

Summary of Comments on King_Seybold_Final.pdf

Page: 1

Sequence number: 1

Author: jking

Date: 9/1/2004 7:42:35 PM

Type: Note

Thanks to the Seybold organization for allowing me to present at their San Francisco conference in August 2004.

And thank you for looking at my presentation. I have annotated each page with a comment like this so that you can read roughly the same words as I would speak while giving this talk in person. Just mouse over the little yellow icon or double click on it to open it.

You might want to print out the comments for reading alongside of your display so you do not have to fool with them on the display. If you own Acrobat Standard or Pro you can get a copy to print out using Document->Summarize comments.

Jim King -- 8/31/2004

Page: 2

Sequence number: 1

Author: jking

Date: 9/1/2004 5:56:34 PM

Type: Note

The Seybold Conferences have always been a major venue for printers, publishers and designer. And Adobe has been an important part of that. It is a pleasure to be here as we continue move on to new paradigms and new challenges.

Success in design and publishing used to be primarily about the work; Today it's about the work... and the workflow.

Page: 3

Sequence number: 1

Author: jking

Date: 9/1/2004 6:06:56 PM

Type: Note

We have spent many years concentrating on tools to increase individual productivity. Now it is time to also turn to tools to increase the group productivity and to help people work together more easily.

Page: 4

Sequence number: 1

Author: jking

Date: 9/1/2004 6:07:13 PM

Type: Note

But why are we now more interested in workflow and collaboration?

Economic pressures are demanding greater efficiencies, streamlined ways to develop jobs, and to accommodate last-minutes changes

Repurposing rich media content requires that skill sets extend across a variety of mediums

More information; more people involved; more versions generated; more platforms to consider

How is this manifest day-to-day?

More time spent searching for and finding assets.

More files because of shorter run jobs or because components have been broken out for re-use.

More collaboration and sharing of assets.

Greater need to define, protect, add value to and archive assets.

There is just a lot more of everything.

Page: 5

Sequence number: 1

Author: jking

Date: 9/1/2004 6:07:20 PM

Type: Note

Here is a picture that a friend of mine claims he recently purchased for \$500.00. It looks a lot like what my children would product with crayons when they were in about the 4th grade.

I think my friend spent too much money.

Page: 6

Sequence number: 1

Author: jking

Date: 9/1/2004 6:07:32 PM

Type: Note

Well most of you probably knew that this is a famous painting done by Picasso. It is certainly worth orders of magnitude more than \$500.

But if you did believed my story about my friend, isn't it amazing how this one little signature or piece of extra information, can, in and of itself, make the same picture worth so much more.

This is also true for our computer files. Sometimes knowing just a little bit more about the files can make them much more valuable.

Page: 7

Sequence number: 1

Author: jking

Date: 9/1/2004 6:08:52 PM

Type: Note

What we are going to spend a lot of time talking about today is "metadata". Data about data. The notion of adding a creation date, or an author or photographer's name to a digital asset has been with us for a very long time. But now it is taking on much increased importance because of the increased volume of files we have to deal with and the wider variety of things we must do with them. We cannot spend the better part of our days running around trying to find things or find out about things.

Page: 8

Sequence number: 1

Author: jking

Date: 9/1/2004 6:09:42 PM

Type: Note

We have learned to make heavy use of little pieces of information like the words found on road signs. This isn't exactly metadata but these simple labels play that role in our everyday lives.

Page: 9

Sequence number: 1

Author: jking

Date: 9/1/2004 6:09:52 PM

Type: Note

Lose a little, lose a lot!

Page: 10

Sequence number: 1

Author: jking

Date: 9/1/2004 6:14:27 PM

Type: Note

Getting back to the title of the talk I find that metaphors and analogies are often helpful to understanding provided we do not take them too far. I find it interesting to compare how I organize information in my physical office and how I organize information in my computer. If I have, say, a technical report in paper form and I want to save it to refer to later, I usually have

two choices. I can file it in a particular filing cabinet in a particular folder. If I am well organized I will be able to remember which cabinet and which folder in a month or two when I come back to retrieve it.

