MPEG-21: Goals and Achievements

Ian Burnett
University of Wollongong, Australia

Rik Van de Walle
Ghent University, Belgium

Keith Hill
Rightscom, London

Jan Bormans
Interuniversity MicroElectronics Center, Belgium

Fernando Pereira
Instituto Superior Técnico, Portugal

MPEG-21 is an open standards-based framework for multimedia delivery and consumption. It aims to enable the use of multimedia resources across a wide range of networks and devices. In this article, we discuss MPEG-21's parts, achievements, ongoing activities, and opportunities for new technologies.

The appetite for content consumption and easier access to information continues to increase rapidly. Access devices are also evolving so that we have a wide choice of terminal and network capabilities. The result is that people, both in their personal and professional lives, are increasingly creators as well as consumers of digital media. They thus demand solutions that deliver accessible and advanced multimedia creation and consumption on many platforms. These new content providers and the traditional media sources share a core set of concerns: management of content, repurposing content based on user preferences and device capabilities, protection of rights, protection from unauthorized access/modification, protection of privacy of providers and consumers, and so on. The MPEG-21 Multimedia Framework initiative aims to meet these challenges by enabling the use of multimedia resources across a wide range of networks and devices. In this article, we discuss MPEG-21, an open standards-based framework for multimedia delivery and consumption by all the players in the delivery and consumption chain.

MPEG-21 context

MPEG-21 is the newest of a series of standards being produced by the Moving Picture Experts Group (see http://mpeg.telecomitalialab.com), more formally known as ISO/IEC JTC 1/SC 29/WG11. MPEG has a long history of producing multimedia standards: MPEG-1 (1993), MPEG-2 (1996), MPEG-4 (1999), and MPEG-7 (2001).

MPEG-1 and MPEG-2 resulted in many successful commercial products and services, such as video-CD, DVD, digital television, digital audio broadcasting (DAB), and MP3 (MPEG-1/-2 audio layer 3).

More recently, MPEG-4[1] standardized audio–visual coding solutions that address the needs of communication, interactive and broadcasting services, and many combinations of these models. At the heart of MPEG-4 technology is an object-based representation approach in which a scene is modeled as a composition of both natural and synthetic objects with which a user may interact. Through this new coding approach, MPEG-4 creates new opportunities for users to play with, create, reuse, access, and consume audio–visual content. MPEG-4 is currently used on the Internet (such as Quicktime 6) and is being adopted in several other communications contexts.

Building on MPEG-4's object-based vision, the MPEG-7[2] standard concentrates on the description of multimedia content. It enables quick and efficient searching, identification, processing, and filtering of multimedia material. This lets users (both humans and machines) use and interact with non-text-based multimedia content directly and efficiently.

Together, the MPEG-1, -2, -4, and -7 standards provide a complete, powerful, and successful set of tools for multimedia representation. Other multimedia-related standards (and proprietary solutions) exist, such as JPEG and JPEG 2000. However, widespread deployment of multimedia applications requires more than this loose collection of standards. The problem is that overall multimedia consumption and commerce remain nontransparent and aren't happening on a large scale. In late 1999 and early 2000, MPEG examined several questions: Do all existing multimedia-related standard specifications fit together? Does anybody know how they fit together? Do we have specifications for all the

1070-986X/03/$17.00 © 2003 IEEE

Published by the IEEE Computer Society

necessary technical elements for multimedia transactions? Which standard activities are most relevant? Who is making the "glue" that will let these standards fit together?

Many high-quality multimedia standards are being created to meet the needs of different communities. However, the result is an array of standards offering parallel or competitive solutions. In contrast to this a transparent multimedia world requires the creation of a single, big picture view of multimedia.

Currently, multimedia technology provides the different players in the value and delivery chain with an excess of information and services. However, no complete solutions exist that let different communities—content, financial, communication, and computer and consumer electronics and their customers (each with their own models, rules, procedures, interests, and content formats)—interact efficiently using this complex infrastructure. Developing a common multimedia framework would facilitate cooperation between sectors and support a more efficient implementation and integration of the different needs of each community, resulting in an enhanced user experience.[3]

One problem, however, is that in the desire to achieve interoperability, the framework may violate the requirement to protect the value of the content and the rights of the rights holders. From the alternate view, digital rights management (DRM) systems—which protect those rights—can prevent interoperability if they use nonstandardized protection mechanisms. Realizing an open multimedia infrastructure thus requires an integrated and suitably normative DRM system.

Thus, MPEG-21 seeks to create the big picture of multimedia standards. It aims to guarantee interoperability by focusing on how the elements of a multimedia application infrastructure should relate, integrate, and interact. Finally, where open standards for elements are missing, MPEG is creating new MPEG-21 parts to fill the gaps.

MPEG-21 objectives

A key assumption of MPEG-21 is that every human is potentially an element of a network involving billions of content providers, value adders, packagers, service providers, consumers, and resellers. Thus, besides client-server-based applications, peer-to-peer networking and the resulting flexibility of user roles has been an underlying part of MPEG-21 thinking since the early days of the standardization process.

