

## Catalogers Group, January 5, 2006 Minutes

Discussion of: *CDL Guidelines for Digital Objects (version 2.0, dated Nov. 2005)*  
[http://www.cdlib.org/inside/diglib/guidelines/cdl\\_gdo\\_v2\\_draft.pdf](http://www.cdlib.org/inside/diglib/guidelines/cdl_gdo_v2_draft.pdf)

Present: S. Layne (moderator), J. Riemer, C. Kincy, K. Wilson, V. Bross, M. Beck (recorder),  
L. Ratliff, B. Allahyar, C. Horning, L. Mendes, A. Riggio, C. Miller

Scope (Section 1.1): Why is e-scholarship excluded from the scope of the guidelines? It seems these would be important to include, if federated searching is to cover all UC-controlled content. Response: maybe the guidelines for e-scholarship are actually stricter than CDL's "sliding scale" guidelines, and if so, are there conflicts between the two? Having separate sets of guidelines for different types of material makes it more difficult to identify when they diverge (or converge). It also raises questions about applying standards consistently across systems. [Clarification from Jan. 10th Digital Library Council meeting: The CDL guidelines govern resources that UC staff will be creating and managing. The e-Scholarship Repository documents will be submitted by authors, and the barrier is therefore lower.]

Despite the "sliding scale" assurances, there really seem to be only two choices for handling metadata, the Basic or the Enhanced service level. Question: Once an institution has chosen one or the other level of service, do they always have to enter metadata at that same level? Or can either level be chosen on a project-by-project (or even record-by-record) basis?

There is virtually no metadata required at the Basic service level. A title is generated from the file name if none is provided. (Comment: it seems almost like a Swiss bank account; you can deposit it, but only the institution and the depositor can ever find it again!) Question: Why even bother with the Basic level? Comments in response: perhaps just to get a record in the system (pure archiving), with the hope that someone will someday enrich the metadata. There is a potential problem in identifying duplicate records if there is so little metadata and thus little if no content to compare from file to file. What metadata will permit identifying and distributing the records for a particular digital library project?

Regarding entering data at the Enhanced service level: Why MODS and not Dublin Core as the preferred schema for entering metadata (and, also, to which "qualified" Dublin Core do the guidelines refer)? Comments in response: Perhaps because MODS is more like the MARC format. Also, Dublin Core does not handle non-filing characters or name-title pairings effectively. In section 3.2.1, it should be clarified that it is not the format that is rich, but the content. Both MARC and DC Qualified can contain either very full or quite skimpy data.

The listed identifiers used in the Enhanced service level may not necessarily be stable or unique. The CDL definition of an "identifier" may not be what we think of as an identifier, mainly a unique match point for any given record.

Question: Do we wish for CDL to always create an ARK (or some kind of system identifier) for every incoming object, regardless of whether the contributing institution has supplied its own

identifier? Comment: CDL should separate the creation of an ARK for use within its own system from the basic need of creating a unique identifier for incoming materials.

The guidelines point to existing databases for name forms, subject headings, etc. There appears to be an attempt at standardization, and CDL lists a variety of established metadata standards from which to choose.

Comment: Sometimes the instructions in the introduction to the Enhanced service level seem more restrictive than the numerous examples that follow. For example, regarding dates, there appears to be too many options as to how to construct them. It might be good to for the guidelines to instruct that once an institution chooses any of the available standards, to consistently apply them throughout the creation of its metadata.

Question: what is meant by the metadata tag "Subject (Title)"? Comments in response: Probably name/title or uniform title? This could be a confusing label to some.

In Appendix B, the full form of "TGM" needs to be corrected to "Thesaurus *for* Graphic Materials." The source code for TGM I is 'lctgm,' while 'gmgpc' corresponds only to TGM II.

Comments re: Appendix D, Linking from Digital Objects to Collection Descriptions:

There are two options for creating the parent (collection level) record, using a MARC or EAD record. CDL needs to find a way to wrap the ARK to the parent record.

Question regarding local (UCLA?) practice: Under whose purview will the creation of these parent records fall? If MARC, Cataloging and Metadata Center staff? If EAD, Special Collections staff?