

Connect

Information Technology at NYU



Search This Site

Spring 2003 Edition

Read the article below, or select an option from this menu:



Current Issue



Archives



About Connect

Academic Computing Services

Using Cumulus to Manage Your Multimedia Assets

By [Philip Galanter](#), [Gary Shawver](#), and [Eric Stedfeld](#)

Once upon a time, academic work was expressed mostly in text, with the occasional illustration, graph, or chart prepared by editorial assistants. Publication meant books or journal articles, and lectures were usually enhanced with simple, hand-written notes on the blackboard or overhead transparencies. Now, university faculty not only publish papers and books, they also create websites, use PowerPoint slides to support lectures, offer students access to multimedia archives, make instructional videos, and create their own illustrations, charts, and graphs as needed.

A given faculty member might have hundreds or thousands of multimedia objects. As their personal collection of images, video clips, audio clips, flash animations, Acrobat .pdf files, maps, diagrams, and other objects grows, it becomes difficult for faculty to find and maintain the multimedia objects they need. In addition, it is beneficial to be able to track each object's source, copyright status, and usage history. Finally, there are times when sharing multimedia assets with students or other faculty is appropriate but a technical hassle.

To address these needs, NYU's Studio for Digital Projects and Research (<http://www.nyu.edu/studio/>) initiated an internal project to identify and support a productivity tool that would allow faculty members to organize, search, and retrieve their materials from a multimedia database. The project team includes Eric Stedfeld from NYU Libraries and, from ITS Academic Computing Services, project leader Philip Galanter (Arts Technology Group), Keith Adams (Faculty Technology Center), Shelly Smith (Arts Technology Group), and Gary Shawver (Humanities Computing Group).

The team's focus is on a system suited to individual faculty members or researchers, not on large, shared, enterprise-wide systems. Although such systems are of interest at NYU--indeed NYU's Digital Library effort is one provider of that functionality--what we seek to provide through this project is a personal tool that is "grassroots" rather than "top-down" oriented.

Cumulus from Canto

We began by surveying database applications with the following criteria in mind. First, we wanted a multimedia database tool with a graphical user interface that was very easy to use. We also felt it important to select a tool from a stable company with a good history of robust support, that would also run on multiple platforms, including Windows, Macintosh, and Unix variants such as Linux. Our ideal tool would be affordable yet powerful, and embrace all manner of media, including images, video, audio, Quark XPress files, PDFs, etc. In addition, even with

our focus on the individual user, we wanted to find a multimedia database application that would grow along with the user's needs, and thus operate in single-user and server-based groupware and enterprise modes. Finally, we wanted to find a tool that would allow users to merge collections or export them to other larger collections such as NYU's Digital Library.

We reviewed a number of products (see sidebar 1), and only Cumulus met all of these noted criteria (<http://www.canto.com>). Right out of the box, Cumulus is immediately useful. Cumulus is able to collect much of the information associated with a digital multimedia object automatically, making it much easier to use than traditional, general-purpose databases. Cumulus ships with a default set of media-related database fields programmed into it, a catalog that contains a subset of those fields, and a default record display which shows a subset of the cataloged fields, providing the user with information about a specific image, movie, or other multimedia object. One can simply drag a folder full of multimedia objects into the application, and Cumulus will automatically create a record for each object and, where possible, fill out fields such as creation date, file type, and image size.

Most users add additional information to their database based upon their needs and content expertise. For example, an art historian might want to digitize his slide collection, and for each slide include the artist's name, the title of the work, the date of the work, keywords describing what is being depicted, and the art movement or period with which the artist is associated. Once this multimedia database is created using Cumulus, it is then a very simple matter to search for, say, all images by Jackson Pollock, or those images identified as Cubist in style, or various combinations of criteria. Having executed the search, the images are immediately available--even if they are physically scattered throughout different folders and drives.

Metadata is Good, But Standardized Metadata is Better

The technical term for the descriptions added to the various Cumulus fields is "metadata". If your collection of multimedia is strictly personal, and you have no intent to share it or later merge it into a larger collection, you can define and enter your metadata however you see fit. Keep in mind that multimedia objects are only valuable if you are able to find them.

For this reason, and to allow for the future sharing, migration, and export of your multimedia database, some attention to metadata standards is appropriate. Metadata standards typically have two aspects. First, the metadata field names for information, e.g., artist name and file creation date, can be standardized. Second, the vocabulary used in those fields can also be standardized.

As part of our project, we developed a customized catalog to help those new to Cumulus. This catalog presents the user with the minimum set of metadata fields necessary to make their multimedia collection useful, which also provides for the future migration of their collection into departmental or NYU Digital Library collections. When the user retrieves information about an object, all of this information appears as shown in fig. #1.