Interoperability is the driving force behind all multimedia standards. It's a necessary requirement for any application that requires guaranteed communication between two or more parties. From a more philosophical view, interoperability expresses the user's dream of easily exchanging any type of information without technical barriers. To achieve this goal, we must standardize both the content structure and a minimum set of communication processes. The key to effective standardization is to create a minimum standard that normatively defines a minimal (but complete) set of tools that will guarantee interoperability. Such minimum specifications provide space and a foundation for competitive, proprietary, and alternate developments (which wouldn't be normative) of tools that don't need standardization to obtain interoperability. This enables the incorporation of technical advances—lengthening the lifetime of the standard—as well as stimulating technical and product competition. A standard also has important economic implications because it allows the sharing of investment costs and the acceleration of application deployment. A further advantage is that an open standard reduces consumer reliance on sole proprietary solutions; this is essential if we want truly transparent multimedia usage and consumption to be a reality.

MPEG has been proactive in identifying current multimedia initiatives and inviting collaboration in the context of MPEG-21. Examples of envisaged collaborations are the Open eBook Forum, International Telecommunications Union, Open Mobile Alliance, and so on. The aim of this process is to maximize interoperability, minimize the overlap between concurrent activities, and share common technologies.

To ensure that the standardization of a big picture encompassed both technological and user demands, MPEG established a formal development process as follows:

1. Define a framework supporting the MPEG-21 vision statement.
2. Identify the multimedia framework's critical components.
3. Understand how the framework's components relate and identify where standards technology gaps exist.
4. Involve relevant (and complementary) standardization bodies in this effort.

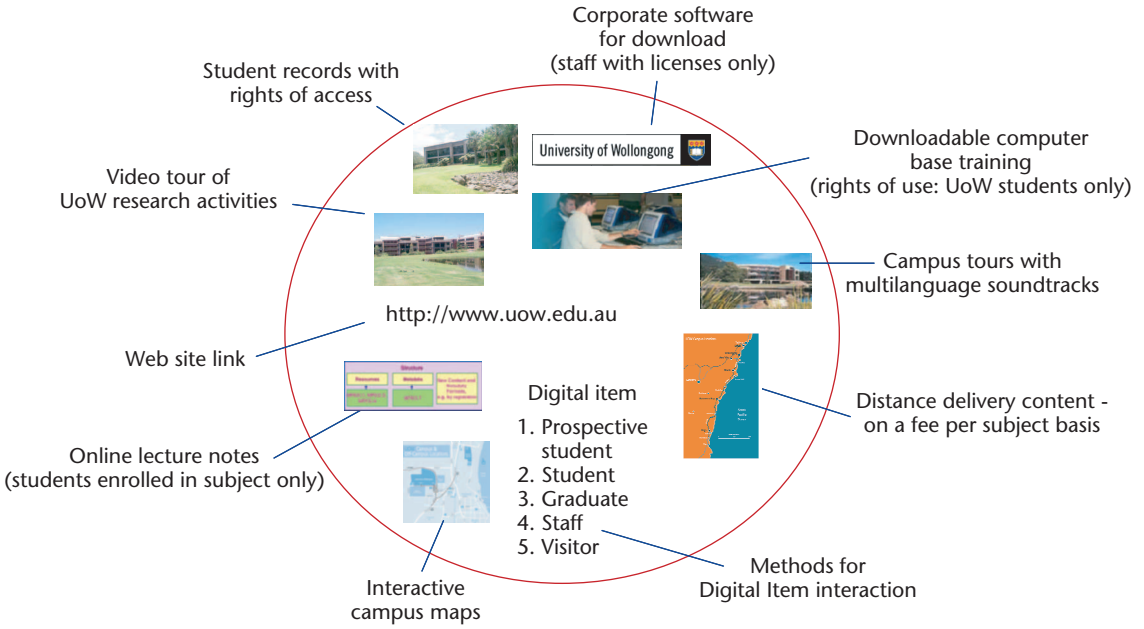

Figure 1. Example Digital Item.

5. Assess each of the unavailable technologies. If they fall under MPEG expertise, then MPEG develops the corresponding standards. If not, engage other bodies to achieve their development.

6. Integrate the relevant available and developed technologies.

At the time of writing, this process has been underway for some two and a half years. To date, several parts of the MPEG-21 framework are fully standardized and several others are in an advanced state of development.

Basic concepts and elements

The basic concepts in MPEG-21 relate to the what and who within the multimedia framework. The what is a Digital Item that's a structured digital object with a standard representation, identification, and metadata within the MPEG-21 framework. The who is a user who interacts in the MPEG-21 environment or uses a Digital Item, including individuals, consumers, communities, organizations, corporations, consortia, governments and other standards bodies, and initiatives around the world. The user roles include creators, consumers, rights holders, content providers, distributors, and so on—there's no technical distinction between providers and consumers.[4] All parties that must interact within MPEG-21 are categorized equally as users. They assume specific rights and responsibilities according to their interaction with other users. All users must also express and manage their interests in Digital Items.

