Response from Jan Eklund to Janis DesMarais regarding the practice of using a relational database and/or a flat file database to manage digital assets. Quoted from VRA-L, 21 February 2008, reproduced with permission.

Janis,

This is a common question posed by administrators so here is the simple answer. For images of cultural objects, there are many arguments for using a relational model (tables linked to other tables via intersection tables) for your cataloging tool and a flat model (rows and columns like a spreadsheet) for your presentation tool, but they all boil down to basically two concepts: **complexity** and **consistency**. Data about cultural objects is often complex and this complexity cannot be captured efficiently in a flat data model because basically you have to leave space *in every record* to accommodate the most complex object you will ever encounter. This adds up to a lot of wasted space, and wasted space means more money and hardware needed for storage, backup, preservation, etc. It's much more efficient to catalog in a relational environment, where data can be entered once and then linked to many other records.

Data consistency is the other compelling reason to catalog in a relational environment. Once there is more than one person doing data entry, the potential for data inconsistency increases exponentially. Differences in opinion, spelling, and transliteration in the source material make it hard enough for one person to keep things consistent. In a relational model based on the Core categories, one *work* or *collection* record is established with appropriate links to one or more titles, dates, artist names, etc. and then individual *image* records, each representing a "view" of the work, may be linked to that one work record. This way, all the descriptive data about the work is entered once, and every image that shows this work inherits the same information. This data consistency insures that when you go to search for things, you get consistent results returned.

Management and service of the digital assets (the image files), on the other hand, is handled quite well by a flat model because the descriptive fields that apply to each *image* (file size, pixel dimensions, photographer who captured it, date it was captured, etc.) are fairly straightforward and not generally subject to scholarly debate. A lot of image metadata can even be harvested automatically by the capture device and the digital asset management tool. If the relational work data can be concatenated or "flattened" and then imported into a select number of descriptive fields in a digital asset management tool, then you have the best of both worlds: consistent, complex, descriptive metadata about works linked to multiple, individual, views of those works in an efficient discovery and access tool.

Janice Eklund
Curator of Visual Resources
History of Art Department
University of California
Berkeley, CA 94720-6020

I am researching what image management system to buy/use to serve the Visual Arts department and eventually the whole college at the College of the Holy Cross in Worcester, Massachusetts. I'm liking CONTENTdm more and more each day because of its flexibility and capabilities. What I'm finding though, is that many in this field use another software tool, other than what comes with CONTENTdm, like a designed FileMaker database, to catalog. Why is that? I'm assuming that what comes with CONTENTdm just isn't capable of all you want to do, but their documentation says that the fields can be personalized to fit VRA Core(s) and Dublin Core. People seem to do this with LUNA Insight too...I'm new to the field so if someone can shed light on this and help me understand why this is done so often I'd appreciate it.

If anyone does use CONTENTdm AND catalogs with their tools, please give me some feedback on that experience and if you'd do anything differently if you could.

Janis L. DesMarais
Visual Arts Resources Coordinator
College of the Holy Cross
Worcester, MA 01610