SIDEBAR ONE

Alternatives for Maintaining Your Multimedia Assets

As part of this project, we looked at a number of multimedia asset management tools including those listed below. While Cumulus is the best fit for our criteria and the package our facilities (such as the Studio for Digital Projects and Research and the Faculty Technology Center) could best support, you may be interested in some of the alternatives we considered:

Content dm: <http://CONTENTdm.com>

Endeavor ENCompass:
<http://www.endinfosys.com/new/encompass.htm>

Ex Libris DigiTool:
<http://www.exlibris-usa.com/digitoolibrary/>

Extensis Portfolio: <http://www.extensis.com/portfolio>

Fotoware Fotostation: <http://www.fotoware.com>

IrfanView32: <http://www.irfanview.com>

iView Media Pro: <http://www.iviewmultimedia.com/>

Luna: <http://www.luna-imaging.com>

Madison DID: <http://cit.jmu.edu/mdidinfo/>

MetaCat:
http://www.metalogic1.com/html/metacat_about.html

Virage: <http://www.virage.com>

Record Name	Bliss.bmp
Thumbnail	
Asset Reference	Mac OS : alba hd:\Users\Shared\Pictures\XP Wallpaper\Bliss.bmp Windows : C:\Users\Shared\Pictures\XP Wallpaper\Bliss.bmp
Image Width	800 pix
Horizontal Pixels	800
Horizontal Resolution	72
Image Height	600 pix
Vertical Pixels	600
Vertical Resolution	72

Figure #1 - Partial Screenshot of Cumulus Information Display

We color-coded the information that Cumulus presents to help in entering information about a digital resource (see fig. #2). Information in blue and green is required should the collection of resources under development have a future in a digital library. This information also provides general best practice guidelines for maintaining any collection of multimedia files. Information in blue is automatically entered by the database when a multimedia file or collection of such files is dragged and dropped into the Cumulus catalog window. Information in green is what the user may be required to enter if Cumulus cannot and if it is relevant to the resource being documented. Information in black and red is either optional or part of Cumulus' internal mechanisms. As is evident in the green section, we have tried to make the database self-explanatory by making the initial value in the field provide instructions for data entry (e.g., LastName, FirstName in the Artist field, and suggestions in the Keywords field). The user replaces this initial value with an actual value. Where possible, we have provided pull-down menus with choices for easier entry.

Our suggestion is that you use the customized catalog we have developed, even if you do not anticipate sharing your multimedia assets or migrating them to another pooled collection. The catalog serves as a concrete recommendation and example of how your multimedia database should be set up, based on our collective experience. Your collection will be much easier to search and use if you follow these parameters.

Record Name	Bliss.bmp
Thumbnail	
Asset Reference	Mac OS : alba_hd/Users/Shared/Pictures/XP_Wallpaper/Bliss.bmp Windows : C:/Users/Shared/Pictures/XP_Wallpaper/Bliss.bmp
Image Width	800 pix
Horizontal Pixels	800
Horizontal Resolution	72
Image Height	600 pix
Vertical Pixels	600
Vertical Resolution	72
Compression	Uncompressed
Duration	
File Data Size	1406 KB
File Format	Bitmap Image
Color Mode	Color
Content MD5	
ICC Profile Name	
Asset Creation Date	1/16/02 1:57:49 PM
Asset Modification Date	1/16/02 1:57:49 PM
Artist	LastName, FirstName; LastName, FirstName; etc.
Composer	LastName, FirstName; LastName, FirstName; etc.
Credits	Role: LastName, FirstName; LastName, FirstName; etc.
Director	LastName, FirstName; LastName, FirstName; etc.
Performers	LastName, FirstName; LastName, FirstName; etc.
Producer	LastName, FirstName; LastName, FirstName; etc.
Writer	LastName, FirstName; LastName, FirstName; etc.
Keywords	Enter consistent values using the appropriate authority for your field.
Keyword Authority	
Keywords 2	Enter consistent values using the appropriate authority for your field.
Keyword Authority 2	
Keywords 3	Enter consistent values using the appropriate authority for your field.
Keyword Authority 3	
Notes	Enter an account of the content of your resource here.
Content Type	Enter "Image," "Sound," "Text," or "Video" here.
Genre	Enter consistent values
Genre Authority	
Source	A Reference to a resource from which the present resource is derived
Content Language	A language of the intellectual content of the resource. Use ISO Language Codes found at http://www.oasis-open.org/cover/iso639a.html .
Copyright	Enter copyright information here.
Capturer	LastName, FirstName; LastName, FirstName; etc. or OrganizationTitle
Capture Device Manufacturer	Enter manufacturer of capture device, e.g. digital camera, scanner, etc.
Capture Device Model	Enter model of capture device, e.g. digital camera, scanner, etc.
Capture Software	Enter software used to capture asset.
Caption	
Caption Writer	LastName, FirstName; LastName, FirstName; etc.
Don't Delete Record	-
Categories	Photographs
Record Creation Date	12/16/02 3:45:56 PM
Record Modification Date	12/18/02 11:46:52 AM

Figure #2 - Color-coded Information in Cumulus

In addition, we have anticipated a future possible need to export or share your collection by creating a mapping from our Cumulus custom catalog to both the international Dublin Core and local NYU Digital Library ZeroDB metadata standards (see sidebar 2). While it would be difficult to "future-proof" any digital asset collection with certainty, this customized catalog represents our best attempt to do so for personal collections.