In practice, a Digital Item is a combination of resources, metadata, and structure. The resources are the individual assets or (distributed) content. The metadata describes (distributed) data about or pertaining to the Digital Item as a whole or also to the individual resources in the Digital Item. Finally, the structure relates to the relationships among the parts of the Digital Item—both resources and metadata. An example of a Digital Item might be a presentation of a university, including photos, videos, animation graphics, textual information, news related to the university's research activities, e-learning material, navigational information driven by user preferences, and so on (see Figure 1). The Digital Item is thus the fundamental unit for distribution and transaction within the MPEG-21 framework.

Parts of MPEG-21

MPEG-21 is organized into several independent parts primarily to allow various slices of the technology to be useful as stand-alones. This maximizes their usage and lets users implement them outside MPEG-21 as a whole in conjunction with proprietary technologies. For example, we use MPEG-2 video together with MPEG-2 systems but not with MPEG-2 audio in the context of the US digital TV system. However, although we may use the various parts independently, they were developed to give optimal results when used together.

The MPEG-21 parts already developed or currently under development are

1. *Vision, technologies, and strategy*: describes the multimedia framework and its architectural elements with the functional requirements for their specification.[5]
2. *Digital Item Declaration (DID)*: provides a uniform and flexible abstraction and interoperable schema for declaring Digital Items.
3. *Digital Item Identification (DII)*: defines the framework for identifying any entity regardless of its nature, type, or granularity.
4. *Intellectual property management and protection (IPMP)*: provides the means to reliably manage and protect content across networks and devices.
5. *Rights Expression Language (REL)*: specifies a machine-readable language that can declare rights and permissions using the terms as defined in the Rights Data Dictionary.
6. *Rights Data Dictionary (RDD)*: specifies a dictionary of key terms required to describe users' rights.
7. *Digital Item Adaptation (DIA)*: defines description tools for usage environment and content format features that might influence the transparent access to the multimedia content—notably terminals, networks, users, and the natural environment where users and terminals are located.
8. *Reference software*: includes software that implements the tools specified in the other MPEG-21 parts.
9. *File format*: defines a file format for storing and distributing Digital Items.
10. *Digital Item Processing (DIP)*: defines mechanisms for standardized and interoperable processing of the information in Digital Items.
11. *Evaluation methods for persistent association technologies*: documents best practices in evaluating persistent association technologies using a common methodology (rather than standardizing the technologies themselves). These technologies link information that identifies and describes content directly to the content itself.
12. *Test bed for MPEG-21 resource delivery*: provides a software-based test bed for delivering scalable media and testing/evaluating this scalable media delivery in streaming environments.

MPEG-21 achievements

Here we provide an overview on currently available MPEG-21 technologies. For those under development (parts 4 and 8 through 12), we consider only those parts that, at the time of writing, are at a high level of maturity.

Vision, technologies, and strategy

The MPEG-21 technical report describes the multimedia framework and its architectural elements with the functional requirements for their specification.[5]

This part is available as a free downloadable document from the International Organization for Standardization Web site at http://www.iso.ch/iso/en/aboutiso/introduction/index.html.

Digital Item Declaration

The Digital Item Declaration part is the second part of MPEG-21 (ISO/IEC 21000-2).[6] It specifies a uniform and flexible abstraction and interoperable schema for declaring the structure and makeup of Digital Items. Through the Digital Item Declaration Language (DIDL), we can declare a Digital Item by specifying its resources, metadata, and their interrelationships.

The ISO/IEC 21000-2 describes this DID technology in four main sections:

- *Model*. The DID model describes a set of abstract terms and concepts for defining Digital Items. Within this model, a Digital Item is the digital representation of a work (for example, a digital music album, an e-book, or a piece of software including setup and configuration information). As such, the Digital Item is the thing that's acted upon (managed, described, exchanged, collected, and so on) within the model.
- *Representation*. The DIDL is based on the terms and concepts defined in the DID model. It contains the normative description of the syntax and semantics of each of the DIDL elements, as represented in XML.

- *Schema*. The complete normative XML schema for DIDL comprises the entire grammar of the DID representation in XML.

- *Some detailed examples*. Illustrative examples of DIDL documents are provided to aid in understanding the use of the specification and its potential applications.

Major concepts in the context of the DID include the following:[6]

- A *container* is a structure that groups items and/or containers. We can use these groupings of items and/or containers to form logical packages (for transport or exchange) or logical shelves (for organization). Descriptors allow for the labeling of containers with information that's appropriate for the purpose of the grouping (for example, delivery instructions for a package or category information for a shelf).

- An *item* is a grouping of subitems and/or components bound to relevant descriptors. Items may contain choices, which can be customized or configured. Items may also be conditional (on predicates asserted by selections defined in the choices). We consider an item that contains no subitems an entity—a logically indivisible work. An item that contains subitems is a compilation.

