Technology and art – putting things in context

Leonardo Chiariglione
Telecom Italia Lab.
Via G.Reiss Romoli, 274
I-10148 Torino (ITALY)
Phone: +39 011 228 6120
leonardo.chiariglione@tilab.com

ABSTRACT

A set of standard technologies already developed or under development by the Moving Picture Experts Group (MPEG) promises to bring back to authors the control of their works. The technologies are those of Content Representation, Digital Item Declaration, Interoperable Intellectual Property Management and Protection, and Metadata.

1. INTRODUCTION

One cannot stop being admirative of the depth and breadth of achievements made by our forefathers in their intellectual endeavours. How could such works as Homer's Iliad, Confucius' Analects, Murasaki Shikibu's Makura no Soshi or Dante's Divine Comedy be produced in those remote times, when life was so harsh and the authors of those works knew that there was no means for them to be compensated of their creations?

The works they so ingenuously created and liberally donated to humankind shaped its evolution. This virtuous example of works freely available to all and influencing generations of learned people has led some to advocate the recreation of an environment where the flow of information is unencumbered, so that the interchange of ideas among peers eventually gives rise to the great idea that some people will exploit.

There is reason to moderate one's enthusiasm. The societies that have produced those cases of excellence were those where a tiny layer of aristocrats ruled over and exploited a mass of people with no rights and only one duty – to serve that tiny layer of aristocrats and provide them with all they needed. The members of the top layer of that society could, if they so desired, afford spending their time in pursuit of intellectual achievement and, indeed, many of them did achieve excellence. Unfortunately or, better, fortunately, this is hardly a model that can be transplanted to the society of today.

There is one element that made the arrangement of those times possible and that was the fact that the author was in complete control of his work. He produced his written work and the more the work got distributed, the happier he was. Cicero, a representative of Roman aristocracy, successfully defeated Catalina's conspiracy and published his famous "Oration against Catilina" as a useful means to further his political ambitions.

Gutenberg's invention of moving characters was the tool that motivated more society layers to produce works by giving a monetary value to the work itself. For the first time authors had the means to gain from their published works. Printing, however, was also the instrument that took control of the work out of the hands of the author and handed it over to a middleman – the publisher. This was possible because of the high cost of the printing equipment that only wealthy entrepreneurs could afford. To this the cost of distribution was added later.

The result is that, in spite of the efforts made by the French Revolution to enshrine "author's rights", it is fair to say that today the "author's rights" may be a noble concept more or less respected depending of one's country, but the control of the work is in the hands of the publisher.

To consume a book no special device was required and the same was true for photography. With the invention of cinematography, sound recording, radio, television etc. devices were needed that allowed the user to view or listen to the content. Today the control of the work is in the hands of the publisher, but the means to consume the content are more and more in the hands of technology providers.

The facts of the last few years are putting a strain to this arrangement. Control of content is nominally in the hands of the publisher, but when digital content reaches the playback device, it is in the hands of the end user who has been convinced to buy the device from a middleman thanks to its ability to copy content. This copying function can now be exercised with an ease that was unthinkable before and because everybody can have a copy of digital content, content loses its value. The publisher loses and, as a reflection, the author loses as well. And if the author loses it is like a spring drying up: for some time people downstream will see water but after a while the river, too, will dry up.

We live in times of transition. So far technology has been used in a piece-meal fashion enabling and disabling opportunities hitting or benefiting users at random. But technology can also be used in a way that allows the author to recover control of the work that Gutenberg took away from him half a millennium ago.

This article presents some of the author's ideas on how the individual's art can be served by technology along with an introduction of some relevant parts of the work carried out by the MPEG group (ISO/IEC JTC1/SC29/WG11) in this area.

2. KEEP IT SIMPLE

The first issue faced by today's author is the sheer complexity of the "work" he can create. In Homer times it was enough to be able to play a zither and "sing" a poem. Setting aside the fact that the poem and the music were likely the result of centuries of spontaneous poetical formations, a single person could do everything. Today, and even more so in the future, a single piece of work may entail the participation of many different actors.

