DLSTF: Activities and Progress (March – December 2009)

The Task Force has successfully:

1. Initiated work in each of the four areas of the charge, as noted.

2. Developed a WBS (Work Breakdown Structure) for the Task Force activities

3. Developed a Working Definition for the UC Digital Collection:

   The University of California (UC) Digital Collection is a collaborative effort of the UC Libraries to strategically create, manage, preserve and enable reuse of authoritative digital collections. It features selected resources both licensed for and created within the UC system, targeted to the diverse needs of a vast and broad ranging academic community. Using advanced technologies, the UC Digital Collection is integrated with and extends the collections and services of the UC Libraries. (Revised 6/12/2009)

4. Initiated two inventories: *Current and Planned Digital Initiatives by Campus*; and *Inventory of Current Technologies Used to Support Digital Collections-Initiatives by Campus*

- Each campus was asked to provide an inventory of current and planned digital initiatives. For each initiative they were also asked to provide the following information: technology used to create the digital content, technology used for providing end-user access, metadata standards used, original format of the content, file format, why/how the initiative was undertaken, and if it was a collaborative project.

- The granularity with which people defined the term "initiative" varied somewhat, but most defined an initiative as a collection. Examples include local history images, a selection of medieval manuscripts, and newspapers created by Japanese American internees.

- As of this writing, there were over 300 initiatives, with all campuses reporting. Most campuses have between 2 and 18 initiatives, with UCSD at 57 (many are one-offs) and UCB with well over 200. Some categories and observations follow, with the caveat that the inventories are a work in progress; not every campus has yet reported on every issue, meaning that there may be some variance in the current report. We will build on and complete the inventories in the coming months.

  a. Subject Area

     The subject area of the initiatives is extremely broad, ranging geographically from local to national to global in scope. The majority of the items are humanities/social sciences, with the exception of UCSF collections, as well as a few science-related collections at UCLA, UCM and UCSD.

  b. Technology Creation

     Campuses reported using a variety of scanner and OCR tools. Only three database systems (Webgendb, HP image database, and DAMS) are currently listed.

  c. Technology Access

     The most reported access method was via local campus website (35 instances reported); after that it was through CDL (12), more specifically via OAC/Calisphere or eScholarship. For most in-process projects, technology access was reported as "yet to be determined."

    d.   <u>Metadata</u>

        The most-used standards are Dublin Core and METS. Other standards used include EAD, TEI, VRA CORE, MODS

    e.   <u>Original Format</u>

        Most of the original formats were printed graphic art and/or texts and images (including maps), followed closely by photographs. There were three audio initiatives reported and only one microfilm initiative.

    f.   <u>File Format</u>

        Jpg/tiff was by far the most common format, followed by PDF.

5. Initiated Survey of Digital Library Definitions and Models of other institutions that have a "UC Digital Collection" concept

One facet of our charge was to explore the notion of and develop a broad definition for the "UC Digital Collection" and its relationship to current/traditional collections. To begin this work, we performed an environmental scan seeking out institutions that had a digital library or a digital collection. By gathering this data, we would acquire a baseline understanding of our peers, enabling our development of the definition of a UC Digital Collection.

Our survey was international in scope and identified a broad range of definitions and models revealing various configurations of digital content delivery and production services. For example, typically single institutions defined themselves as dedicated to the production, maintenance, delivery, and preservation of a wide range of high-quality networked information resources supporting education. A single institution in the population we surveyed additionally defined themselves as providing leadership in digital library development.

We also surveyed consortial models, which typically define themselves as collecting, preserving and making available digital content in the support of education, the common good and future generations. A single organization defined itself as providing world-class leadership in innovative use of information and communication technology. Although we did not include the California Digital Library (CDL) in the survey group, CDL would easily be grouped with the consortial organizations, as it supports the assembly and creative use of the world's scholarship for the UC Libraries and provides tools and services enabling the sharing and use of these materials by said libraries.

6. Completed SWOT (Strengths, Weaknesses, Opportunities, Threats (Charge: 1D, 2A, 3B, 3D, 4):

The Task Force conducted a SWOT analysis in order to help us more deeply understand the strengths, weaknesses, threats and opportunities of creating the UC Digital Collection. The analysis will assist the Task Force in making recommendations regarding collaboration within and beyond UC and in developing a strategy for the "Second Phase."