- A *component* binds a resource to a set of descriptors. These descriptors are information related to all or part of the specific resource instance. Such descriptors will typically contain control or structural information about the resource (such as bit rate, character set, start points, or encryption information) but not information describing the content within.

- An *anchor* binds descriptors to a fragment, which corresponds to a specific location or part of a resource.

- A *descriptor* associates information with the enclosing element. This information may be a component (such as a thumbnail of an image or a text component) or a textual statement.

- A *condition* describes the enclosing element as being optional and links it to the selection(s) that affect its inclusion.

- A *choice* describes a set of related selections that can affect an item's configuration.

- A *selection* describes a specific decision that affects one or more conditions somewhere within an item.

- An *annotation* describes a set of information about another identified element of the model without altering or adding to that element.

- An *assertion* defines a full or partially configured state of a choice by asserting true, false, or undecided values for some number of predicates associated with the selections for that choice.

- A *resource* is an individually identifiable asset such as a video or audio clip, an image, or a textual asset. A resource may also potentially be a physical object. All resources must be locatable via an unambiguous address.

- A *fragment* unambiguously designates a specific point or range within a resource. Fragments may be resource-type specific.

- A *statement* is a literal textual value that contains information, but not an asset. Examples of likely statements include descriptive, control, revision tracking, or identifying information (such as an identifier as described in MPEG-21 part 3[7]).

- A *predicate* is an unambiguously identifiable declaration that can be true, false, or undecided.

Figure 2 shows some of the most important elements within the DID model and how they're related. Figure 3 (on p. 66) shows a simple example of a DID, based on the Digital Item in Figure 1. Here, the DID is a container consisting of one global descriptor and one item. The latter consists of two descriptors, one choice, and two subitems. The choice lets the user choose between a video presentation and a textual presentation. When the user chooses the video presentation (and only then, due to the condition in the first subitem), the user will receive an MP4 file (research.mp4). When the user chooses the textual presentation (and only then, due to the condition in the second subitem), the user will receive a text file (lecturenotes.txt).

Digital Item Identification

The third part of MPEG-21 (ISO/IEC 21000-3)[7] specifies

- how to uniquely identify Digital Items (and parts thereof),
- how to uniquely identify intellectual property related to the Digital Items (and parts thereof),
- how to uniquely identify description schemes,
- the relationship between Digital Items (and parts thereof) and existing identification systems (such as the International Standard Recording Code, International Standard Book Number, International Standard Serial Number, digital object identifier, and so on), and
- the relationship between Digital Items (and parts thereof) and relevant description schemes (such as the International Standard Musical Work Code, International Standard Text Code, and so on).

Note, however, that the DII standard doesn't specify

- new identification systems for the content elements for which identification and description schemes already exist and are in use (for example, it doesn't attempt to replace the ISRC, as defined in ISO 3901,[8] for sound recordings) and
- normative description schemes for describing content.

The DID and DII parts are closely related; we can associate identifiers covered by DII with Digital Items, containers, components, and/or fragments thereof by including them in a statement in the DID (see Figure 4 on p. 67). Examples of likely statements include descriptive, control, revision tracking, and/or identifying information.

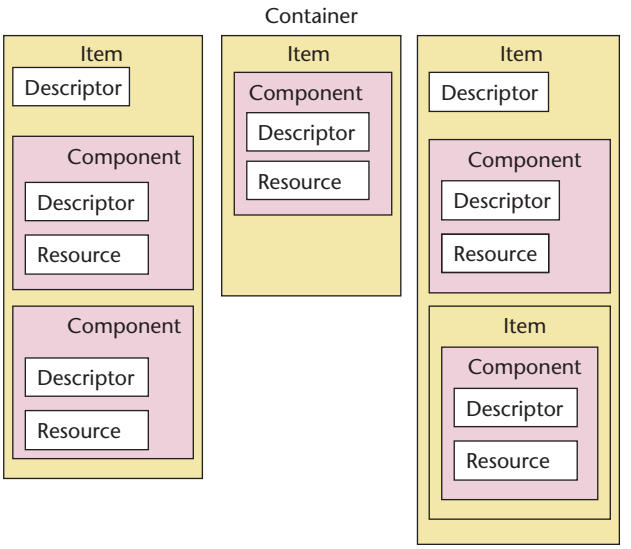

Figure 2. Some Digital Item Declaration model elements and their relationships.

Rights Expression Language

The fifth part of MPEG-21 (ISO/IEC 21000-5)[9] explains the basic concepts of a machine-interpretable language for expressing the rights of users that act on Digital Items, components, fragments, and containers. It doesn't provide specifications for security in trusted systems, propose specific applications, or describe the details of the accounting systems required. Nor does it address the agreements, coordination, or institutional challenges in building an implementation of this standard. However, it does describe the language's syntax and semantics.

