

MPEG-7 and DITA: Dovetailing of Two Industry Standards

Interview Excerpt

In his [2008 interview with Dave Kellogg](#), CEO of Mark Logic, The Content Wrangler (TCW) asked Mr. Kellogg to predict the course of video in XML formats such as [Darwin Information Typing Architecture](#) (DITA):

...

TCW: Do you see a markup language like DITA playing a role in video. Movies are modular—each segment is filmed independent of one another. These video content components should—hypothetically—be able to be managed like components of text content and reassembled dynamically, on demand. Do you see what we see? Should we be able to pull video components from a repository and reassemble them as needed? Or, can we do that now?

DK: Yes, it's all possible. MPEG-7 is all about XML and indexing metadata about the content. To find videos and clips, I think it will provide a great advance. Personally, I'd rather search metadata about a video than the captured speech-to-text. In fact, I think metadata is generally the key to multimedia indexing and re-use of all kinds. So, I think it will happen.

...

Video Challenges

Video has the potential of playing a major role in how we process and understand complex ideas. It can more thoroughly reach visual learners and complement technical media that has largely been transmitted through the written word and still images, but topic-based video is not without its hurdles. Any given video sequence is largely a "black box" because the contents are rarely described with the necessary level of precision. For example, one might attempt to make video management more tenable by chunking video segments at a relatively high level of granularity; however, few users of online documentation would watch 10 minutes of video to find 20 seconds of relevant material. There would have to be a way to move to the appropriate marker and view the context-sensitive video segment efficiently. Alternatively, one could chop video into many component pieces, but this effort would essentially create many black boxes if the segment descriptions were not managed in a systematically. Unfortunately, current multimedia search and retrieval methods lag behind text retrieval methods, making the task of video segment retrieval difficult to implement and manage. MPEG-7, as Dave Kellogg suggests, allows for transparency of video description information, thus providing a solution to the search and retrieval problem.

Topic-based Video

MPEG-7 is a multimedia (video, audio, voice, images, graphs, 3-D models, etc.) content description standard. MPEG-7 data can be passed into, or accessed by, a device or a computer code. This description standard is not aimed at any one application in particular; rather, it supports as broad of a range of applications as possible. MPEG-7 uses XML to store metadata that references multimedia instances. It provides fast and efficient searching, filtering, and content identification. MPEG-7 is not to be confused with other standards developed by the Moving Picture Experts Group (MPEG) such as MPEG-2 or MPEG-4, as these deal with the actual encoding of moving pictures. MPEG-7 makes it possible to separate the information about video (metadata) from the video itself. This separation of

concerns provides the key to the multimedia search conundrum, using a standard XML toolset and workflow.

Elements of the MPEG-7 schema can be attached to video time-code segments in order to tag particular events within video. These video segments can then be retrieved and viewed through the corresponding metadata. Video components described by MPEG-7 can be integrated into online documentation and made accessible to the end user in a similar way as are, for example, text-based components in DITA. From an author's perspective, adding video to a documentation set is as easy as inserting a resource id into the XML source. Like their text-based counterparts, video components may be reused across topics and documentation sets.

Those familiar with the [Semantic Web](#), the framework in which semantic markup (i.e., document annotation) and metadata are its core, will recognize many parallels and intersections with [MPEG-7](#). Integration between multimedia standards (e.g., MPEG-7) and Semantic Web standards (e.g., [RDF and OWL](#)) are on the cutting edge of processes that allow video integration within a wide variety of Web contexts.

The success of Web-based documentation is due in part to the development of standards that allow interoperability in open environments. As predicted, we are now seeing the emergence of applications that rely on the MPEG-7 standard to integrate video with the output of structured documentation systems. Not only will we continue to see more video integrations with DITA but with Docbook, S1000D and other structured formats as well.

Resources

- [Overview of MPEG7 - Superior Handling of Metadata Content](#) is an article that introduces some of the concepts of MPEG-7 from a 30,000 foot vantage point.
- The [MPEG-7 & MPEG-21 community](#) aims at bringing together experts from research and industry in the area of multimedia meta data interoperability for collaborative working environments. To accomplish this goal the portal provides information about MPEG-7 & MPEG-21 related events, participating research groups, as well as industry partners to share expertise on multimedia meta data interoperability.
- While the folks at [chiariglione.org](#) have described this page as an [MPEG-7 overview](#), be aware that it gets into the thick of MPEG-7 quite quickly.
- [Download the MPEG-7 Schema](#)