

Metadata Goes Mainstream

Metadata from the world of librarians and database searching is moving to center stage in our everyday lives. And the metadata “revolution” is coming to us through pictures—those cute, happy, funny shots of kids, parents, neighbors and workmates that we love to share and post on the internet.

Fred Brown



Fred specializes in indexing and information access strategies for technical documentation, intranets and web sites. Fred also

teaches indexing skills to technical writers. He started Allegro Technical Indexing in April, 2000. He has over ten years of technical communication experience working largely with high-technology firms specializing in software and telecommunications.

Fred coauthored *Index It Right! Advice from the Experts (Vol. 1)*. He is also a recipient of the Australian and New Zealand Society of Indexers (ANZSI) Web Indexing Award (<http://www.aussi.org/prizes/webindexawards.htm>) for the Writer's Block web index (<http://www.writersblock.ca/common/index.htm>).

Fred is writing an article series on information retrieval—from back-of-the-book indexes to digital libraries—with special attention to open standards and open source for KnowGenesis International Journal for Technical Communication (IJTC). Contact him at fred.brown@allegrotechindexing.com.

Photo gallery services and software allow us to organize and share hundreds of photos. And organizing those photos so that you, your family, and friends can find pictures of your week by the beach, the dog eating your hat, or the smiles of a newborn baby, requires metadata. For example, you may group all your beach shots under the heading “vacation,” annotate your dog photo with the keywords “dog” and “hat,” and add a description to that immortal picture of the latest addition to the family with “first smile for daddy!”

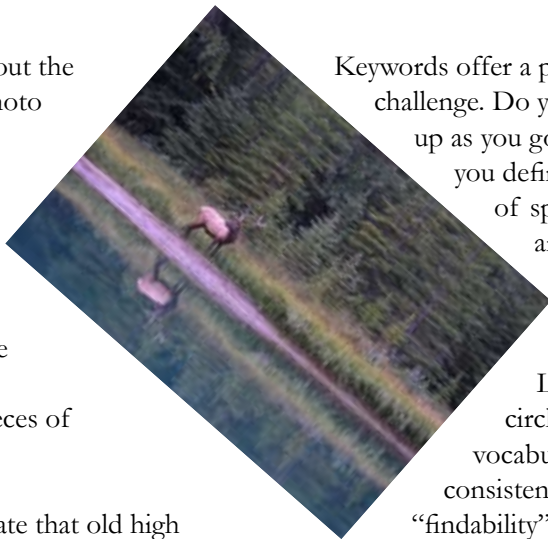
Metadata can be defined as information about a collection or items in that collection. The “items” can be topics in a content management system, drawings in an art gallery, or temporary workers in a personnel agency. Metadata allows us to manage, use, and find items within a collection. Metadata can serve a number of functions: administration, identification, and technical. Metadata can determine who has access to view or update topics in a help system. Organizing art works by period, such as “Medieval,” “Renaissance,” and “Romanticism” helps art gallery visitors locate the art works they wish to see. A database of temporary employees would list technical information about specific skills and experience in order that prospective employers can find just the right fit.

We typically organize collections of photos into albums. We need to store information about each album, such as title, description, and main subject. More specific information may be stored for each photo, such as date, keywords, and image format. The metadata enables a number of tasks:

- organizing a presentation
- finding a specific photo
- displaying a photo correctly
- describing the contents of a photo or album
- giving credit to the photographer
- controlling access to or use of a photo

The metadata can be stored in a database separately from the photos themselves. This is particularly true for metadata about albums. Other metadata, such as date and file type, can be stored within the photo itself, and is often added by the digital camera at the time the photo is taken. That way, if you move a copy of the photo from your collection to a different location, the metadata travels with it. The photo becomes self documenting. Keeping the metadata separate from the photos has the advantage of enabling the retrieval of the metadata without having to access the photos directly—particularly useful if there are restrictions on accessing the photos. On the other hand, having metadata embedded makes the photos more transportable without losing key

information about the photo. Many photo gallery services and software programs use a combination of both database storage and embedding for different pieces of metadata.



Keywords offer a particular challenge. Do you make them up as you go along? Or, do you define a “picklist” of specific keywords and choose from within that list. Picklists, also known in Library Science circles as “controlled vocabulary,” improve consistency and “findability”. If you do create a picklist and have several members of your family putting up photos plus their metadata, do you let them add new keywords to the picklist too? Or, do you keep careful control over the picklist yourself, again to ensure consistency?

through sites on the internet, too. You’ll see basic metadata, such as title, description, and date added along with subject category and keywords. Certain videos may be restricted to site members or those over 18 years of age. The site can also collect metadata from viewers, such as video rating and comments.

You may annotate that old high school scrapbook using your own particular system, but if you want to share your photos with others, it helps to use a common metadata system. The International Press Telecommunications Council (IPTC) (<http://www.iptc.org/pages/index.php>) has developed a metadata schema for photographs using the Extensible Metadata Platform (XMP).

The International Press Telecommunications Council (IPTC) (<http://www.iptc.org/pages/index.php>) has developed a metadata schema for photographs using the Extensible Metadata Platform (XMP). Metadata can be stored in XML database files or in JPEG, TIFF, or PSD photo files.

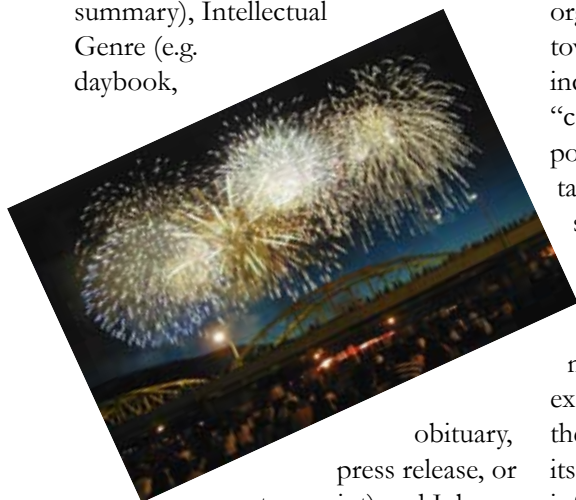
like it’s becoming “metadata madness” out there. And it doesn’t appear that it’s going to slow down any day soon.

Metadata can be stored in XML database files or in JPEG, TIFF, or PSD photo files. The IPTC standard defines four “panels” of metadata: photographer contact information; photo content; and abstract image information; and status information relating to management, work flow, and copyright. Specific metadata fields include Creator, Headline (brief publishable synopsis/summary), Intellectual Genre (e.g. daybook,

Picklists created by a group of people are sometimes called, “folksonomies” (taxonomies created by common folk). When it comes to vocabulary control, keeping it simple goes a long way, because things can get messy very quickly.

The IPTC standard has a metadata field for the Subject News Code taxonomy (<http://www.newscodes.org/>). This set of codes is oriented towards news organizations and includes terms such as “fashion,” “consumer issue,” and “water pollution.” Published subject taxonomies may or may not be suitable for your particular purpose.

The JPEG 2000 photo standard makes provision for extensive metadata, including the IPTC metadata schema just discussed. For example, JPEG 2000 has fields for the history of how the image got to its present state, camera and software information, focal plane resolution, exposure time, and subject distance. We can now share home videos



obituary, press release, or transcript) and Job Identifier.