More often, I cannot determine which cabinet and folder so I just set it down on my desk usually into a pile of material that has suffered a similar fate.

Finding things in those piles is often pretty efficient. They are usually sorted chronologically so if I remember about the time I received it I can make a good guess as to where in the pile it might be or which pile it might be in.

I have thought about these two choices and have concluded that it really depends upon how many times I will try to retrieve a given item. If I will be retrieving it often and I need it quickly then filing it in an established place (cabinet and folder) is the best choice. It may take a little longer to put away the first time, but I gain that back during the quick retrievals. On the other hand, if the chances that I will really want to look at something even once again, then just dropping it in a pile is most efficient. I only pay the time, if and when I go to find it.

The organized filing system fails me usually for two reasons. I do not remember into which cabinet and folder I put something, and I often need to put something away under two or three categories. For example, if I get a note from a friend, John, who works for Apple, and asks to have lunch with him on Thursday, I would like to file it under John, Apple and Meetings at Seybold. I often actually copy computer files so that I can file them in multiple places. But then I update one and not the others.

When multiple people are accessing same filing system these things compound. We must have a well established taxonomy for cataloging that we all agree to and understand.

Of course, these same asset archiving issues arise inside my computer.

Sequence number: 2

Author: jking

Date: 9/1/2004 6:15:47 PM

Type: Note

Computers should be better at searching through piles of things to find what I want. Right?

Well for text documents that has been proven to be true. We can index a pile of documents and then very quickly find each of the documents that contain a certain phrase or word.

The problems arise when I have non-text documents. What properties can I search on. In stacks of paper I have some visual cues like paper color, binding, size, shape, etc.

We can use some of these same properties with our computer files: date created/modified, author, size, file-type. But to do this really well we need to turn to adding more metadata to those files. If we can get the right indexing information added to all of our files, especially the non-text files then we can unleash the great power of the computer to find things for us quickly.

So this is my view of how it all comes down to metadata, metadata, metadata.

Page: 11

Sequence number: 1

Author: jking

Date: 9/1/2004 6:16:28 PM

Type: Note

OK. I have attempted to give you some insight into why metadata is so important and necessary in today's new world.

So what are the challenges of obtaining and effectively using metadata?

Page: 12

Sequence number: 1

Author: jking

Date: 9/1/2004 6:16:42 PM

Type: Note

The first challenge is establishing universal standards.

Suppose that I am working on a print piece for the Seybold organization that I have been calling the "Seybold Agenda".

If I add this title as metadata using the IPTC (International Press Telecommunications Council) standards for metadata I will call it the "description" of the piece. If I add the metadata using the DISC (Data Image Submission Criteria) standard it is to be called the "caption". And I am sure that other standards have yet other labels for this particular piece of metadata.

If I want to reliably search through a large set of diverse assets I need to have these simple labels standardized and across standards organizations.

Everyone like standards, especially if the one that is adopted is theirs.

Page: 13

Sequence number: 1

Author: jking

Date: 9/1/2004 6:19:43 PM

Type: Note

At Adobe we have analyzed the standards situation and concluded an effective metadata standard needs to have 5 properties as shown.

Open-- It needs to be open itself and built upon other open standards.

Extensible -- The universe of useful metadata will never be a closed book. New needs will constantly be arising so we need to be able to support extensions over time.

Universal --This refers to the need for the same metadata standard to be supported widely in applications and file formats. Having specialized pockets of standards is not really having a standard.

International -- This is usually nicely satisfied by having the material expressed using Unicode and XML and by making sure that multiple languages are supported.

