M

MPEG-21 MULTIMEDIA FRAMEWORK

Christian Timmerer and Hermann Hellwagner
Department of Information Technology, Klagenfurt University, Austria

Synonym: *Multimedia Framework; ISO/IEC 21000.*

Definition: *The MPEG-21 Multimedia Framework enables transparent and augmented use of multimedia resources across a wide range of networks, devices, user preferences, and communities, notably for trading (of bits).*

Vision, Strategy, and Tools

The aim of the MPEG-21 [1-24] standard (ISO/IEC 21000), the so-called Multimedia Framework, is to enable transparent and augmented use of multimedia resources across a wide range of networks, devices, user preferences, and communities, notably for trading (of bits). As such, it represents a new step in MPEG's standards evolution, namely to support the *transaction of Digital Items among Users*.

A *Digital Item* is a structured digital object with a standard representation and metadata. As such, it is the fundamental unit of transaction and distribution within the MPEG-21 multimedia framework. In order words, it aggregates multimedia resources together with metadata, licenses, identifiers, intellectual property management and protection (IPMP) information, and methods within a standardized structure.

A *User* (please note the upper case "U") is defined as any entity that interacts within this framework or makes use of Digital Items. It is important to note that Users may include individuals as well as communities, organizations, or governments, and that Users are not even restricted to humans, i.e., they may also include intelligent software modules such as agents.

The MPEG-21 standard currently comprises 18 parts which can be clustered into **six major categories** each dealing with different aspect of the Digital Items: *declaration (and identification)*, *rights management*, *adaptation*, *processing*, *systems*, and *miscellaneous aspects* (i.e., reference software, conformance, etc.) which are described in the following.

Declaring and Identifying Digital Items

A Digital Item is a structured digital object with a standard representation, identification, and metadata. The standard representation of Digital Items is defined by a model which describes a set of abstract terms and concepts and is expressed by the XML Schema based Digital Item Declaration Language (DIDL) [4]. The resulting XML document conformant to DIDL is called Digital Item Declaration (DID). The DID may contain several building blocks as defined in DIDL which defines the structure of the Digital Item. A brief overview of the most important building blocks is given in this paper, for further details the reader is referred to [2-4].

The *Item* comprises a grouping of sub-items or components. In general, an item can be considered as a declarative representation of a Digital Item. Note that an item without sub-items can be considered a logically indivisible work and an item that does contain sub-items can be considered a compilation.

The *Component* defines a binding of a multimedia resource to a set of descriptors which provides information related to all or parts of the resource. These descriptors will typically contain control or structural information about the resource such as bit rate, character set, start points, or encryption information.

A *Descriptor* associates information with the enclosing element, i.e., its parent (e.g., item) or following sibling (e.g., component). The information can itself be a component (e.g., thumbnail of an image) or a textual statement.

A *Resource* is defined as an individually identifiable asset such as a video, audio clip, image, or textual asset. Note that the resource must be locatable via an unambiguous address.

Digital Items are configurable through the so-called choice/selection mechanism. A *Choice* describes a set of related *Selection*s which can affect the configuration of an item. As such it provides a generic and flexible way for multimedia content selection based on certain criteria defined by the Digital Item author. Such criteria may include rights expressions and/or usage environment constraints.

Another important aspect of MPEG-21 is the identification of Digital Items. The Digital Item Identification (DII) standard provides means for uniquely identifying DIs and parts thereof [4][10]. However, it is important to emphasize that DII does not define yet another identification scheme; in fact, DII facilitates existing schemes such as International Standard Book Number (ISBN) or International Standard Serial Number (ISSN) and specifies means for establishing a registration authority for Digital Items.