REL can declare rights and permissions using the terms as defined in the RDD. It provides mechanisms to support the use of digital resources in publishing, distributing, and consuming of digital movies, digital music, electronic books, broadcasting, interactive games, computer software, and other creations in digital form, in a way that protects digital content and honors the rights, conditions, and fees specified for digital contents. It also specifies access and use controls for digital content in cases where financial exchange isn't part of the terms of use and supports exchange of sensitive or private digital content.

REL also provides a flexible, interoperable mechanism to ensure that personal data is processed in accordance with individual rights. REL supports guaranteed end-to-end interoperability, consistency, and reliability between different systems and services. As such, it offers richness and extensibility in declaring rights, conditions, and obligations; ease and persistence in identifying and associating these with digital contents; and flexibility in supporting multiple usage/business models.

Rights Data Dictionary

This sixth part of MPEG-21 (ISO/IEC 21000-6)[10] specifies a RDD for use within the MPEG-21 framework. The RDD forms the basis of all expressions of rights and permissions as defined by REL.

```
<?xml version="1. 0" encoding="UTF -8"?>
<DIDL
xmlns="urn:mpeg:mpeg21:2002:01 -DIDL-NS"
xmlns:xsi ="http://www.w3.org/2001/XMLSchema -instance"
xsi:schemaLocation ="urn:mpeg:mpeg21:2002:01 -DIDL-NS E:\Users\RVdW\Temp\DIDL.xsd">
    <Container >
        <Descriptor >
            <Statement mimeType ="text/plain">
                This information package was developed by University Records Unlimited
            </Statement >
        </Descriptor >
        <Item>
            <Descriptor >
                <Statement mimeType ="text/plain">
                    Copyright owner: University Records Unlimited
                    Permission: Read  Only
                </Statement >
            </Descriptor >
            <Descriptor >
                <Component >
                    <Resource ref="http://www.uruweb.org/logos/uow.jpg " mimeType ="image/jpeg "/>
                </Component >
            </Descriptor >
            <Choice choice_id ="INFO_PICKER ">
                <Descriptor >
                    <Statement mimeType ="text/plain">
                        Choose the information you want to receive:
                    </Statement >
                </Descriptor >
                <Selection select_id ="VIDEO">
                    <Descriptor >
                        <Reference target="#VIDEO_TITLE "/>
                    </Descriptor >
                </Selection >
                <Selection select_id ="TEXT ">
                    <Descriptor >
                        <Reference target="#TEXT_TITLE "/>
                    </Descriptor >
                </Selection >
            </Choice >
            <Item>
                <Condition require="VIDEO"/>
                <Descriptor id="VIDEO_TITLE ">
                    <Statement mimeType ="text/plain">
                        Research over view video
                    </Statement >
                </Descriptor >
                <Component >
                    <Condition require ="VIDEO"/>
                    <Resource ref="http://www.uruweb.org/video/research.mp4 " mimeType ='video/mp4 "/>
                </Component >
            </Item>
            <Item>
                <Condition require="TEXT "/>
                <Descriptor id="TEXT_TITLE ">
                    <Statement mimeType ="text/plain">
                        Lecture notes
                    </Statement >
                </Descriptor >
                <Component >
                    <Condition require ="TEXT "/>
                    <Resource ref="http://www.uruweb.org/text/lecturenotes.txt " mimeType="text/plain"/>
                </Component >
            </Item>
        </Item>
    </Container >
</DIDL >
```

Figure 3. Example Digital Item declaration.

The RDD comprises a set of clear, consistent, structured, integrated, and uniquely identified terms to support REL. The RDD specifies the structure and core terms and specifies how further terms may be defined under the governance of a registration authority.

Use of the RDD will facilitate the accurate exchange and processing of information between interested parties involved in the administration of rights in, and use of, Digital Items.

The RDD also supports the mapping and transformation of metadata from the terminology of one namespace (or authority) into that of another namespace (or authority) in an automated or partially automated way, with the minimum ambiguity or loss of semantic integrity.

The RDD is a prescriptive dictionary, in the sense that it defines a single meaning for a term represented by a particular RDD name (or headword), but it's also inclusive in that it recognizes the prescription of other headwords and definitions by other authorities and incorporates them through mappings. The RDD also supports the circumstance that the same name may have different meanings under different authorities. The RDD has audit provisions so that we can track additions, amendments, and deletions to terms and their attributes.

The RDD recognizes legal definitions as, and only as, terms from other authorities that we can map into the RDD. Therefore, terms that are directly authorized by the RDD neither define nor prescribe intellectual property rights or other legal entities.

The RDD standard and the RDD are distinct things. The standard provides, first of all, a methodology for making the dictionary. The standard then contains the initial group of terms—the RDD standardized terms—defined according to this method. Finally, the standard also provides for a registration authority that will manage the ongoing work of adding to the dictionary.

For its part, the dictionary also has the characteristics of an ontology. It provides a set of well-defined terms for use in rights expressions for Digital Items. In recognition of the great diversity and complexity associated with multimedia content, it also represents as many different definitions of terms as its users require. Further, it shows their relationships in a structured way to support the mapping and transformation of terms between different schemas and systems. For these reasons, the standardized terms in the dictionary aren't the definitive list, but the foundations. The standard is the starting point for a dynamic dictionary and not the finished article.