Robust -- Suppose that I have three applications in a workflow. Work starts at A moves to B and then to C. Suppose A and C need and work with the metadata but in B metadata is not needed, not used. That does not mean that the application B does not have to be aware of and process metadata. It must maintain the metadata obtained from A and pass it on to C. So all applications, platforms and file formats have to actively support metadata preservation and handling even if they do not process it themselves.

Page: 14

Sequence number: 1

Author: jking

Date: 9/1/2004 6:41:00 PM

Type: Note

So the first challenge is establishing far ranging standards.
The second challenge is capturing the metadata in the first place.
Of course the most straightforward way to get metadata is to have each of us add it manually to our assets (files).
Perhaps we should stage "Metadata Mondays" where we all sit at our computers and enter the metadata to all of the files we have created in the last week.

Page: 15

Sequence number: 1

Author: jking

Date: 9/1/2004 6:42:31 PM

Type: Note

There are things that we are doing to avoid Metadata Mondays.
We can add metadata automatically, at least certain kinds of metadata that are normally and automatically available to our applications. Things like: dates, file size, and file-type. But going further, within digital cameras we are adding things like the camera brand and lens settings and even the GPS coordinates of where the camera was, when the picture was taken.
When we put our minds to it, we can find more and more ways to harvest valuable metadata from our environment and do it automatically. For example, in one of Adobe's back rooms we have people looking at facial recognition. I might label one of the pictures of my son and the computer might find and label all other pictures that also include his image. Well at least it could bring up candidates for me to OK. And then that metadata (my son's name) will be added to all those pictures automatically.
And the earlier in the process that we add the metadata the better. It will be more accurate and more workflow steps will have it available to them.

Page: 16

Sequence number: 1

Author: jking

Date: 9/1/2004 6:44:45 PM

Type: Note

Again, we have to stretch our imaginations to find more sources of automatic metadata.
Another possible source of valuable metadata can be provided by the context in which I am working. Suppose that I am using a workflow product that assigns me jobs and tracks the time I spend on each job for accounting purposes.
Then perhaps I can label each asset I touch during that period with that project label. So if I am now working on the Seybold project and the computer "knows" it, then it can add a client field of "Seybold" to my metadata on all the files I create.
What other contextual information might we be able to take advantage of?

Page: 17

Sequence number: 1

Author: jking

Date: 9/1/2004 6:47:22 PM

Type: Note

So the two challenges to metadata are:

(1) standards and (2) capturing the metadata in the first place.

Let's see what tools and applications we can apply to these problems.

Of course, the talk is going to turn a little more Adobe centric at this point. These are the tools and applications that I know the most about.

Page: 18

Sequence number: 1

Author: jking

Date: 9/1/2004 6:50:56 PM

Type: Note

Almost all modern metadata standard proposals recommend using XML to hold the metadata. For the last 3-4 years I have been on a personal campaign to convince people that just saying something is XML does not convey much information and is not very helpful.

The reason is that, even though the name XML stands for "Extensible Mark-up Language," XML is NOT a mark-up language. Rather, it is a set of rules and tools for creating mark-up languages. Very specialized and powerful mark-up languages -- thousands of them, customized for special uses.

It is a little like saying that I have written my report using the Roman alphabet. I guess that can be true but it is not very interesting. However, if I tell you that my report is written in American English and uses the Adobe standard technical report template, that is much more useful information.

So, I have a rule that I have imposed within Adobe and I hope to impose it on you, now. You may not use the term "XML" by itself anymore. You must always use the phrase "XML for ...". Like saying I am using the Roman alphabet for expressing English poetry.

This makes for much more meaningful conversations and exchange of information. Indulge me, please!

Page: 19

Sequence number: 1

Author: jking

Date: 9/1/2004 6:52:55 PM

Type: Note

As I noted earlier, the use of metadata has been with us for a long time, both with everyday things and in our computers. So there are already a lot of standards.

One that is of increasing interest is the JDF or rather "XML for Job Definition Format." This is an XML mark-up language for recording all kinds of metadata about a print job from start to finish.

