is a major factor that is impacting decisions; also staffing (retirements) are diminishing - need to retain some form of/semblance of the collection as new people come into the dept. Want to set up some form of structure to effectively

Many of us are sending things to storage based on electronic availability; need to obtain a copy of that cataloged item as an archival copy; good place to start is with what is coming on shipping lists.

How will this shared collection affect other shared collection agreements that have been established outside the UC/Stanford group, i.e. UCB Latin American Collection has agreements with UT Austin. Linda emphasized that it is really a local option to determine what will be a shared copy. A good example is UCB and the Congressional Globe.

Community users will still have access to the materials that are deposited at RLF's. In fact, as Linda pointed out, this actually expands their access, because they will have access to all materials that are deposited at RLF (as long as the delivery is good)

Shared collections and restrictions of use: material that is designated as a shared item, can be requested by a department, and designated as building use only (i.e. user cannot check out item)
Discussion on local collecting - what do the campuses have that are unique? We need to continue to collect local areas and share (i.e. EIR of a local forest - only one campus might collect this)

General Comment regarding this plan: Keep in mind that this was top down decision to which we are responding. They [UL's] must save processing time, money and space. We therefore really need to think of a contingency plan to deal with this. There are many titles that would fit into this plan very well - serials and data (e.g. Federal Register); in many cases these titles have good online access as well.

The archival issue: Dim vs. Dark Archive (Trisha)

Complicates the issue if we say we need a pristine copy for use, to digitize, and preserve. If it's really important, then we should preserve it. We, i.e. UC/Stanford, don't have to be the last copy. Same thing with California documents; because there is a fair amount of redundancy; e.g. federal register, or other materials permanently archive on GPO access

Dark copy - cannot be accessed in any way (example are the Elsevier journals that are being withdrawn from collections and sent to an RLF because we have complete full-text access). Materials cannot be lent, because they go into "deep storage"; but this doesn't mean that we can't have an electronic copy.

Dim Archive - this means we digitize material; then can loan the print copy under certain stipulations or circumstances. Materials "can be lent". What kind of restrictions fall onto a dim archive?

Discussion:

How could we ever have a dark archive, since the material is really the government's and should be available for loan? It is possible to have a dark archive with different forms of access - can make copies in various formats; keep a print copy

Subscription serials are being put into a dark archive, that is fine, but the materials we are talking about are the records of our country and government. We need to make sure we have access to this material. We should allow for some titles to have different status, usage, restrictions. It shouldn't be a matter of what is cheapest. When dealing with subscription serials, users are typically after one article. When using government, however, you need to look and use the whole body of information.

How do we want to identify the shared copy?

In order to make sure local options are preserved, we have to clarify and work out the shared copy method. (see page 4 of Draft, paragraph at top that begins with "By ensuring...")

Suggestion - start with the big ticket, large titles (don't have to change so many records)

Current vs. Retrospective Materials -where are we going to place the emphasis? Keep in mind that retrospective materials tend to be non-digital.

Definition of Retrospective: Shared older print not available on the web - dead serials. Are they low use or are they not.

Definition of Current: materials you are now receiving, e.g. if we are going to share a particular item number, certain campuses will emphasize portions of states

Suggestion: Can we move to designate anything that is ALREADY at the SRLF or NRLF will become part of the shared repository? AND anything that gets sent to RLFs will become part of the shared collection?

Discussion:

We also have to consider that the RLFs are also under the gun for space, so is it important to designate these copies as dim, dark, or copy of last resort? Keeping in mind that if something is designated as dim or dark, you cannot rely on lending/borrowing this material.

What do we need beyond the current RLF status for government information? Do we need some special designation that wouldn't allow it to be loaned in the normal way? Can we designate shared copies for building use only? Would the loan requirements have to be changed, if it's the shared copy, and cannot be checked out from the borrowing library? Would this change your local retention decisions?

For example, UCD and UCR have strong agriculture collections; if this isn't particular collection strength for say UCSD, they could opt to discard their agriculture collections and make requests from UCD and UCR; they would be assured that there would be a shared copy in the system. But if something is the shared copy and a high use item, then this might change decisions.