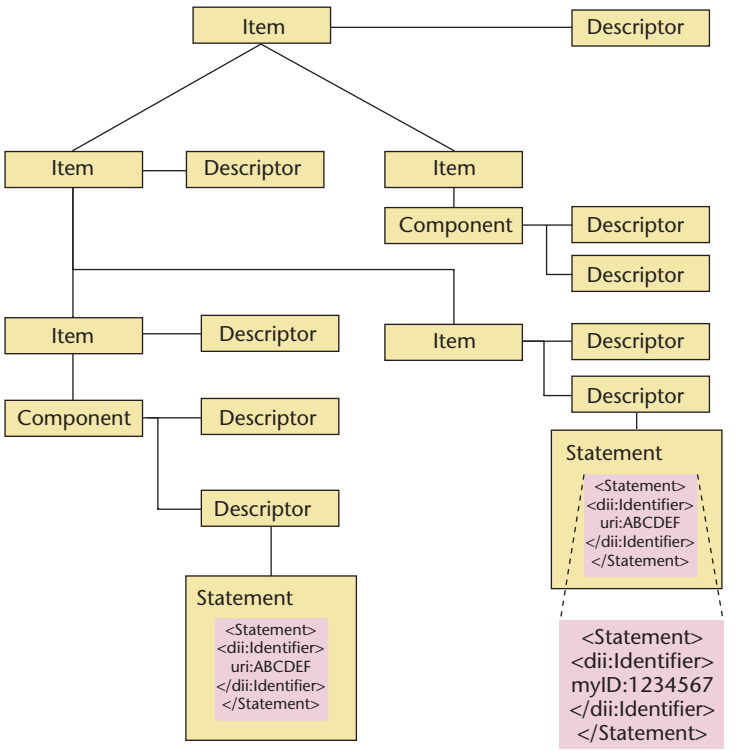

Figure 4. Relationship between the MPEG-21 Digital Item Declaration (gold boxes) and Digital Item Identification (pink boxes) parts.

MPEG-21 ongoing activities

Besides the MPEG-21 parts previously discussed, work on other MPEG-21 parts has begun.

Digital Item Adaptation

This seventh part of MPEG-21 (ISO/IEC 21000-7) specifies tools for the adaptation of Digital Items.[11]

One of the goals of MPEG-21 is to achieve interoperable transparent access to (distributed) advanced multimedia content by shielding users from network and terminal installation, management, and implementation issues. To achieve this goal requires the adaptation of Digital Items (see Figure 5, next page). As shown in this conceptual architecture, a Digital Item may be subject to a resource adaptation engine, a description adaptation engine, or a DID adaptation engine, which produces the adapted Digital Item. For the adaptation, it's essential to have available not only the description of the content but also a description of its format and of the usage environment so that content adaptation may be performed to provide the user the best content experience for the content requested. While MPEG-7[2] dealt with the

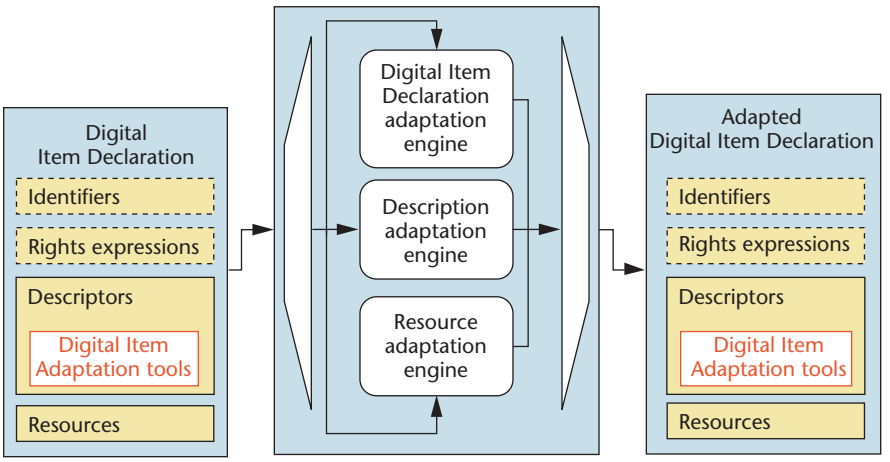

Figure 5. Digital Item Adaptation architecture.

content description problem, the description of content format and usage environments wasn't addressed, although it's now the target of the MPEG-21 DIA specification.

Recently, multimedia adaptation has become a reality in the wireless application protocol and 3rd Generation Partnership Project mobile communication arenas. However, we'll need to eventually focus on problems of adaptation for a broader range of content and conditions. Since MPEG-21 Digital Items are a package of varied multimedia resources, this leads to the problem of adapting single multimedia resources and maximizing the user experience of complete resource packages in a variety of contexts.[12]

Digital Item Processing

One of the important aspects related to the Digital Item concept is that the DID is a static declaration of only the information (structure, resources, and metadata) in the item; in the declaration itself there's no implied processing of that information. This is important because it contrasts with, for example, HTML-based Web pages where the presentation of the information is intermingled with the information itself. While the separation of the information from the usage of the information in Digital Items has significant advantages (different users can receive different presentations of the same information), it also means that on receipt of the static declaration, a user has nothing that indicates how the information should be processed.

