

# Thought Leaders Hammer Out Metadata Standard

By Ron Roszkiewicz

Organizers described the Content Metadata Summit 1.1 as “an intimate interactive gathering of senior, metadata-savvy content technology thought leaders to debate and prioritize industry requirements.” Their goal is “to build a foundation for the XMP-Open industry initiative.”

Open standards are tricky. What is gained through openness and collaboration can be lost through inconsistency and shifting priorities. Without centralized control, there can be a lack of focus. Standards committees can be passive, rubber-stamp operations and contentious, self-serving political bodies each trying to make sure their interests are served, with less regard for the good of the industry. At least that was what professionals involved in such initiatives have told me over the years.

With that as context and background, I was not sure what I would find at my first Content Metadata Summit 1.1 in New York. The summit leaders for the event were Andrew Salop, ex-Adobe senior product manager for DAM and content management systems and founder and first evangelist at the company for the extensible metadata platform (XMP), and David Steinhardt, president of IDEAlliance and the lead for this working group. Salop currently runs Metaseed.net, a consulting company specializing in the framing and marketing of digital content lifecycle management technologies.

The organizers described the event as “an intimate interactive gathering of senior, metadata-savvy content technology thought leaders to debate and prioritize industry requirements.” The goal of their activities is “to build a foundation for the XMP-Open industry initiative.”

Participating organizations on the customer side included American Express, Capps Digital, Hachette Filipachi, Leo Burnett, McCann Worldgroup, Ogilvy & Mather, Pearson, Showtime Networks, TBWA, TransContinental and Turner Broadcasting. On the developer side were ClearStory, EMC, Industrial Color, Interwoven, MediaBeacon, Pound Hill and Snap.com.

This meeting was a followup to a Jan. 11 summit in San Francisco that included customers Boeing, Landor Associates, Leo Burnett, Mattel, McCann Worldgroup, Showtime Networks, Transcontinental and Warner Bros., and developers ClearStory, EMC, Industrial Color and MediaBeacon. As such there were some conference calls conducted in the interim as a result of action items from the first meeting.

As a metadata true believer and someone who’s written about and developed tools for this market, I have some investment in the success or failure of the XMP

platform. Reviews of XMP and off-the-record feedback has, by and large, been positive, but always with qualifications. XMP is an untested technology with identifiable missing pieces, and the push has always been back at Adobe to complete the specification and satisfy the market need before expecting others to adopt it. Since I’ve been very critical of Adobe’s lukewarm support of this fledgling technology in the recent past, I was curious whether Adobe was trying to make progress or put the baby up for adoption.

Before diving into the XMP labyrinth, kudos are in order to the co-leads for organizing the event and controlling its pace. Spending an entire day listening to and talking about any single topic — even one as stimulating as metadata — is not for everyone. For the most part, the time was well used and the host for the event, Industrial Color, was supportive and involved.

## How Open Is Open

In terms of openness, Adobe would not simply hand over the source code and let the developer community have at it. Rather, it’s more a process where Adobe would allow the core code to be extended in much the same way that PDF has been extended through the /X1, /X2 and /X3 initiatives for prepress. Part of the process will be carried out on a day-to-day basis by IDEAlliance as technical and business working groups hash out issues and make suggestions

---

**For Adobe and XMP, “open” primarily means open to suggestions, with an official channel to process them.**

---

or engineering change orders are discussed before being executed in the code by Adobe. As the developer of a suite of applications that relies on XMP as the vehicle for managing metadata, Adobe has too much invested in its development to allow any substantive changes by outsiders. So “open” primarily will mean open to suggestions, with an official channel in place to process them.

As the principal conduit to Adobe for changes to XMP, IDEAlliance will act as a gateway and support

organization to the user community — a role for which it is well-suited. IDEAlliance's long-term support for XML technology, standards and conferences has earned it much respect. As a sponsor-supported, not-for-profit organization, IDEAlliance can serve as a credible buffer for Adobe to the user community and synchronize and standardize third-party development efforts. The principal unanswered questions at this point are: Will the stakeholders represent all of the key industries; will Adobe provide timely support for considering user input and updating the XMP Toolkit; and can Adobe, IDEAlliance and IDEAlliance workgroups manage all of the responsibilities that will fall upon them when the deal is struck. The hand-over doesn't seem to have taken place yet, and we are still examining the scope and feasibility of the proposal.

In his thoughtful presentation, David McAllister, director of Platform Standards at Adobe, defined some of the various forms openness can take. An open standard is based on:

- publicly available specifications;
- description is complete enough to allow independent implementations to interoperate;
- no business restrictions; and
- interested parties can participate in the creation of the standard.