UCR is able to digitize materials through their ILL - BUT what happens once you digitize something - need to make sure this digital copy archived. BUT this also shifts costs - to ILL; burden of work gets shifted to another department.

Also have to keep in mind that digitizing various formats of materials does not always improve access and usability of the information; statistical publications are a good example, but there are many other formats.

This is why Linda would like to try this out with a pilot project. She suggests that dead serials are a good place to start. Others in the group asked if we could also think about current materials that would be well suited to this. There are many documents that we no longer use in print.

Once something becomes shared copy - it cannot be removed as shared copy.

Distributed shared collection vs. all in one place

Can there be a shared copy that is not at a RLF?? Is this reasonable? If we don't feel it is, then Linda won't propose. Folks seemed to think we could possibly move forward on this.

The UC/Stanford Libraries already have Shared Purchase Model that we can use as a basis for our shared copy model. Cataloging records are key; this was one of the weaker points in the shared purchases - it was up to the local campus to catalog. There has been a problem with accessibility to these collections.

Advantages to having one place - easier to manage; RLFs can digitize on demand. CDL could then become the depository of the digitized copies.

Recommendation 2 - Developing a collection framework

We need to develop what it is we are going to share; and then continue to evaluate/revise.

Conducting a pilot project (dead serials and big current titles mentioned earlier) could help with this. Most agreed that we need create an overall collection statement or set of ground rules; Need a framework that nails down the procedures. Need to make sure everyone is on board with the "process". A pilot project might help us iron out the details. Later on, perhaps we could sit and delineate subjects once we have a better grasp of how this all works.

Do we have a timeframe? Could this become part of annual selection and offer a way to evaluate what titles can be added.

Jim Church suggested we name the project: Government Information Management Initiative (**GIMI**) or was that **GIME** - Government Information Management Environment?

A number of the government information departments send portions of their collections to other libraries or departments (health related materials at UCLA go to Biomed). We need to make sure we consult with these other areas to find out about their usage, as well as make sure they are aware of the collection framework. Also, need to make involve and make this clear to other departments involved with processing government information, e.g. RLF's, cataloging departments, etc.

Summation of discussion: Designate what is shared; what is our ability to obtain materials; emphasize local options; enhance marc records

Discussion of Pilot Project: (Recommendation 4 on Draft)

The purpose of this project is to find a way to make the process seamless to the user; so they don't encounter any problems with accessing the information they need. We should gather input from users about the pilot project. As was noted earlier, most users want it now. In doing the pilot, can find out about their behaviors and communicate the results of pilots to them.

Action: Yvonne would like to identify major serial sources for California documents as pilot project? Are there any other CA docs librarians interested in discussing?

(break for lunch 11:50)

Recommendation 7 - Tangible products - handheld digital materials (floppies, cds, etc.)

Still some core titles that we want to make available; can we do this cooperatively; UCB cooperative project, part of Counting CA project - make data available on Sunsite as well; 1950, 60, 70 Census; CountyCity Data Book back to 1944: social science & government data site; 2000 census and forward will not have the ftp-ability; make uc data available thru a similar project of counting ca - where the data is available via ftp for download

Examples of shared data that Linda can include in the report: foreign trade, agriculture data.

Trisha reminded us that these projects (UCSD, UCB, etc.) are about access not about preservation. We need to devise a way to preserve and make data persistently available. The UCSD project is trying to preserve. This varies from the University of Indiana because you can use UCSD data with current statistical programs; the U of I project have to use the old technology/programs. Counting California is trying to preserve the "digital aspects of collections"; but at this point there isn't anything formal that says the data is preserved.

Action: Linda will add the need for preservation of CD-ROM data in the report

Action: The group asked if UCSD was given the support (and funding), would they want to take on the 3.5"floppies? Elizabeth Cowell responded: perhaps, if they had the necessary support; BUT would have to consult with the people who did it first; but think they could make the project happen. It was noted that the UCSD project also helps with the retention decisions on other campuses.