An example DID is shown in Figure 1. The Digital Item is appropriately identified utilizing DII and provides three selections. Note that each of the three selections may contain further DIDL elements with more detailed information regarding each selection but this is omitted here due to space limitations. The sub-items conditionally refer to one of the selection identifiers and comprise the actual reference to the media resource.

```
<DIDL xmlns="urn:mpeg:mpeg21:2002:01-DIDL-NS"
  xmlns:dii="urn:mpeg:mpeg21:2002:01-DII-NS">
 <Item>
  <Descriptor>
   <Statement>
    <dii:Identifier>
     urn:uma:tutorial:digital_item:T800
    </dii:Identifier>
   </Statement>
  </Descriptor>
  <Choice minSelections="0" maxSelections="1">
   <Selection select_id="TERMINATOR_1"/>
   <Selection select_id="TERMINATOR_2"/>
   <Selection select_id="TERMINATOR_3"/>
  </Choice>
  <Item id="Terminator_1_Item">
   <Condition require="TERMINATOR_1"/>
   <Component>
    <Resource mimeType="video/mpeg"
     ref="rtsp://my.movies.com/arni/terminator1.mpeg"/>
   </Component>
  </Item>
  <Item id="Terminator_2_Item">
   <Condition require="TERMINATOR_2"/>
   <Component>
    <Resource mimeType="video/mpeg"
     ref="rtsp://my.movies.com/arni/terminator2.mpeg"/>
   </Component>
  </Item>
  <Item id="Terminator_3_Item">
   <Condition require="TERMINATOR_3"/>
   <Component>
    <Resource mimeType="video/mpeg"
     ref="rtsp://my.movies.com/arni/terminator3.mpeg"/>
   </Component>
  </Item>
 </Item>
</DIDL>
```

Figure 1. Example DID declaring a Digital Item of the fictive Terminator trilogy.

Expressing Rights

Digital Rights Management (DRM) support within MPEG-21 can be divided into three parts, namely the *Rights Expression Language (REL)*, the *Rights Data Dictionary (RDD)* and *IPMP Components*.

The *REL* is a machine-readable language that can declare rights and permissions on digital resources [5][12]. The main goals of REL can be formulated as supporting guaranteed end-to-end interoperability by providing a standard way to express rights/interests and a standard way to express grants of rights. The former is used for protection of digital content as well as privacy and use of personal data. The latter specifies the access and use of controls for digital content by honoring the rights, conditions, and fees specified in the rights expressions. The REL data model is shown in Figure 2 which contains four basic entities. The right defines the action (or activity) or a class of actions that a principal may perform on or using the associated resource under given conditions, e.g., time, fee, count, territory, freshness, integrity, marking, signed-by, and so forth.

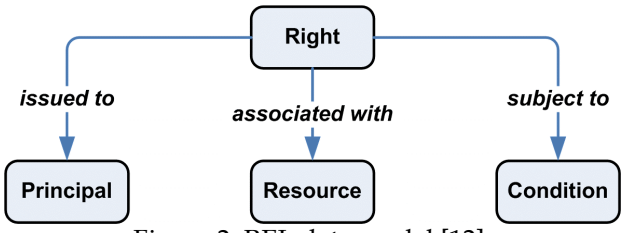

Figure 2. REL data model [12].

The *RDD* comprises a set of clear, consistent, structured, integrated, and uniquely identified terms to support REL [5][13]. The goals of the RDD are twofold. On the one hand the RDD provides a standard way to describe the semantics of terms based on their relations to other terms. On the other hand, the RDD supports mapping/transformation of metadata from the terminology of one namespace (or authority) into that of another namespace (or authority).

The *IPMP components* specify how to include IPMP information and protected parts of Digital Items in a DIDL document [11]. It deliberately does not include protection measures, keys, key management, trust management, encryption algorithms, certification infrastructures or other components required for a complete DRM system. Currently, the IPMP components consist of two parts, the IPMP DIDL providing a protected representation of the DID model, and IPMP information schemes defining structures for expressing information relating to the protection of content including tools, mechanisms, and licenses.

Adaptation of Digital Items

A vital and comprehensive part within MPEG-21 is Part 7 of the standard, referred to as Digital Item Adaptation (DIA), which specifies normative description tools to assist with the adaptation of Digital Items [6][14]. In particular, the DIA standard specifies means enabling the construction of device and coding format independent adaptation engines.