Thus, DIP is the area of MPEG-21 that encompasses all the aspects of processing a static Digital Item from a user (and application) perspective.[13] It includes such processing as downloading the Digital Item, requesting IPMP and rights processing of the Digital Item, downloading and presenting individual media resources, and performing tasks (methods) that an author requires for a particular Digital Item. While Digital Item processing is still in the early stages, Digital Item methods have been defined in more detail.

The concept of Digital Item methods is that, on receipt of a Digital Item, users will have a list of processes (methods) available that they can apply to the Digital Item. These methods provide a mechanism for a user to specify a preferred set of procedures by which the Digital Item should be handled. A simple illustration of this functionality would be a music album Digital Item with an add track method. The method would let a user add a new track to the Digital Item in the preferred format (that is, in the correct place in the hierarchy and with the correct descriptors). MPEG intends for the methods to be run on the Digital Item method engine (DIME), which supports standard base operations (Digital Item base operations, which are analogous to a programming language's standard library of functions). Users could then create methods for their Digital Items from these base operations by using a standardized syntax (Digital Item Method Language). Figure 6 shows how the DIME will be used as the core of an MPEG-21 processing node. The figure indicates the interfaces between the DIME and various parts of MPEG-21, the user, and the platform's facilities.

Outlook

Based on the MPEG requirements[14] and the enhanced requirements resulting from new usage cases,[4] MPEG-21 continues to standardize relevant new technologies, and for more investigative areas, generate technical reports. The following issues are under consideration:

- Definition of an interoperable framework for IPMP[15] that improves and adds to already available IPMP tools defined within the context of the MPEG-4 standard.
- Evaluation of methods for persistent association technologies (MPEG-21 part 11).[16]
- Definition of MPEG-21 reference software in general and a software test bed for MPEG-21

resource delivery in particular (MPEG-21 parts 8 and 12).[17]

- Definition of an MPEG-21 file format.[14]
- Specification of MPEG-21 event reporting mechanisms that will monitor and communicate among peers and users the events relating to Digital Items and/or the programs and devices that operate on them at any given time.[18] This will let users, for example, monitor and record the use of Digital Items, resources, and their descriptions.
- Exploration of requirements and technology for highly scalable audio and video coding. In this context, MPEG is considering how these developments can be optimally aligned with MPEG-21 in general and with MPEG-21 DIA in particular.[19]

Opportunities for technologies

One of the key aspects of MPEG-21 is that it's a standard framework and not a complete, implementable solution. Hence, the multimedia and signal processing communities have many opportunities to use novel techniques and solutions within the framework. Part 7 provides an example of this: DIA.[11] This part of the standard gives the community a rich set of metadata that describes the context of delivery of resources (for example, terminal, network, natural environment, and personal preference information). We can use these as input to signal processing algorithms that adapt the media itself or higher level software that chooses appropriate media streams based on content and context metadata.[20] In all these spaces, MPEG has left the algorithms and the innovation open to researchers and developers while providing the standardized infrastructure, since they don't impact interoperability.

Hence, while a number of parts of MPEG-21 are in place, and others are following close behind in the standardization timetable, researchers and developers should find the framework a useful place to start when building new and innovative research tools. The advantage of creating compatible solutions with the MPEG-21 framework is the broad market base that instantly becomes available from interoperability.

Conclusion

Although electronic content creation, distribution, consumption, and trade are already possible today, no current end-to-end solutions let different user communities interact in an interoperable and efficient way.

MPEG-21 guarantees such interoperability by focusing on how the elements of a multimedia application infrastructure should relate, integrate, and interact. Key MPEG-21 parts have already reached the status of international standard (or will reach it during 2003) while new technologies continue to be considered for inclusion in new MPEG-21 parts. The strength of MPEG-21 is that it's a standard grown from clear multimedia industry requirements for interoperability. As such, the requirements, existing standards, and current work are firmly based on usage cases and scenarios provided by that industry. **MM**

Acknowledgments

We'd like to thank all the MPEG-21 members for their interesting and fruitful discussions, which have substantially enriched our technical knowledge. Because we wrote parts of this article on the basis of the official MPEG documents, the MPEG community is in a way a coauthor.