Recommendation 8 - Cooperative Cataloging project for Federal Docs:

This is based on Cooperative Cataloging of California Documents. Those records many of us get from Marcive, LC, or DDM

Suggestion: Jim Church asked that international documents also be inclusion in this project; international documents are overwhelming their cataloging department, which has suffered from poor staffing. International might not be included because this is the SOPAG Task Force, therefore they have to focus on federal.

Action: Linda suggested that we bring up this International issue with GILS. Becky Culbertson responded that any campus can now do the cataloging for these materials. Jim Church could write up a report/request for this, using the CA proposal as a model. He will coordinate with Trisha.

Questions: Is there interest in doing shared cataloging via Marcive?

Advantage is that you get one set of marcive records once. (keep in mind - UCSD only does electronic titles)

Problem with a single Marcive approach is that individual campuses might have different local policies that might not work well with these records.

What happens if all 9 campuses start cooperative cataloging? How would records be efficiently distributed to all campuses? Would each campus have to go to 9 different places to retrieve records?

Recommendation 9 - Microforms issue:

We've seen the use of microforms dropping. Why?

A large amount of materials are available only in this format. Due to merging of departments, knowledge of the collections was lost. Finding documents within larger microfiche sets is also difficult. If we could get catalog records for individual fiche, there would probably be a significant increase in request for them.

Fiche duplication provision that enables us to share fiche across campuses. Would we support fiche sharing/duplication at RLF's or on a campus. Put the materials at RLF; have RLF get fiche digitization equipment; facilitate desktop delivery (or cdrom delivery). Several raised the issue of copyright, since many of these fiche collection are not produced by the government. We can generate an electronic copy under fair use; we just can't put it up on the web.

Recommendation 13 - Reference Sharing for Government Information:

Does the group support something of this nature (digital reference type program)?

Most supported this idea. The people who are new to the group and to government information all chimed in that it would be very helpful to a list of this nature for several reasons: help with reference questions, but also help them to identify who the "specialists" in our group are.

Suggestions for how to do this:

- Use the GovInfo-L - already in place. A few noted that we run the risk inundating the list with too many questions, then people won't read;
- Set up another list specifically for this purpose or Web board; use digital reference software - when it works.
- Set up a "volunteer list serve" - designate one government document librarian from each campus to help monitor questions.
- Trisha noted that there is new software that will collect all of users questions; they can evaluate, send to campuses, etc (to help with new melvyl systems) this could potentially be used by campuses for a discussion forum

How can we, as the UCGILS group, provide support to people who are completely new to this area of reference and information, especially since many of our government documents reference desks are gone? List-servs, cross training?

Action: Linda will explore and expand this idea.

Action: Kris K. will add a list of collection specialists and library collection strengths to UCGILS web site; Linda will add some of the strengths, how reference is given; how to contact; in the appendix

[\(TOP\)](#)

PRESERVING DIGITAL DOCUMENTS: UPDATE ON THE CDL MELLON GRANT (Trisha Cruse)

<http://www.cdlib.org/libstaff/sharedcoll/gii/MellonGovInfo/>

Handouts:

" Appendix A from the Environmental Scan, Preliminary Survey Results; Interim Report pp. 10-11 available at:

http://www.cdlib.org/libstaff/sharedcoll/gii/MellonGovInfo/CDL_Mellon_environmental_scan_ver.3.pdf

" Neutral Tools for capture, curation, and preservation (ask Trisha for copy)

Note: Appendix A contains most of the information discussed for this portion - administration, capture, curation, and preservation.

Is there currently a place where (born) digital government documents are preserved?

Trisha: The Mellon Grant will help us address the issue of archiving and preserving digital government information. We will need to develop a framework. This might not be worked out until summer.

Concerns with digital documents - the material is here one day and gone the next. A very good example is the Pete Wilson website. Once he left office, his web site vanished and nobody had archived or preserved it.

Trisha asked if there are any campuses who could work on this until a framework is in place? UCLA will try to find out if we are able to gain server space for born digital local documents.

For those digital items that have already been cataloged, there is a possibility that the actual document no longer matches the cataloging or that the digital item (content) is no longer available. Trisha felt we should put our energies towards new titles that are being selected and not go back to the electronic California documents that have already been cataloged.