The tools allowing for *device independence* are generally referred to as Usage Environment Description (UED) tools which include terminal capabilities and network characteristics as well as user characteristics and the characteristics of the natural environment. Such descriptions provide a fundamental input to any multimedia adaptation engine.

In order to cope with today's diversity of existing scalable coding formats, e.g., MPEG-4 or JPEG2000, a generic adaptation approach for these coding formats is desirable, i.e., an approach that works *independently of the coding format*. DIA's response to this desire is the Bitstream Syntax Description (BSD) tool providing means for describing the high-level syntax of a media bitstream, e.g., how the stream is organized in terms of frames, layers, or packets, utilizing the Extensible Markup Language (XML). Therefore, the Bitstream Syntax Description Language (BSDL) based on XML Schema defines restrictions and extensions taking the structure of multimedia formats into account.

The actual bitstream adaptation based on BSDs can be divided into two logical steps. The first step transforms the BSD (e.g., using the Extensible Stylesheet Language for Transformations (XSLT)) according to the parameters derived from the usage

environment properties. The second step adapts the bitstream by means of the transformed BSD according to the definition of the BSDtoBin processor as specified in the DIA standard. Please note that both steps can be and should be combined for efficiency reasons.

However, the BSD-based adaptation approach is only one step towards coding format independence. An integral part of media adaptation is providing the optimal adaptation parameters with respect to the UED, taking into account QoS information of the multimedia content. Therefore, DIA specifies two tools that meet the above requirements, namely the *AdaptationQoS* and *Universal Constraints Description (UCD)* tools. AdaptationQoS specifies the relationship between, for example, device constraints, feasible adaptation operations satisfying these constraints, and associated utilities (or qualities) of the multimedia content. The UCD enables users to specify further constraints on the usage environment and the use of a Digital Item by means of limitation and optimization constraints; e.g., the UED might describe a 1,280 x 1,024 pixel resolution display and the UCD constrains this further by informing the adaptation engine that only 70% of this area is available while the frame width and height of the multimedia content should be maximized.

Processing of Digital Items

The declaration of a Digital Item defines its structure, but still a DID is static. The question what happens when a DID arrives at a terminal remains unanswered so far. The Digital Item Processing (DIP) standard, MPEG-21 Part 10 [7][17], allows Users to add functionality to a DID: on receipt of a DID a list of Digital Item Methods (DIMs) that can be applied to the Digital Item is presented to the User. The User chooses a Method which is then executed by the DIM Engine (DIME).

DIMs provide a way for Users of a Digital Item to select preferred procedures by which the Digital Item should be handled. Note this is done at the level of the Digital Item itself and it is not intended to be utilized for implementing the processing of media resources themselves. As such, DIMs are basically a "list of operations" specified in the normative DIM Language (DIML) for which ECMAScript has been selected. DIMs may utilize Digital Item Base Operations (DIBOs) or Digital Item eXtension Operations (DIXOs). The former are normative basic operations on which DIMs are built analogous to a standard library of functions of a programming language. DIBOs are atomic operations and are defined using a normative, high-level interface. The latter are used as an extension mechanism enabling interoperable execution of user-defined operations for which bindings to Java and C++ exist.

MPEG-21 Systems Aspects

The MPEG-21 systems aspects include the MPEG-21 File Format (.mp21), the Binary Format for XML-based MPEG-21 descriptions, and Digital Item Streaming (DIS).

The DID declares a Digital Item using XML which includes textual assets such as metadata or licenses and references to the actual multimedia content. Nonetheless, the DID is declarative and does not provide a physical container including all the assets of a Digital Item. Additionally, it is not possible to embed binary content within XML in an

efficient way – base64 encoding results in approximately 33% overhead. The *File Format* [16] has been defined to solve these issues and is based on the ISO base media file format [25] which has been extended by specific MPEG-21 requirements such as the 'meta' box with the mp21 metadata handler.