What Comes Next?

Currently, we are working with a small set of pilot faculty projects to test our customized catalog. Once these pilot studies have been completed, we will move on to fully support Cumulus use by NYU faculty at both the Studio for Digital Projects and Research (<http://www.nyu.edu/studio/>) and the Faculty Technology Center (<http://www.nyu.edu/its/ftc.html>).

You are welcome to contact those facilities, the authors of this article, or any of the project members mentioned above with questions about Cumulus and multimedia database usage at NYU.

SIDEBAR TWO

About Metadata Standards

The customized Cumulus catalog we've created includes mapping for the international Dublin Core metadata standard and the NYU Digital Library ZeroDB metadata scheme. By using our customized catalog and filling in the required metadata, you get a good deal of standards compatibility "for free". Before you begin entering your metadata, however, you may also want to research and conform to any standardized "controlled vocabulary" used in your discipline, as explained below.

Dublin Core

The origin of the Dublin Core standard can be traced to a hallway discussion about semantics and the difficulty of finding things on the Web that took place at the 2nd International World Wide Web Conference at Chicago in October of 1994. This led to the NCSA and OCLC sponsoring a workshop to discuss metadata semantics in Dublin, Ohio, in March of 1995. There, more than fifty people participated in formulating "a core set of semantics for Web-based resources" that could be used to categorize "the Web for easier search and retrieval" (<http://dublincore.org/about/history/>). This eventually became known as "Dublin Core metadata."

Dublin Core provides a set of semantics to describe the core attributes of an object. These are expressed in its fifteen standard elements (Title, Creator, Subject, Description, Publisher, Contributor, Date, Type, Format, Identifier, Source, Language, Relation, Coverage, and Rights) plus one alternate element (Audience), and in its element refinements and element encoding schemes. The Dublin Core standard continues to evolve. It is flexible and extensible, and can co-exist with other standards. It is not intended to be exhaustive, but rather to provide basic information applicable to most, if not all, objects.

Since its inception as a semantic for describing Web-based resources, Dublin Core has grown to encompass many applications and disciplines. While its roots are in the World Wide Web and HTML, Dublin Core can be applied to databases and XML. Some of the disciplines now using Dublin Core are Archives and Museum Information and Automation Systems, Digital Information Management, Digital Libraries, Educational Institutions, Electronic Collections, Governments (Local, Regional, State, and National), Informatics, Knowledge Management, Library Technology Services, Networking, Product Research and Development, Publishing and Media Syndication, and Resource Discovery. The ITS Humanities Computing Group at NYU uses Dublin Core to encode arts and humanities data. Additional information about Dublin Core is available at <http://dublincore.org/>.

ZeroDB for NYU's Digital Library

ZeroDB is a digital asset management system being created by the Digital Library Development Team of Bobst Library. It is designed for the loading, long-term storage, and retrieval of digital objects. It is capable of handling representations of real-world assets comprised of a variety of digital formats and is structured to be compatible with the Metadata Encoding and Transmission Standard (METS), a file format for encoding complex digital objects developed by the Digital Library Federation and maintained by the Library of Congress. For locally developed collections of digital assets that are candidates for long-term preservation, using our customized Cumulus catalog will ease future accession into ZeroDB.

Controlled Vocabularies and Metadata Support at NYU

Because each discipline has its own technical language and standards of best practice, there is typically a controlled vocabulary used for metadata specific to each field of study. Depending on the discipline, the controlled vocabulary may be highly formalized, casual and open to variation, or anything in between. If you use Cumulus to organize your multimedia materials, you should consider the use of the controlled vocabulary specific to your field along with the custom catalog we can supply.

Although there are many sources of expertise on this issue at NYU, faculty are strongly encouraged to direct their questions regarding metadata to the Studio for Digital Projects and Research, which is a designated center of expertise. ITS and Library staff affiliated with the Studio can answer your questions, or refer you to an appropriate source. See <http://www.nyu.edu/studio/metadata/> for some draft guidelines on metadata, or send e-mail to digital.studio@nyu.edu

For specific information regarding controlled vocabularies, please contact Sherman Clarke, Head of Original Cataloging (sherman.clarke@nyu.edu).

Author Biographies

Philip Galanter is Associate Director of the Arts Technology Group at ITS Academic Computing Services. He can be reached at galanter@nyu.edu.

Gary Shawver is a Humanities Computing Specialist at ITS Academic Computing Services. He can be reached at gary.shawver@nyu.edu.

Eric Stedfeld is an Information Technology Specialist with the Digital Library Development team at Bobst Library. He can be reached at eric.stedfeld@nyu.edu.

Page last reviewed: April 9, 2003. All content © New York University.
Questions or comments about this site? Send e-mail to: its.connect@nyu.edu.