

Opening Address: Introduction to CDWA Lite

Ken Hama, Executive Director, Digital Policy & Initiatives, J. Paul Getty Trust

Three years ago a small team at the Getty set out to see if we could create an easier model for sharing images with aggregators like ARTstor.

We wanted to end the model of one-off production efforts to contribute to AMICO, ARTstor, etc.

Ideally we wanted to contribute one time *to the network* and let anyone anywhere aggregate at least meta-data, but also images, depending on rights and the Getty programs' openness to aggregation.

We wanted to do this with a relatively low profile, not make the project unrealizable by adding every requirement and every meta-data possibility.

- We wanted to describe a lower the bar approach for how to catalog for resource-discovery, which would have a small required data set.
- We wanted to lower the bar for technology by finding existing community solutions, as well as roles for vendors and aggregators, in supporting shared technology.
- We wanted to get close to an expectation of universal discovery in which, unless everyone is participating, the real promise of cultural heritage online is simply not viable.

We wanted to use existing standards and protocols that would be applicable for collections of unique objects across the various management styles of archives, libraries and museums.

This describes the values we were looking for in creating CDWA-Lite and its implementation in the Open Archives Initiative Protocol for Metadata Harvesting.

Let me also say that the 'we' in this included staff from the Getty's Web team: Jack Ludden, Ted Dancescu, and Cris Grunca; the Getty Research Institute: Murtha Baca, Karim Bougida, Deb Lenert, Patricia Harpring and Joe Shubitowski; the Getty Museum: Erin Coburn; and from ARTstor: Bill Ying, Nancy Allen, Ameer Ahmed, Dustin Wees, Emerson Morgan, and Eric Li.

I know that is a long list of names – and even at that I've probably left a few out – but as this project moves forward we should take care to record the attribution history of the developed data standard in the standard itself.

Using two collections – paintings in the Museum and a collection of photographs of tapestries in the Library – we set out to automate the contribution of images and metadata to ARTstor. And eventually any other aggregator.

- The data exchange protocol was OAI-PMH with images either included (thumbnails) or by reference (pointers to and FTP server).
- The data structure standard was CDWA-Lite, a small number of data elements and yet smaller number of required data elements drawn from CDWA.
- The data content standard was CCO, providing an understanding of content across libraries, archives, museums and visual resource collections.
- The data value standards we turned to were those applicable to the collections we were working with: AAT, ULAN. Naturally a wide array is eventually to be included here.

This migrated outside the Getty first with a prototype, and that was done with lots of assistance from ARTstor. The prototype effort resulted in the presence of Getty collections in ARTstor that are available today.

The Getty also worked closely with OCLC to modify OAI Server Software to work in conjunction with CDWA Lite, which is called OAICATMuseum, and which Guenter will share with us more shortly.

Next, through meetings like this and an online survey we asked for broad input on the usability and usefulness in a community review. The CDWA Lite Web survey received 144 responses representing fields as diverse as Metadata Librarians, Architectural Historians, Professors, Catalogers and Curators to name just a few. The results of this survey have given us an improved understanding of user community, institutional, and individual needs.

The responses contain valuable data regarding CDWA Lite data elements, XML schema, institutional adoption of CDWA Lite, CDWA Lite in the Ecology of Standards, and improvements that can be made to documentation.

The most difficult issue arising from the review was the fact that the XML schema that we had created was not – but just by a little bit – compliant with the CIDOC CRM. This discussion was initiated by Regine Stein and Monika Hagedorn-Saupe earlier this year and continued at the ICOM meeting in Vienna last August. As we look at the current work of RLG Programs today, the current state of thinking about the core data standard, and the work of system vendors and others who will help support this exchange model for museums, I hope the conversations keep in mind the value of moving towards one single data standard.

Building on the vision that was initiated and early work that as done at Getty and ARTstor, and included OCLC, the CRM compliant version called ‘museumdat’ critically preserves essential values: low bar for resource discovery, low burden for core cataloging, and ease of use.