Another important issue for MPEG-21 is the efficient storage and transport of MPEG-21-based metadata. Therefore, MPEG's Binary Format for Metadata (BiM) [26] has been adopted for the *Binary Format* of the MPEG-21 standard [22]. BiM defines an alternative schema aware XML serialization format which adds streaming capabilities to XML documents, among other useful features.

Finally, the incremental delivery of a Digital Item in a piece-wise fashion and with temporal constraints such that a receiving peer may incrementally consume the Digital Item has been acknowledged within MPEG-21. Therefore, the *Digital Item Streaming* [24] standard provides answers on how to fragment XML documents which are heavily used within MPEG-21, how to assign time information to those XML fragments, and how to specify the streaming of different components of a Digital Item.

Miscellaneous

The remaining MPEG-21 parts that do not fall into one of the previous categories are briefly outlined here:

- **Part 8: Reference Software** [15]: The reference software of ISO/IEC 21000 serves three main purposes:
  - Validation of the written specification of the MPEG-21 parts.
  - Clarification of the written specification of the MPEG-21 parts.
  - Conformance testing for checking interoperability for the various applications – which aims to be compliant to MPEG-21 – against the reference software.

  In other words, the reference software implements the normative parts of the written specifications. Additionally, so-called utility software is provided which demonstrates how to use the reference software in various application scenarios.
- **Part 11: Evaluation Tools for Persistent Association** [18]: The term "persistent association" is used to categorize all the techniques for managing identification and description with content. This includes the carriage of identifiers within the context of different content file and transport formats, including file headers and embedded into content as a watermark. It also encompasses the ability for identifiers associated with content to be protected against their unauthorized removal and modification.
- **Part 12: Test Bed for MPEG-21 Resource Delivery** [19]: The test bed is mainly composed of a streaming player, a media server, and an IP network emulator. This document describes the API of each components of the test bed to facilitate a component oriented development process. This platform provides a flexible and fair test environment for evaluating scalable media streaming technologies for MPEG contents over IP networks
- **Part 14: Conformance Testing** [20]: This part represents the conformance testing specification for various MPEG-21 parts.
- **Part 15: Event Reporting** [8][21]: This standard specifies:

  - How to express Event Report Requests (ER-R) that contain information about which events to report, what information is to be reported and to whom;
  - How to express Event Reports (ER) which are created by an MPEG-21 Peer in response to an Event Report Request when the condition of the ER-R is met; and
  - How to use Digital Items – as specified in Part 2 of MPEG-21 – to package Event Reports and Event Report Requests.
- **Part 17: Fragment Identification for MPEG Resources** [23]: This part specifies a normative syntax for URI Fragment Identifiers to be used for addressing parts of any resource whose Internet Media Type is one of: audio/mpeg or video/mpeg.

Note that the purpose of Part 13 of MPEG-21 was to define new Scalable Video Coding (SVC) technology with high compression performance. Examples of potential applications that can benefit from such improved scalable coding technologies are: Internet video, wireless LAN video, mobile wireless video for conversational, VoD, and live broadcasting purposes, multi-channel content production and distribution, surveillance-and-storage applications, and layered protection of contents. However, during the course of the development of SVC it turned out that it will become an extension of Advanced Video Coding (AVC) and, thus, SVC has been moved to MPEG-4, Part 10 [27][28].

Concluding Remarks

The aim of MPEG-21 was – as the name implies – to standardize a multimedia framework for ubiquitous multimedia consumption across heterogeneous applications and domains, e.g., networks and devices. All parts of MPEG-21 address a distinctive set of requirements which allows implementers of the standard to design and implement a system or application that goes beyond simple multimedia content delivery in an interoperable way. In practice, however, it has turned out that only relatively small portions of the whole framework have been adopted by industry so far. This ultimately leads to the question whether MPEG has addressed the requirements in a vital way and what needs to be done to foster adoption of the MPEG-21 concepts on a broader scale. Efforts are underway to bring MPEG-21 into the mainstream of multimedia technologies, yet these questions remain open.