As most of the material we handle comes to be represented digitally, so must the famous job

jackets that many of us are so familiar with.

So you can think of a JDF file as a digital job jacket.

At the DRUPA conference last Spring, Adobe demonstrated some technology for creating and tracking JDF in Creative Suite workflows. Stay tuned!

Page: 20

Sequence number: 1

Author: jking

Date: 9/1/2004 6:53:54 PM

Type: Note

Another XML based standard is RDF or Resource Description Framework. This one falls more into the rules and tools category, laying out some standards to use when creating XML mark-up languages dealing with resources and their interrelationships.

It is a W3C recommendation and was targeted for adding semantic information to the Web.

We bring this up because this is an important standard to use in subsequent XML based metadata standards.

Page: 21

Sequence number: 1

Author: jking

Date: 9/1/2004 7:21:42 PM

Type: Note

Adobe has created a standard XML mark-up language for metadata and a way in which to package that metadata into arbitrary file formats in an unobtrusive manner. The whole business is called the Extensible Metadata Platform (XMP). More on the platform part later.

But now I want to talk a bit about the XMP metadata mark-up language itself.

We have tried to apply the 5 properties that we determined were required for a standard.

We based XMP upon the rules of RDF and the list of labels established as the Dublin Core vocabulary. We are building upon existing standards.

XMP standardizes on a vocabulary that covers most of the things for which we need metadata, including the obvious, creation date, creator name, description of the work, etc.

And we have made provisions so that when new vocabulary terms are needed they can be neatly added to the standard.

We want everyone, especially all the application developers in this industry to standardize on XMP. And we need your help to spread the word.

Page: 22

Sequence number: 1

Author: jking

Date: 9/1/2004 7:25:31 PM

Type: Note

We have done a couple of things that we feel are pretty clever and should make it much easier for others to agree to use this standard.

First is the way we recommend that XMP packets be included into files. Of course, you can have just an XMP XML file of metadata that contains the metadata for some other file, but we all know that that is a hard situation to control. We forget to keep the two files together and we lose the metadata.

We do support that "side-car" approach, but the better way to do it is to include the XMP packet inside the file to which it belongs. We have chosen a strange string of characters that are extremely unlikely to appear as a string in any file you own. That string is shown here as `id='W5MOMp ... '`.

The XMP packets can be put anywhere into a file format that is convenient for that format, but they are not to be compressed or encoded in any way. Then I can write a program to search the file for the unusual string of characters and when I find it, it is very likely that I have found an XMP packet. I can then read the packet and process it as I might any XML file.

The beauty of this way of including the XMP into file formats, is that I can write one single program to search for and find XMP packets, and that program will operate on any file.

Adobe has included XMP packets of metadata in all of our standard file formats including: PDF, PSD (Photoshop), ai (Illustrator), in (Indesign), etc. Any document management system that wants to access the metadata in any of these file formats does not have to learn anything about those formats but just has to have the standard XMP search and process program. And Adobe provides a Software Developers Kit (SDK) to help others develop these programs and to incorporate them into their workflow products.

Page: 23

Sequence number: 1

Author: jking

Date: 9/1/2004 7:30:23 PM

Type: Note

Everyone loves standards as long as it is their own.

Well we love XMP and want everyone else to convert to using that as the sole standard. And as I said earlier, we need your help to evangelize this single standard.

But we have done one more thing to make its adoption by other vendors and standards groups more palatable.

In Adobe's Creative Suite products we have a standardized tool for displaying metadata and that tool can be customized by the customers themselves.

Here we have an example where in XMP the name "John Doe" is labelled as the "creator" and the title for this piece of work is "Seybold Agenda" and in XMP that is labelled "subject". So "creator" and "subject" are two of the XMP vocabulary names that we want everyone to standardize upon. However the IPTC has long ago standardized on the labels "author" and "description" for these two items. And the DISC has standardized on terms more meaningful to photographers of "photographer" and "caption" for these same items.

