Developing a successful metadata schema

Chris Porter

is the Director of Search Data at Getty Images, the world's leading provider of visual imagery. Chris is responsible for developing scalable metadata schemas and the keywording of all Getty Images' assets to support the customer web search experience.

Keywords: *metadata model, business decision, customer, consistency, communication, change*

Abstract Metadata is an important and highly valuable part of any digital asset management (DAM) system. The metadata turns a digital file into an asset by allowing the digital file to be retrieved from the system, shared between users and reused by the business. Before you start to define the metadata model for your DAM system there are several key business decisions that need to be understood, such as the purpose of your DAM system and the amount of time and money you are willing to spend on metadata capture. The timing of the model definition is critical to the overall success of the DAM system deployment and the model should continue to meet the changing needs of the business-users. What is the best way to create a sustainable metadata model?

THE VALUE OF METADATA

When digital asset management (DAM) systems were first developed, early users of DAM were primarily focused on ingesting the digital files into the system. Users did not seem too worried about how they were going to retrieve the assets from the system in order to realize their value to the business. But over the past several years there has been a notable shift and users now understand that the right metadata can turn a digital file into a reuseable asset. Identifying exactly which metadata will be useful is a daunting task as mistakes can be expensive. So here are some guidelines that will hopefully demystify metadata definition and help you define a successful metadata model that will meet the current and changing needs of your business.

Metadata is the descriptive information used to identify a digital file within a DAM system. Metadata turns a digital file into a digital asset by allowing it to be retrieved from the DAM system, shared between users and reused by the business. Metadata is usually held in specific fields and includes data such as a unique asset identification number, a caption, rights and permissions data and keywords. A metadata schema or model is a consistent set of data fields used to describe all your digital files and it should support the business purpose of your DAM system.

SOME KEY DECISIONS

When the metadata schema has been defined, the capture of the data for each digital file can be a time-consuming and expensive task. It is therefore important

Chris Porter
Director of Search Data
Getty Images
601 N.34th Street
Seattle, WA 98109, USA
Tel: +1 206 925 7122
Email: chris.porter@gettyimages.com

to make some key decisions at the very beginning of the project to guide the scope of your schema. First, you should be clear about the purpose of your DAM system. If you are using DAM just to track assets between departments then your metadata model could be very simple. If your need is to supply data to a customer-facing website to drive search and retrieval then it is likely that your model will be complex and perhaps even beyond the scope of a generic DAM product.

How much money are you willing to spend on capturing the metadata for your digital files? It is an inherently time-consuming process and there will be the temptation to get sucked into the pure science of consistent and accurate metadata. It needs to be understood that there will be business constraints such as the time per image, and decisions will need to be made within those requirements.

Should every digital file get the same metadata applied? Do you have some assets that are more valuable to you than others? Don't build a complex and expensive model for low-value content. You can have one metadata model which can encompass all levels of content, but with a small subset of fields appropriate for low-value content.

It is important to map out the ingestion workflow into the DAM system and to identify the correct stage at which the metadata will be assigned to the digital files. It might be necessary to have different pieces of data assigned at different stages in the workflow and by different groups of users. Be careful of designing a workflow to accommodate every known exception to the normal process. This is where the project scope starts to get out of control and the system

ends up being over-designed to compensate for failures in the business process. If you are going to do any process reengineering, make sure it is carried out and signed off prior to starting the DAM project.

The metadata schema should be developed very early on in the project life cycle. The early timing of the schema development is a key success factor. Always develop the metadata schema and have metadata applied before any content is loaded into the DAM tool.

PUT THE USER FIRST

So, now we have covered those business decisions, how do we start gathering the requirements for the metadata schema? The best place to start is with the user of the DAM system. A 'user' is anybody who interacts with the system, whether they are referencing or creating the metadata. It is important to spend time gathering the user specifications to make sure that each piece of required data has been documented. In addition, if your workflow spans several departments, get key people together early in the project to determine consensus on the nomenclature of the data fields and what they should contain. That way an asset ID, for example, means the same to everyone.

It may be that your user group and workflow span different countries now or in the future. If you can have an international team working together from the beginning you will avoid the problems of trying to retro-fit a tool designed exclusively for English-speaking users.

As you work through the project, communication with the users is essential to keep them up-to-date with key decisions and to allow them to feel part of

the project. Typically a DAM implementation is a long project, so regular communication about the project status is important. Within organizations, some departments are more influential than others and you need to ensure that decisions are not weighted towards the needs of one group.

APPLYING THE METADATA

By now you should have a good idea of how many metadata fields you are going to create, what each field should be called and the type of data in each field. Now would be a good time to take a reality check. If your metadata schema is going to be unnecessarily complex then it will be expensive to ingest new content into the DAM system. In addition, significant training will need to be given to employees. Training is an important issue because if it takes a month to train a new user then it will be difficult to respond quickly to changing business needs.

Once the metadata fields are defined then you should develop a method of consistent application of data into the fields. This is so you can have several different people in different locations all adding data in the same way. Make sure everything is documented as you go along so that training materials are easy to pull together.

You should give thought to a quality control process and whether it would be a valuable step in your workflow. As part of this decision you need to decide what parameters are going to be important. For example, if you are only using the DAM system to transfer assets between departments, then if a caption has a typographic error will it negatively impact the user?

Several metadata standards have been

developed by independent organizations such as IPTC (International Press and Telecommunications Council) or Dublin Core. You might find these standards a useful place to start and to prevent you from reinventing the wheel. In addition using IPTC or Adobe's Extensible Metadata Platform (XMP) you can imbed the key pieces of metadata into the digital file so that the metadata travels with the digital file. This is particularly important if you intend to exchange digital files with external parties.

To save time with the creation of the metadata you should look to leverage data that is available when the digital file is ingested into the DAM system. If the photographer or artist supplies any information with the digital file then it should be used where possible. In addition, most DAM systems can identify certain pieces of metadata and automatically apply it to the digital files as they are ingested into the DAM system, for example, the aspect orientation of the image.

When building the user interface for the metadata creation it is a good policy to design the tools to make them easy and intuitive to use so that you reduce the required training and knowledge transfer.

CONTENT MANAGEMENT

Another key function of the metadata schema is the management of the assets within the DAM system. A good metadata model should allow you to make assets inactive in the system and to indicate why an asset is no longer useable by the business. The DAM system should be able to track the assets as they move through the workflow and to allow activity reports to be generated to support key performance indicators.

TRANSITION AND CHANGE

Business transition planning plays a very large part in the implementation of the new metadata model. Changes to working practices need to be rolled out in a controlled fashion to minimize the impact to the business and to ensure that all users are ready for the new system.

It is important to embrace the concept of change as you are defining your metadata schema. Because DAM projects have a long lifecycle it is realistic for the business needs to shift during the project cycle. You need to be responsive to those changes and adapt the model accordingly. In addition, as your business evolves after the deployment of the DAM system, the metadata model needs to have been built in such a way that it can support the

new business direction. A key feature of a metadata capture interface is the ability to change one or multiple fields for a large number of images simultaneously, these “bulk” updates will help to modify metadata records.

SUMMARY

It is apparent that the key to a successful metadata schema does not lie in the actual definition of the individual data fields — that would be far too straightforward! The success comes from the right business decisions at the start of the project, the early definition of the metadata schema, the level of collaboration between the system users and the project team, comprehensive documentation and training, planning for the system implementation and embracing change as part of the process.