

MODS/Dublin Core Record Utilization Task Group
Preliminary Report
Philip E. Schreur, Chair
6/5/2009

At the November 2008 PoCo meeting, a discussion ensued over the possibility of contributing records to the BIBCO and CONSER programs via non-MARC metadata standards, in particular MODS and Dublin Core. The Committee decided to recommend the formation of a task group to explore the issues surrounding this possibility. Soon afterwards the MODS/DC Task Group was formed and, in February of 2009, the PCC Steering Committee formally approved the charge. The Task Group has been charged with investigating the current use of DC and MODS by PCC libraries, any representation of these records in WorldCat, and recommendations on whether and how the records in these schemas could be shared.

The Task Group began its investigation by creating a targeted survey and distributing it to the official BIBCO and CONSER representatives of participating PCC libraries. Thirty-six institutions responded. Approximately 50% of those institutions create Dublin Core records and 25% create MODS records; however, only 9% contribute those records to OCLC. The most common method of contribution is to allow OCLC to harvest the records via OAI. For the most part, records are discovered locally outside of the traditional library catalog, either through a digital asset management system (CONTENTdm, DSpace, ArtStor, Fedora, etc.) a locally developed discovery platform, or DLF Aquifer. A few of the institutions make these records discoverable along with the items in their traditional catalog through commercial products (Aquabrowser, Primo).

OCLC was approached to discover how members could contribute records, how the records themselves are handled and how they are made discoverable. Currently, individual Dublin Core records can be contributed to OCLC via Connexion. When these records are contributed, they are converted from Dublin Core to OCLC's Common Data format, flipped to MARC, and contributed to WorldCat. Records contributed via batchload must go through the same process.

In moving content from a rich format to a simpler format, OCLC has experimented with three methods of preserving data. The first is to create an application profile consisting of the simple format, extended with elements from the detailed format. The second is to map complex elements into the best place in the simple format and modify it with an attribute that identifies a special encoding. And the last is to map the inputs to various standards as they flow through OCLC's systems, but preserve the original record. By using one of these three methods, complex data can be preserved in moving from a rich format to a simpler one, however, data originally created in a simple format cannot automatically be mapped to differentiated elements in a more complex one.

The Task Group has finished its examination of the BIBCO standard and what is possible to express in MODS and DC. It is now completing the same examination of the CONSER standard. Once this is complete, it will begin to pull together its recommendations based on all the above findings.

The Task Group is still on track to have the final report to you by September 20th. The recommendations will most likely consist of a best path, questions to be resolved, and next steps in pursuing that goal. The topic has been as fascinating as I had hoped and the Task Group looks forward to completing its charge in the time frame allotted.