For the most part, Adobe has satisfied all of these requirements in keeping with its precedent with other standards of releasing but still controlling.

McAllister also explained the degrees of openness:

- API, SDK: As long as the interfaces are standard (or published), the toolkit can be considered open.
- Published specifications: read-only and read-write.
- Formal Standards Organization: voluntary and regulatory.
- Open Source (in some senses, open source can be the opposite of a regulatory standard): licensing and transparency.

According to McAllister's explanation and the way the Adobe/IDEAlliance collaboration is developing, most of the options mentioned in his lists are covered. The one that's not obvious is regulatory, and while it seems unnecessary for control over the XMP platform standard, it could be necessary to manage schema and related data, as explained later in this article.

## It Takes a Village

In a previous article in TSR ("The Brief Tortured Life of XMP," Volume 5, No.10, Aug. 17, 2005), we said it was sad to see such a precocious technology receiving such lukewarm support. We cited a lack of developer support, inconsistent and flawed support in Adobe's products and low adoption rates. The news has gotten better since then. First of all, Adobe has consistently and comprehensively expanded its support for XMP in its own products in CS3. Second, Adobe is sponsoring the IDEAlliance effort in an attempt to find out as much as possible about where and how XMP is being used in order to make it a better technology. Third, Adobe states that it is involving top XMP and standards people and spending many dollars in this development and engagement.

As far as participants are concerned, the "village" is missing some key inhabitants. While end users are extremely important for drawing up user cases and making the business case for a technology and an implementation, engineers must be involved to implement it in code so it can be used. Absent from this meeting were Apple, Microsoft, Google, Novell, IBM, many DAM vendors (MediaBeacon, ClearStory, Interwoven and EMC were there), CMS vendors and many others that are extremely important for making the toolkit something they can adopt and use in their solutions. These companies must be actively involved.

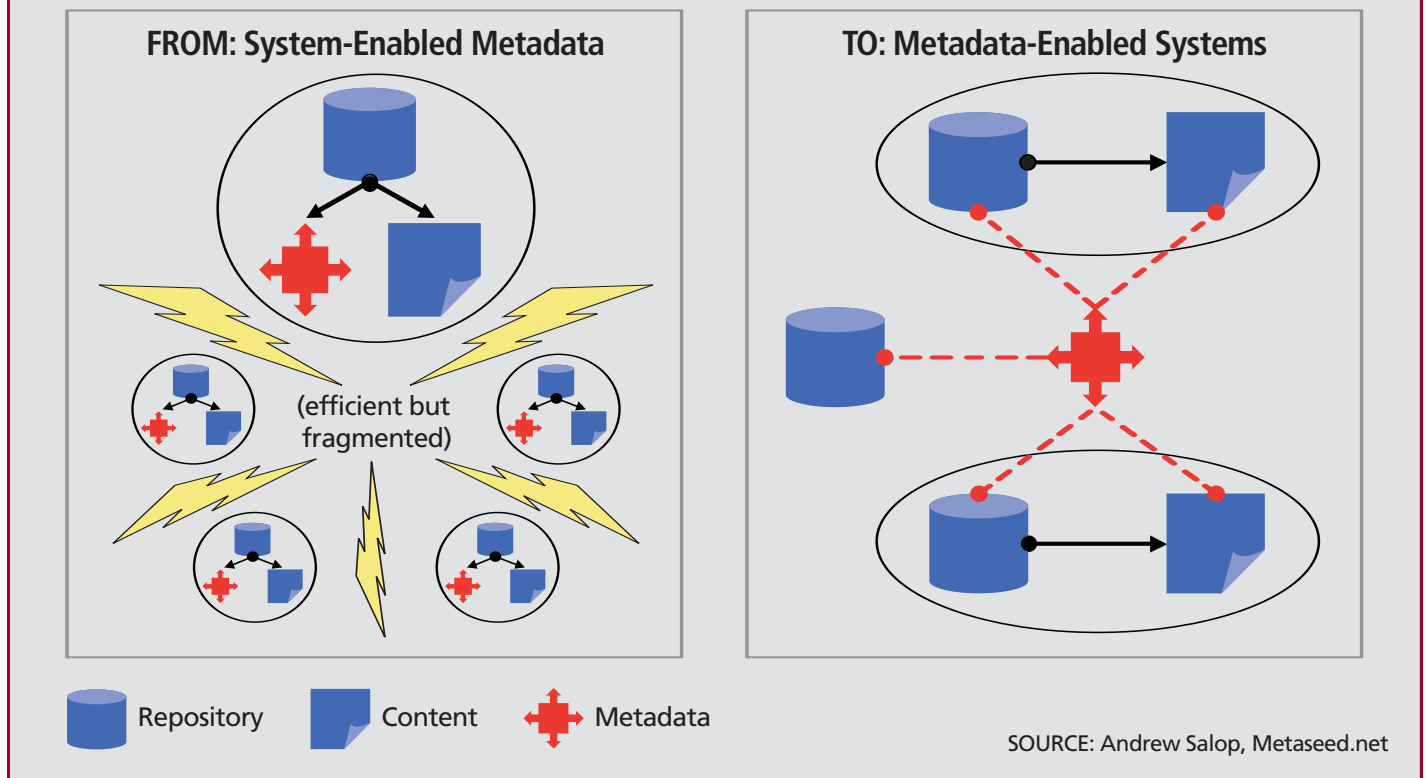
In the division of responsibilities between Adobe and the consortia (IDEAlliance and the stakeholders), Adobe will retain control over the data model, copyright, XMP brand, current SDK, basic XMP library support, FLV (motion metadata) and SWF (Flash file format). The consortia will be responsible for advising Adobe on alternative-language SDKs, compliance tests, new file type enablement, operative system file stream enablement, side-car enablement and extensions to the data model. The consortia will have ownership responsibility for name space registry, cross-map definition and support, standards group outreach and coordination, compliance certification and logo and the "XMP Open" brand.