Let's take the example of a "music compilation". In a way that is already possible today, even if the result can be rather clumsy, one can only imagine the number of different players providing such elements as

- The music tracks (obviously)
- The lyrics of the music
- The scores
- The MIDI files for use on cell phones

- Some photos of singers and performers
- Ditto for videos of their performances
- Ditto for interviews
- News related to the song as soon as they become known
- Statements by opinion makers as soon as they become known
- Ratings of agencies as soon as they become known
- Navigational information driven by user preferences
- Bargains related to songs/artists
- Etc.

After, and in some cases during, the production of these pieces of information it is necessary to digitally encode them to match the delivery media of today or those to be used a few years from now.

This is the time to dispel one of those media encoding legends, viz. that since compression is constantly improving,, many content formats must be supported to keep up with the constant flow of these new content formats.

That media encoding progresses is not untrue, but the fact that a new scheme comes up providing some improvement compared to existing schemes is not a sufficient reason for throwing away a scheme that is already widely deployed and works just well enough.

This legend is kept alive by some of the technology middlemen whose role is enhanced by keeping fuzzy the surroundings of the content consumption device. Indeed every new codec implies the creation and distribution of new devices – it does not matter if they are hardware of software. If the author is to be empowered with the ability to distribute his content, he cannot be expected to be concerned with distribution of new content consumption devices, the only thing he cares about is having his content playable by his prospective customers.

The second issue concerns the need to organise all components of the multimedia work in such a way that access and consumption of the different pieces of content become possible. This translates into the definition of an architecture of this multimedia object. In the MPEG-21 parlance, multimedia objects are called Digital Items (DI). A DI is a collection of multimedia resources (content) that are held together based on certain relationships and which may have metadata associated with them.

The Digital Item Declaration (DID) is the standard form this "multimedia work representation" takes in MPEG-21. The Digital Item Declaration (DID) standard describes a set of general terms that allow the modeling of a broad range of multimedia content. The aim is to allow a work to be expressed as a Digital Item and thus it allows for the description of aspects ranging from the specific media resources through to abstract concepts such as choices available to a user.

3. MANAGING AND PROTECTING CONTENT

If the springboard of progress is the ability of humans to be amazed by things remarkable happening around them, it is fair to say that the last few years of Internet craze have provided ample opportunities for people to be amazed by unexpected phenomena. It is debatable, however, whether this has also provided opportunities for progress.

The most striking example is provided by the discovery made in those Wild Web years that such low-bandwidth information sources as MP3-encoded audio files could be delivered at no cost using the Internet. This discovery made self-appointed Internet prophets declare that the Internet would bring the role of the middleman on the web to an end, because producers of information would be able to reach their customers without the need of somebody in between them.

Easier said than done, not only because it soon turned out that information could indeed be delivered at no cost over the Internet, but mostly because the information became accessible to anybody and infinitely replicable: as a result, the information itself lost all monetary value. This "discovery" was just something that other business domains – specifically pay television – had discovered some 20 years before.

The pay TV business decided that, for the business to exist, the delivery of pay TV should happen in encrypted form. Today some people think that the same approach should be adopted for delivery of content on the Internet. There are, however, a few reasons why this is unlikely to work. The first is that this method must still be proved to work for the environment for which it has been designed. Indeed most digital pay TV operators are in deep trouble and some of them are outright bankrupt. And we are talking here of a domain where users are already somehow accustomed to receive TV programs using a special set top box provided by the pay TV operator.

There is one more reason why this approach is unlikely to work on the Internet. This is the fact that Internet users have been accustomed to access information in a seamless way using a standard browser. Users are now unlikely to accept something that requires clumsy interactions, much less something that requires users to switch devices depending on the URL that their browser is pointing at.

Still without the ability to prevent unauthorised people from accessing content, there is no business possible on the net.