Discussion of using a PURL server - it was noted that such a server doesn't necessarily solve the preservation problem, just the cataloging problem. The best way to archive these documents is to save it to CDL servers.

Developing an information base for this:

-Series of surveys (phone surveys) of high level people - who's doing what; pitfalls; challenges.

-Groups who are actually doing this asked same questions

Results: pretty much same challenges & nobody is talking to each other

CDL Mellon group will have a meeting of those who are working with these kind of collections. GPO will sponsor a meeting to bring together these people. Elisabeth Cowell at UCSD will be hosting the meeting, April 22nd meeting. This information will be presented to a group of librarians to figure out the tools they need, concerns they have.

The Technical Survey was conducted at the San Diego Supercomputer Center. From this they categorized various approaches to capture, curation, and preservation by using different crawlers and analyzing the results.

Webmetrics

What is the size of the .gov - how big is it; file types; other intricacies;

This is the most difficult to determine, because .gov is not a private entity there aren't any webmetrics

Surface web vs. deepweb (invisible web)

Top level pages vs. what is contained in databases

Blue Planet, a web think tank, estimates that the deep web is 4-5 times larger than surface web, with 90% of deep web being government information.

Qualitative analysis - what is the rate of change by key users of .gov info - what is their threshold of information loss?

Test Crawl - partnered w/ Stanford @ comp sci group

Given list of websites to capture: US State Interior, US Senate after election, EPA CA Energy Commission, California State Water Resources Control Board. Crawler goes to a level of 10 in the web site (designated as an adequate level)

Stored info; crawl took about a week and yielded 10million pages - only surface web pages.

Now working with Stanford to fund positions to further their development and go after deep web information.

San Diego Supercomputing Center - work with them to look at various file formats (.pdf, .gifs) look at replication of these files (ex. 200 happy face gifs, do we want to reproduce the .gifs over and over?). They have technology/we have the need, and thus was built a relationship.

They are also the recipients of DLI 2 grants from NSF.

Issues with Capture: hoovering up the web or sucking up the web

- Bulk crawling doesn't work; can't get to the deep web; too many throw away items.
- Hand crawling won't work - pages change too fast. Cornell group is examining at how to predict the instability of web sites.
- Authenticity problems (Judy Russel at GPO) - they don't have anything that designates something as authentic from their website; GPO is working on a digital certificate. Role of the Mellon Grant is to develop a trusted repository; trust that info that comes from the repository is trustworthy. RLG has a great deal written on trusted repository
- Metadata development is most time consuming and costly; or are there adequate procedures for auto generation of metadata?
- How do you serve diverse community needs? University needs, agency needs
- Need to establish middle ground or theory neutral approach - neither bulk or hand; what ever best fits the need of the local community can be practiced.

Discussion & Questions:

Q: Can you go thru password protected sites? Is this really possible?

A: Not going through password protected sites, rather a site that requires the user to do something to look at data or information; way to automate queries at these sites; example EPA reports in database.

This proposal is mainly for web based materials, but it offers a good possibility to "free the fiche"

Q: Is the technology such that you can get all of the versions/editions of the web sites as they get revised?

A: Not a lot of technology out there at the moment that will do this; but will want to explore this. Some examples of technology that already exists: Way back machine; internet archive; elections 2000 project - only get about 1/3 of the material. See: <http://www.archive.org/> and <http://www.loc.gov/minerva/collect/elec2002/index.html>

Do we really want to collect snapshots of all dynamically generated sites (e.g. bulletin boards)? Any type of site that has an underlying database, is going to be difficult to obtain.

CDL is going to rely upon us/campuses to define what it is we want - do we want the facts that something exists or the actual material?

Practical applications for library? Enables us to capture what we want from .gov web sites and then can link up to catalog.

Q: Will this software be able to generate content channels for portals.

A: San Diego is working on datagrids and dynamic ontologies to create portals based on user needs.

Q: Can we capture that which has already been cataloged?

A: We will be looking at capturing actual titles and entire web sites. This project is actually looking at all levels of capture.