## Representing for Standards

Quite a few standards have been proposed and supported over the past 20 years. In the graphic arts and publishing world, many of these specs have come from Adobe and have evolved through the efforts of standards bodies. Two that spring to mind are the tagged image file format (TIFF) and the portable document format (PDF). Like XMP, TIFF requires a "schema" or collection of tags and values. There have been extensions over the years to support color, for example. The core code was so right on the mark from the start that the spec for it has survived since 1992, with changes to support new technology. During the days of TIFF standards development, many of those involved were professional standards participants. It was all they did for their company and it required a combination of technical and political savvy to pull it off. The mix of participants today is much broader. End-users are typically added and consultants participate on behalf of their clients. Ultimately, Adobe manages and maintains the TIFF standard.

The XMP-Open committee seeks to balance the business case and technical issues to arrive at a comprehensive standard. For example, everyone agrees that there should be a way to embed rights information in the XMP data and also a way to protect embedded data through some form of encryption. The business cases for these are easy to make. Both involve technology that does not exist in the current XMP specification. The business cases to support the manner of implementing rights information or types of security are not complete, however. The rights needs of a multilingual publisher are certain to be different from the rights of a single-language publisher,

## Extending Metadata Beyond the Context of a Database



**Evolution** Even the most well-defined standards are not static; they evolve through an ongoing process of filtering and forming at the hands of the industry they serve. Adobe's XMP emerged from the lab in 2001 and was tossed into the market as built-in functionality for Adobe Creative Suite applications and as a developer's kit. The process to help this technology evolve was not in place at the time and industry participation was muted. Six years after the introduction of the technology, a process supported by IDEAlliance, an ever-widening group of developers and users and sponsored by Adobe is being put in place to take what has since become a stable technology and give it a means to grow.

and the security needs of a government agency will most likely exceed typical commercial uses.

Summit participants attended business or technical breakout meetings. The technical working group also discussed the state of the platform, issues surrounding the completion of the XMP Toolkit 4.x release, and new features and functions.

### XMP's Labyrinth

For all of us to benefit, a chicken-and-egg dilemma with metadata will have to be sorted out while the committee does its work on the technology. Schema, the collection of properties (labels) and values (the information filled into form fields) are being created everywhere daily by businesses, institutions and government agencies. The few available low-level professional tools for managing this data are not suited for use in a typical office. So while the standard for XMP might be defined, the data that will be fed into files is not, for want of an IDEAlliance-like standards management body to filter and rationalize the many into a few. That schema should be managed by a government agency such as the Library of Congress,

which could manage the dictionaries and schema, certify them, register the namespace and provide a centralized location to distribute them.

**Management tools.** The tools for managing these schema, taxonomies and values can stand alone or be embedded in digital asset management systems, content management systems, etc. Companies such as stock photo agencies, where the level of detail of the controlled vocabulary for searching is part of the value added to the success of the customer experience, would not lose anything but would gain every time they acquired a new collection because the core schema would match the one they were using. Google and Yahoo would benefit because the search criteria embedded in the uploaded image files would be consistent across applications and Web sites. Microsoft and Apple could easily extend their metadata operating search capabilities without much difficulty.