Our concession to them is that they can create UI Windows with which to view the standardized XMP metadata using the terminology they have already standardized upon. There are Custom Panel Description Files that are easily created and define a mapping between the single XMP standard terminology and a more specialized terminology that a industry group or even a project team might chose.

So the data remains standardized but the views of it can be customized.

Page: 24

Sequence number: 1

Author: jking

Date: 9/1/2004 7:31:15 PM

Type: Note

We are aggressively working with trade associations, and standards groups to reconcile their existing standards and XMP. And most that we have worked with are anxious and willing to adopt XMP as their base standard.

We have created and ship with the Creative Suite several standard Custom Panel Description Files that capture the terminology in use by these groups and sanctioned by them. Their users can then view and edit the metadata using the terminology they are used to.

Page: 25

Sequence number: 1

Author: jking

Date: 9/1/2004 7:32:08 PM

Type: Note

What you are about to see now is how we are putting metadata to work in innovative ways to deliver a platform for the future of design and publishing.

The Adobe Creative suite features the standardized metadata framework (XMP) and uses an application called Version Cue as a convenient file versioning tool.

Page: 26

Sequence number: 1

Author: jking

Date: 9/1/2004 7:34:46 PM

Type: Note

Gunar did a fine demo, too bad we cannot reproduce that in these notes as well.

So I hope I have been able to show you why there is such an increased emphasis on workflow and metadata--how metadata is a key to effective workflows.

Now for a little philosophizing. I have been working with computers for longer than I want to admit. And I have been a witness to the computerization of many industries. We have been living through a computerization of the printing and publishing industry for the last 15-20 years.

I think that computerizing takes place in three steps.

Page: 27

Sequence number: 1

Author: jking

Date: 9/1/2004 7:35:30 PM

Type: Note

In the first step we provide computers to individuals together with appropriate software and attempt to improve that individuals work.

But we do not change the job that person is to perform nor do we change how that person interacts with others.

Page: 28

Sequence number: 1

Author: jking

Date: 9/1/2004 7:36:22 PM

Type: Note

The second step is to take the collection of computers that we have provide to the individuals and connect them together electronically.

Here we find some great savings because we can move many of our paper artifacts onto the computers and pass them around on the network. We can set up common file servers and make heavy use of e-mail.

But as in step one, at this stage the individual jobs are essentially the same, the relationships with others are the same and the workflow is the same.

Page: 29

Sequence number: 1

Author: jking

Date: 9/1/2004 7:38:29 PM

Type: Note

I think our industry is moving from the second stage to the third stage. We have the computers and the networks.

Once we begin to see the real power of a set of networked computers and become more familiar with working with electronic artifacts, we are ready to move on to the third and final stage.

In this stage we throw everything up in the air for review, including the individual job descriptions, the workflow, what is done when and how. We find that two jobs can be combined into one, freeing up one of those people to go attack problems that we have been holding off working on.

This is the stage where very significant economies can be realized because we are finally moving to take maximum advantage of the electronic environment we have created.

And key to this step: if your competitor finds new workflows that offer big savings and you don't

Page: 30

Sequence number: 1

Author: jking

Date: 9/1/2004 7:40:10 PM

Type: Note

This is a brave, new, multi-dimensional world, with assets coming and going, last minute changes, versions of versions of versions, and re-purposing the is order of the day — not to mention increased demands for sharing and collaboration.

One key way of addressing all these issues and managing your business and its sometimes hidden assets better is through metadata. And I'd like to think that Adobe is helping show the way with tools like XMP, Version Cue, and integration of both into the Creative Suite.

In a rapidly changing world where the unimpeded and productive flow of work is almost as important as the work itself, knowledge management is the hidden strength that designers and publishers are coming to depend on to add value to their businesses by extracting the maximum value from the knowledge assets they possess. Metadata is the "management tool" that makes it happen.

Thank you and enjoy the rest of the conference.

Jim King 9/1/2004