First and foremost among the strengths identified was that UC collectively has a rich and diverse collection. Our instruction and research programs rely heavily on our collections. We have a strong technological foundation in place, including the experience handling and managing large files of catalog records (i.e., Melvyl). In terms of funding, UC has an excellent track record of receiving grants because we are considered a leader among libraries.

Some of the identified strengths, we realized, represented a "double-edged sword." For example, we have a long history of initiating and implementing successful collaborative projects, but we have little

experience partnering with UC non-library organizations and private entities. We have a system-wide committee structure in place, but it has not been tested with an initiative such as this one. We have many standards for creating and sharing digital items, but they do not address all material types. Many of our campuses and the CDL have robust technology and infrastructure in place, but the levels of technology varies across campuses. Finally, being as large as we are, somewhere across UC we have expertise in all areas needed, but the trick—especially given the current budget constraints—is to successfully harness that expertise.

Areas of weakness include lack of project management skills; inconsistent user experience/interface skills across campuses; lack of a system-wide authentication system; lack of funding in general and in particular for storage and for digitization; lack of a sustainable model that would make digital projects part of our daily process, rather than one-off projects; and finally, lack of a system-wide mechanism for managing and documenting copyright.

The threats that were identified ranged from the current government lawsuit with Google; lack of legislative support for higher education; need for all ten campuses' buy-in; the current financial situation and the uncertainty of when it will improve; cost savings that may not realized; the complexity of copyright issues; and, finally, the challenges of decommissioning legacy technology and staying ahead given the rate of technology change.

As a system, we recognize several opportunities, many of which address the weaknesses identified above. We have the opportunity to develop of new technology tools; to build capacity beyond the OPAC; collaborate with new partners; improve the end user experience with our systems; and even, as a result, generate new revenue streams. There is the possibility that we may serve new audiences and provide wider access to UC's collections. We may realize significant cost savings from pooling resources and through the development and implementation of more standardized processes, as well as find increased opportunities for funding. In addition, we could build on and improve current system-wide services such as CDL's DPR, OAC, Melvyl/WorldCat Local, eScholarship, and HathiTrust.

This UC digital library initiative should be considered critical in order for us to maintain our respected position in the library world. It will also provide us with an opportunity to show off UC's best and brightest and attract additional faculty and students.

7. Used a number of methods and tools to work as a task force

- Held 20 conference calls and one in-person meeting

- Used collaborative software: GoToMeeting, ReadyTalk, and Mind42.

- Established the Task Force Wiki on the UCSF Confluence Wiki site to record agendas, minutes, working documents, background documents, and inventories.

- Used project management strategies and tools, including WBS, SWOT, and inventories.

8. A number of the calls in the Fall were devoted to discussions with other UC groups on the charge similarities and future opportunities.
- There were a couple of calls between the DLSTF and CDC. During the call with Chuck Eckman (UCB), CDC Chair and the DLSTF we talked about the intersection of that CDC document with our work. Other ideas explored were the creation of a systemwide collection development policy for digital resources, This topic was further explored on a call when two DLSTF members joined a regular CDC call. The discussion included the recognition of areas of overlap with collection

development and other other groups. The question was posed where does our the DLSTF work end and theirs begin? It was concluded that the DLSTF was acting in a temporary "big-picture" framework that will make recommendations for standing groups that have more of a direct role. A major priority for CDC is born-digital, especially ETDs

- The DLSTF also conducted a call with Trisha Cruse and John Kunze (CDL).They presented a slideshow on UC3 A New Brand for the Digital Preservation Program Preservation. The DLSTF reflections on the presentation were to make sure the DLSTF and any future digital library group work closely with UC3 to help campuses figure out the cost benefit of hosting services locally or contributing to the DPR. There should also be a formal process or strategy for each campus to consult with UC3 to define something that meets their needs

9. There were many discussion by the task force about the lack of clarity surrounding CDL services and how they help/augment/intersect with what we do locally. We expressed the need for clarity on these questions:

- How do the WAS, DPR, OAC, eScholarship fit together and relate to each other?
- Why would you use the three above? Would you ever put something in both, and when?
- And once you know what there is at CDL, do campuses know how to go about getting to them? Are people really clear what the path to the different CDL services are?
- Who gets the information about CDL services?
- How do we strengthen communication so everyone at UC knows about this, doesn't find out about it at an outside meeting?
- Who recommends and maintains standards? CDL's role?
- A consensus came out of this discussion that what is needed is a "Roadmap" or "Manual" for Digital Libraries at UC

10. During the Fall the DLSTF lost two valuable members Gordon Theil,UCLA and Robin Dale, UCSC