The answer lies in a technology that allows people to access protected content in a seamless way by using a device that is capable of decrypting content that has been protected, possibly in a number of different ways, depending on where it comes from. This technology is being standardised by MPEG with the name of Intellectual Property Management and Protection (IPMP) Extension (IPMP-X). With this technology it will be possible for anybody to protect his content knowing that, if certain conditions will be satisfied (e.g., that a financial transaction has taken place), an end-user device will be capable of playing back protected content.

In addition to this key technology more is needed. If I am an author and I want to give somebody else access to a DI, I need to identify that particular DI that I want to give access to. This is a similar problem to the one the book industry had to face and which led to the definition of the International System of Book Numbering (ISBN). Similar identification systems exist for other resources (audio, video, photographs, periodicals…). The system that MPEG-21 has developed with similar functionality is called Digital Item Identification (DII).

But there is more. When a transaction occurs in the physical world, either there is a general understanding of the rules governing the transaction or, especially when the transaction involves a large amount of money, a contract is stipulated between the parties. If somebody buys a compact disc (CD) he will pay the same amount of money, irrespective of the number of times he will listen to that CD. It would be in the interest of both the seller and the buyer to be able to attach specific usage rules to a piece of content, depending of the value of the transaction.

Rights Expression Language (REL) is another standard that is being developed in the MPEG-21 suite of standards. This will enable the expression of rights associated to a work. The language is very generic and can be used to express rights to any type of

objects. Rights Data Dictionary (RDD) is a collection of data specific of the content business that will be referenced by a specific rights expression statement.

4. DESCRIBING CONTENT

The original role of the publisher was to convert a work to a physical object that could be distributed to interested buyers. In due time, this "seller market" role had to extend because so many competing offerings were available that a "buyer market" situation began to form. Today the media market has several faces, one of them being made of big hits, requiring a special type of promotion, and another of a multitude of other offers, both small and big, where a different kind of promotion is needed. Today promotion of works is a fundamental component of the content business.

Giving authors the ability to produce their works in a way that leaves the control of the consumption in their hands but still does not require special consumption devices is a necessary, but far from sufficient, condition. What is the use of something that exists, if no one is aware of its existence?

The MPEG-7 suite of standards produced by MPEG provides some key technological components to enable authors to produce the necessary information to be used for the purpose of pushing descriptions to users or allowing interested users to search databases of descriptions or to browse through descriptions.

MPEG-7 offers a comprehensive set of audiovisual Descriptors that can be composed to create Description Schemes. The Description Definition Language (DDL) is used for the syntactic definition of MPEG-7 Description Tools and for allowing extensibility of Description Tools. The standard also provides BiM (Binary Format for MPEG-7) for the bit-efficient representation of audio-visual metadata.

5. WHERE WE ARE

It is important to know how much of the grand plan outlined above can be relied on, first in terms of availability of standards and then in terms of the actual adoption in products and services.

The table below gives the dates of expected/actual adoption for some of the major standards that are needed to implement the vision.

Table 1: standards adoption timeline

		Availability
Audio-coding	MP3	1992/11
	AAC	1997/04
Video-coding	MPEG-1	1992/11
	MPEG-2	1994/11
	MPEG-4	1998/10
Multiplexing	MPEG-1	1992/11
	MPEG-2	1994/11
	MPEG-4	1998/10
Media composition	MPEG-4	1998/10
Audio metadata	MPEG-7	2001/07
Video metadata	MPEG-7	2001/07
Multimedia metadata	MPEG-7	2001/07
IPMP Extension	MPEG-4	2002/07
Digital Item Declaration	MPEG-21	2002/05
Digital Item Identification	MPEG-21	2002/07
Rights Expression Language	MPEG-21	2003/07
Rights Data Dictionary	MPEG-21	2003/07

6. CONCLUSIONS

This paper has introduced the vision of the author who is finally given back control of his work through a set of interoperable technologies designed to fit in a general scheme.

It will indeed be possible to put the author back in the driving seat, but thinking that middlemen will be wiped away is repeating the same mistake of the early Internet years when people thought that "delivery" was the end of the story.

The role of middlemen will stay as important as ever in their function to create the bridge between a sea of offers and a population of prospective users.