References

1. F. Pereira and T. Ebrahimi, ed., *The MPEG-4 Book*, Prentice Hall, 2002.
2. B.S. Manjunath, P. Salembier, and T. Sikora, eds., *Introduction to MPEG-7: Multimedia Content Description Language*, John Wiley & Sons, 2002.
3. MPEG Requirements Group, *MPEG-21 Overview*, ISO/MPEG N5231, Oct. 2002.
4. MPEG Requirements Group, *MPEG-21 Use Case Scenarios*, ISO/MPEG N4991, July 2002.
5. *Multimedia Framework—Part 1: Vision, Technologies and Strategy*, ISO/IEC 21000-1:2002, 2002, http://www.iso.ch/iso/en/ittf/PubliclyAvailableStandards.
6. *Information Technology—Multimedia Framework (MPEG-21)—Part 2: Digital Item Declaration*, ISO/IEC 21000-2:2003, Mar. 2003.
7. *Information Technology—Multimedia Framework (MPEG-21)—Part 3: Digital Item Identification*, ISO/IEC 21000-3:2003, Mar. 2003.
8. *Information and Documentation—International Standard Recording Code (ISRC)*, ISO 3901:2001, 2001.
9. MPEG MDS Group, *MPEG-21 Multimedia Framework, Part 5: Rights Expression Language* (Final Draft Int'l Std.), ISO/MPEG N5939, July 2003.
10. MPEG MDS Group, *MPEG-21 Multimedia Framework, Part 6: Rights Data Dictionary* (Final Draft Int'l Std.), ISO/MPEG N5842, July 2003.
11. MPEG MDS Group, *MPEG-21 Multimedia Framework, Part 7: Digital Item Adaptation* (Final

Committee Draft), ISO/MPEG N5845, July 2003.
12. F. Pereira and I. Burnett, "Universal Multimedia Experiences for Tomorrow," *IEEE Signal Processing* (Special Issue on Universal Multimedia Access), vol. 20, no. 2, Mar. 2003, pp. 63-73.
13. MPEG MDS Group, *MPEG-21 Multimedia Framework, Part 10: Digital Item Processing* (Working Draft), ISO/MPEG N5855, July 2003.
14. MPEG Requirements Group, *MPEG-21 Requirements*, ISO/MPEG N5873, July 2003.
15. MPEG Requirements Group, *MPEG-21 Architecture, Scenarios and IPMP Requirements*, ISO/MPEG N5874, July 2003.
16. MPEG Requirements Group, *MPEG-21 Multimedia Framework, Part 11: Evaluation of Persistent Association Tools* (Working Draft), ISO/MPEG N5875, July 2003.
17. MPEG Requirements Group, *MPEG-21 Multimedia Framework, Part 12: Resource Delivery Test Bed* (Working Draft), ISO/MPEG N5640, July 2003.
18. MPEG Requirements Group, *Current Vision on Event Reporting in MPEG-21*, ISO/MPEG N5871, July 2003.
19. MPEG Requirements Group, *Scalable Video Coding Requirements*, ISO/MPEG N5880, July 2003.
20. P. van Beek, et al., "Metadata Driven Multimedia Access," *IEEE Signal Processing* (Special Issue on Universal Multimedia Access), vol. 20, no. 2, Mar. 2003, pp. 40-52.

Ian Burnett is an associate professor at the University of Wollongong, Australia. His current research interests are in multimedia processing and delivery, speech and audio coding, 3D audio, and audio separation. He received his BSc, MEng, and PhD in electrical and electronic engineering from the University of Bath, UK.

Rik Van de Walle is a professor of multimedia systems and applications and head of the Multimedia Lab at Ghent University, Belgium. His current research interests include multimedia content delivery, presentation and archiving, coding and description of multimedia data, content adaptation, and interactive (mobile) multimedia applications. He received his MSc and PhD in engineering from Ghent University.

Keith Hill is a founding partner of Rightscom, London, a strategy consultancy for digital content. He has contributed to the major standards activities related to management and protection of intellectual property and identification systems for rights holders. He was a primary contributor to the launch of MPEG-21 in 1999 and continues to play a leading role in defining standards that enable the secure delivery of content within a multimedia framework. He has a BMus (honors) from the University of Sheffield.

Jan Bormans heads the Interuniversity Microelectronic Center's Multimedia Image Compression Systems group in Leuven, Belgium. His main research interests are the efficient design and implementation of embedded systems for advanced multimedia applications. He's the Belgian head of delegation for ISO/IEC's MPEG and SC29 standardization committees. H's also the MPEG-21 requirements editor and chairman of the MPEG liaison group. He earned his PhD in applied sciences at the Vrije Universiteit Brussel (VUB), Belgium.

Fernando Pereira is an associate professor of telecommunications at the Instituto Superior Técnico, Lisbon, Portugal. His current research interests are in video analysis, processing, coding and description, and multimedia interactive services. He has been participating in the work of ISO/MPEG for many years, notably as the head of the Portuguese delegation, chairman of the MPEG Requirements group, and by chairing many ad hoc groups related to the MPEG-4 and MPEG-7 standards. He received his MSc and PhD in electrical and computer engineering from Instituto Superior Técnico.

Readers may contact Fernando Pereira at Instituto Superior Técnico, Instituto de Telecomunicações, Av. Rovisco Pais, 1049-001 Lisboa, Portugal; fp@lx.it.pt.