The difficult question is where to put all of the data? Disc space at UCB, San Diego Supercomputer Center, Stanford???

US Interior web site was given as a suggestion to preserve the different changes to the web site over time.

Q: How do all of these items relate to individual OPAC? What will the user access, interface look like? How will this get into a local system? Does this meet the basic need of campuses to get digital item and then get it into OPAC? Could you convert records into marc format?

Infomine - moving into a focused crawling method; this includes adding LCSH; Melvyl records were used as training data; This already provided a way to look at surface web site and extract the key phrases and generate metadata; if you have this group of keywords - this is most likely the subject of the site. Robot generated records, metadata created; human looks at; robot generated lists given to human selector with a given profile to review records.

Action: Trisha would like to have input from us on the kind of tools we need to capture, curate, and preserve? Even a paragraph of what we would like to do would be helpful.

LOCKSS Docs project with Chuck Eckman - Lots of Copies Keep Stuff Safe; peer to peer communication to make sure that information is copied and shared and preserved. See: <http://lockss-docs.stanford.edu/> for project background. See also: <http://www.lockss.org/>

The LOCKSS project would help to add to the duplication of shared collection materials. It would help keep the base of shared copies retained at 2 copies rather than a base of 1.

[\(TOP\)](#)

COUNTING CALIFORNIA UPDATE (Trisha Cruse)

<http://countingcalifornia.cdlib.org/title/sf32000.html>

SF3 - last big dataset to be loaded; they managed to put up SF3 within a week - as compared to 1990 - which took 8 months..

Data documentation initiative - metadata standard; nobody else using this because "it's simply too hard"; Trying to decide which metadata standard to use. CDL - can't afford the kind of resource that requires so much time and person-power. Harrison attended the metadata meeting with Trisha.

Margaret (??) wants to form a group to work on changing the interface; Formation of a group - one person to head up interface design. Front end of CA Rand Statistics might be a good model (portion when you can select all different variables) - they have the public in mind when designing an interface.

They will be adding new datasets - 1947-77 from County Databooks; and 1950, 1960, and 1970 Censuses (making 6 Decennial Censuses available online). The hope is to add track level data as well. University of Minnesota has census data back to beginning, including maps; see: <http://www.ipums.umn.edu/>

[\(TOP\)](#)

CALIFORNIA STATE DOCUMENTS SHARED CATALOGING UPDATE(Becky Culbertson)

Handout distributed - includes a list of agency websites that are searched monthly for electronic documents.

Becky requested and received priorities of agencies to mine for web government documents.

Action: If you have other agencies, titles, or databases that you would like mined or feel would benefit the entire system, please send them to Becky, rculbertson@ucsd.edu

Action: Becky will send out via govinfo-l the procedures for adding agencies/titles.

A cataloger is mining these sites on a monthly basis for material. The California Publications Catalog is also used so as not to duplicate efforts or records that are already on OCLC. They are using OCLC bibpurl service; same kind of purl server service at CDL; enhances OCLC records. A broken purl report every week is received every week for these materials.

Becky interested in cataloging databases especially those that have replaced serials. An example is from the CA Dept. of Education, Dataquest. See: <http://data1.cde.ca.gov/dataquest/>

Cataloging records are being distributed to libraries each month. System liaisons have been notified that they can come to the UCSD site and download these records (each campus has their own port; records will be available for 3 month.)

Working on distributing monographic sets of records; review file of print records; add electronic resource; note that says electronic reproduction. Example are the Early American Imprints: Either UCSD will be shipping them out or Readex will distribute records. Bev French wants them to code if something is either free or if there is a cost involved. More to come on this. Will soon be adding free or for cost notes.

[\(TOP\)](#)

ALA TASK FORCE ON RESTRICTIONS TO ACCESS

Handout: from web site that has charge for TF

http://www2.library.unr.edu/dept/bgic/Duncan/ragi_qtc.doc

All voiced concern about government information that is being removed; but more importantly - we need to ask for title lists that document WHAT was removed (even if we cannot get the content - but a way to document was once available).

Question: Do we want to submit a group letter for UCGILS on this topic?

What are people doing to inform people about the Patriot Act?