### Conclusion

If the only issues at hand were the collaboration of Adobe and IDEAlliance and the divvying up of respon-

sibility for managing the XMP metadata platform, XMP issues could be settled without much further ado. Unfortunately, it's more complicated than that. The main bones of contention are connected to metadata — not specifically XMP — and are not under the control of Adobe or IDEAlliance.

**Schema glut.** Without some regulatory body managing schema and perhaps even dictionaries, no platform will be successful. Every company wants, or thinks they want, custom schema. The truth is that the level of change they need involves controlled vocabularies (the values used to fill in the form fields) and not the labels or fields themselves.

**Conflict with existing implementations.** The need to adapt to custom labels and properties, sometimes on the fly, is one of the main reasons DAM vendors don't fully support XMP today. Having a neutral body responsible for certifying schema and distributing them would be ideal. For further evidence of today's schema glut, look at [Taxonomywarehouse.com](http://Taxonomywarehouse.com). Imagine how many unpublished schemas created by companies for their own use are not listed.

**Tools.** Resolving issues that arise when different schema are blended must be done in the host application (database) or in an independent third-party tool. Given the right tool, many users would be able to manage their own schema and controlled vocabularies. To date, the only tools available are Protégé 2000 ([Protege.stanford](http://Protege.stanford)) and Synaptica (from [Factiva.com](http://Factiva.com)). Both of these professional-level tools are meant to be used by taxonomists.

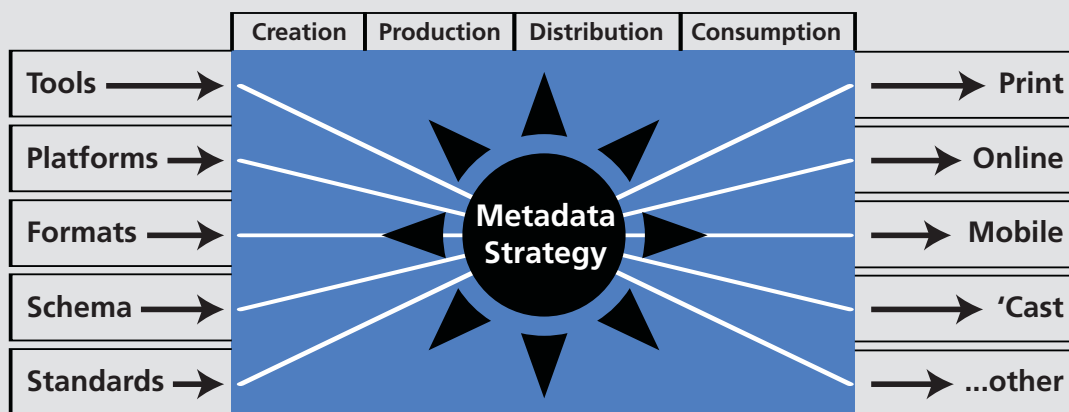


**Leadership.** The issues being sorted out now through the summit and IDEAlliance will be resolved eventually. This is a critical moment, however, and the current palpable level of enthusiasm and momentum must be maintained. The eyes-on-the-objective, no-nonsense team of Steinhardt and Salop seems to be up to the job. With Steinhardt's establishment of IDEAlliance procedures to support and grow the standard, Adobe's flexible and genuine interest in making this work, and Salop's industry connections, evangelism and event coordination, the future looks bright.

**TSR**

Technical group breakout session. Alan Lillich, chief XMP architect at Adobe, making presentation. Working group participants included EMC, MediaBeacon, Interwoven, Industrial Color, Pound Hill and TBWA. Photo courtesy of Andrew Salop

## A Holistic Approach to Content Lifecycle Management



SOURCE: Andrew Salop, Metaseed.net

**Standards** This slide presents some of the sober realities of the marketplace. The market will do what it wants to do when it wants to do it. Unlike standards for file formats such as TIFF and JPEG, which brought order to format chaos and alleviated user frustration, XMP is more discrete and less obvious in its impact and seems to require many levels of justification. In this slide, the points are: Market (business use justification for adopting the standard); Interoperability (requires that the standard be flexible and adaptable); Maturity (built on mature technology and tested in the market); and Diversity (ubiquity of XMP throughout Adobe applications may trump opportunities for new standards to emerge).