There has long been the promise that multimedia authoring will be as pervasive as word processing, the assumption being that people will want to routinely author rich content, and multimedia authoring tools will make it possible. But it’s not turning out that way. People want to keep it simple. Despite technical challenges being addressed from every perspective across tools, standards, and delivery platforms, rich multimedia content authoring for the mainstream is still elusive. Instead, users want to do less complicated and more instantaneous things such as tweeting freshly captured photos, posting unedited video files, applying simple effects to shared pictures, and pinning existing content. Although digital media is clearly becoming the new currency online, rich multimedia authoring is still in the domain of experts.

Substantial efforts have been made to empower average users to do more. One of the early standards for creating multimedia presentations is the Synchronized Multimedia Integration Language (SMIL, www.w3.org/AudioVideo), developed by the World Wide Web Consortium (W3C). SMIL can create spatially and temporally synchronized presentations of media objects that include animations and transitions. Researchers have developed numerous SMIL authoring tools, but despite efforts to create multiple improved versions of the standard since 1998, SMIL has not been widely adopted.

Similarly, a powerful coding scheme for complex multimedia scenes was standardized in MPEG-4 Part 11 that supports complex spatio-temporal presentation of media content, 2D and 3D object encoding, and descriptions of complex user-interaction behavior. Originally providing a Binary Format for Scenes (BIFS), MPEG-4 Part 11 later provided a textual format called the Extensible MPEG-4 Textual Format (XMT) that made it easier to author and edit MPEG-4 multimedia presentations (http://mpeg.chiariglione.org/technologies/mpeg-4/mp04-bifs/index.htm). Unfortunately, BIFS and XMT have not attracted much interest.

More recently, the W3C has been creating HTML5 (www.w3.org/TR/html5), a markup language that promises significant improvements in media handling for the Web. This includes addressing a fundamental gap in Web standards today in that multimedia must be handled through opaque HTML objects and proprietary browser plug-ins. Although HTML5’s support for complex presentation is not as extensive as SMIL and MPEG-4 BIFS/XMT, standardizing how audio, video, and graphics are provided as Web content is an important step forward. Still, it is unclear whether HTML5 will make us all rich multimedia content authors.

It is overwhelmingly apparent, however, that interest from Web users goes beyond just consuming media content. Online and mobile users are extremely active in contributing, manipulating, and redistributing media objects. YouTube recently reported hitting a new high watermark of receiving on average 72 video-hours of video content per minute across its tens of millions of channels (www.youtube.com/t/press_statistics). There is also tremendous interest in photo-sharing sites such as Flickr, which recently surpassed 6 billion photo uploads.

Although media content is clearly at the center of YouTube and Flickr, social networking sites also receive a tremendous number of media objects. Facebook recently reported that it receives 6 billion photos per month. The total number of photos uploaded to Facebook to date is estimated to be more than 100 billion. Foursquare reported that it has received billions of location-based check-ins (https://foursquare.com/about). Many of them include geotagged photos from mobile users.
Clearly, average users are contributing media objects in large volumes. They upload everything from photos of friends and family members to videos that capture their travels and even daily routines. But this is poor media—it lacks the dynamics, interactivity, and composition expected to be at the heart of multimedia authoring for average users by now. Nevertheless, people increasingly want to express creativity with media objects, even if it is with simple manipulations. For example, Instagram has captured tens of millions of mobile users simply by making it easy to stylize and share photos. Without a doubt, helping users share the content they like boasts creativity. For example, Pinterest lets users “pin” interesting photos and then automatically compose them into boards that can be shared and commented on. Although the boards are not what we would call rich media, pinning makes for pure, simple routine authoring.

This growing interest in stylizing content and pinning media objects is blazing a new path of creativity. For now, we’ll need to be content with poor content. At least there’s a lot of it. And we can keep our hopes alive for a richer future.

References

John R. Smith is a senior manager of Intelligent Information Management at IBM T.J. Watson Research Center. Contact him at jsmit@us.ibm.com.

Selected CS articles and columns are also available for free at http://ComputingNow.computer.org.