Mention of the Santa Cruz Public Library, posting of signs, and their resolution;

(Posting of signs can create a scare in public users; but shouldn't they be aware of what is happening?) See:

<http://www.santacruzpl.org/> and for their resolution:

<http://www.santacruzpl.org/libraryadmin/ljpb/patres.shtml>

Susana Hinojosa is part of a group on that is addressing this issue; do OPACs completely delete circulation files? What about information retained in backup tapes and campus servers? She encourages everyone to see the video on this topic.

Action: UCGILS Steering committee will draft responses for passing a resolution; with a statement opposing the passage of the Patriot Act II; There is a good chance that this could be funneled up to CDL top level.

LAUC-SD has already written a letter to legislators including Karrie Peterson's LA Times Op-ed

Has anyone been approached by FBI recently? Not to anyone's knowledge.

General discussion ensued - we need a unified voice that states we think this is wrong and give clear reasons why.

UC Davis Government Information Library Web Site has an Update on Iraq link:

<http://www.lib.ucdavis.edu/govdoc/About/iraq.html>

Many universities are creating pages like this to increase public awareness of legislation, international documents, etc.

Action: Kris will add major issues on Iraq to UCGILS web site

[\(TOP\)](#)

OTHER ISSUES:

CQ Electronic Library - who has it?

UCB decided not to select; UCI though favorably of it, doesn't have it;

Anyone canceling print in lieu of CQ Electronic Library? No.

Election Data will be another component of the CQ E L

People were interested in learning more about the content of UCLA Campaign Archive - Jenifer Abramson addressed this.

Web site: <http://www.library.ucla.edu/mgi/campaign/index.htm>

Information on the content: <http://www.library.ucla.edu/mgi/campaign/background.htm>

Covers mainly California and Los Angeles primary and general election materials, from 1908-present. Early materials are scanned images of print flyers, current materials include captures of candidate's web sites; realia is not included.

LIST OF ACTIONS BY PERSON

Sherry DeDecker	Post Chuck's files from the Foreign Documents Survey to the UCGILS website (this closes a chapter on the 10 year Foreign Documents Survey) CONTEXT
Everyone	Trisha would like to have input from us on the kind of tools we need to capture, curate, and preserve government information? Even a paragraph of what we would like to do would be helpful. CONTEXT

Everyone	If you have other agencies, titles, or databases that you would like mined or feel would benefit the entire system, please send them to Becky, rculbertson@ucsd.edu CONTEXT
Elizabeth Cowell	The group asked if UCSD was given the support (and funding), would they want to take on the 3.5"floppies? Elizabeth responded: perhaps, if they had the necessary support; BUT would have to consult with the people who did it first; but think they could make the project happen. It was noted that the UCSD project also helps with the retention decisions on other campuses. CONTEXT
Becky Culbertson	Becky will send out via govinfo-l the procedures for adding agencies/titles to the list of Shared Cataloging Project for California Documents. CONTEXT
Jim Church Linda Kennedy Trish Cruise	Linda suggested that we bring up this International issue with GILS. Becky Culbertson responded that any campus can now do the cataloging for these materials. Jim Church could write up a report/request for this, using the CA proposal as a model. He will coordinate with Trisha. CONTEXT
Linda Kennedy	Linda will add the need for preservation of CD-ROM data in the report. CONTEXT
Linda Kennedy	Linda will explore and expand the idea of reference sharing for government information. CONTEXT
Kris Kasianovitz Linda Kennedy	Kris K. will add a list of collection specialists and library collection strengths to UCGILS web site; Linda will add some of the strengths, how reference is given; how to contact; in the appendix. CONTEXT
Kris Kasianovitz	Kris will add major issues on Iraq to UCGILS web site. CONTEXT
UCGILS Steering Committee	UCGILS Steering committee will draft responses for passing a resolution; with a statement opposing the passage of the Patriot Act II; There is a good chance that this could be funneled up to CDL top level. CONTEXT
Yvonne Wilson	Yvonne would like to identify major serial sources for California documents as pilot project? Are there any other CA docs librarians interested in discussing? CONTEXT