See: Universal Multimedia Access, MPEG-21 Digital Item Adaptation, Multimedia Content Adaptation in MPEG-21

References

1. ISO/IEC TR 21000-1:2004, *Information technology — Multimedia framework (MPEG-21) — Part 1: Vision, Technologies and Strategy*, 2004.
2. I.S. Burnett, F. Pereira, R. V. de Walle, and R. Koenen, Eds., *The MPEG-21 Book*, John Wiley & Sons, 2006.
3. F. Pereira, J. R. Smith, and A. Vetro, Eds., *IEEE Transaction on Multimedia, Special Section on MPEG-21*, vol. 7, no. 3, June 2005.
4. I.S. Burnett, S.J. Davis, G.M. Drury, "MPEG-21 Digital Item Declaration and Identification—Principles and Compression", *IEEE Transactions on Multimedia*, vol. 7, no. 3, pp. 400–407, June 2005.

5. X. Wang, T. DeMartini, B. Wragg, M. Paramasivam, and C. Barlas, "The MPEG-21 Rights Expression Language and Rights Data Dictionary", *IEEE Transactions on Multimedia*, vol. 7, no. 3, pp. 408-417, June 2005.
6. A. Vetro and C. Timmerer, "Digital Item Adaptation: Overview of Standardization and Research Activities", *IEEE Transactions on Multimedia*, vol. 7, no. 3, pp. 418–426, June 2005.
7. F. De Keukelaere, S. De Zutter, and R. Van de Walle, "MPEG-21 Digital Item Processing", *IEEE Transactions on Multimedia*, vol. 7, no. 3, pp. 427–434, June 2005.
8. A. Tokmakoff, F.-X. Nuttall, and K. Ji, "MPEG-21 Event Reporting: Enabling Multimedia E-Commerce", *IEEE MultiMedia Magazine*, vol. 12, no. 4, pp. 50-59, October-December 2005.
9. ISO/IEC 21000-2:2005, *Information technology – Multimedia framework (MPEG-21) – Part 2: Digital Item Declaration*, 2nd edition, 2005.
10. ISO/IEC 21000-3:2003, *Information technology – Multimedia framework (MPEG-21) – Part 3: Digital Item Identification*, 2003.
11. ISO/IEC 21000-4:2006, *Information technology – Multimedia framework (MPEG-21) – Part 4: Intellectual Property Management and Protection Components*, 2006.
12. ISO/IEC 21000-5:2004, *Information technology – Multimedia framework (MPEG-21) – Part 5: Rights Expression Language*, 2004.
13. ISO/IEC 21000-6:2004, *Information technology – Multimedia framework (MPEG-21) – Part 6: Rights Data Dictionary*, 2004.
14. ISO/IEC 21000-7:2007, *Information technology – Multimedia framework (MPEG-21) – Part 7: Digital Item Adaptation*, 2nd edition, 2007.
15. ISO/IEC 21000-8:2006, *Information technology – Multimedia framework (MPEG-21) – Part 8: Reference Software*, 2006
16. ISO/IEC 21000-9:2005, *Information technology – Multimedia framework (MPEG-21) – Part 9: File Format*, 2005.
17. ISO/IEC 21000-10:2006, *Information technology – Multimedia framework (MPEG-21) – Part 10: Digital Item Processing*, 2006.
18. ISO/IEC TR 21000-11:2004, *Information technology – Multimedia framework (MPEG-21) – Part 11: Evaluation Tools for Persistent Association Technologies*, 2004.
19. ISO/IEC TR 21000-12:2005, *Information technology – Multimedia framework (MPEG-21) – Part 12: Test Bed for MPEG-21 Resource Delivery*, 2005.
20. ISO/IEC 21000-14:2007, *Information technology – Multimedia framework (MPEG-21) – Part 14: Conformance Testing*, 2007.
21. ISO/IEC 21000-15:2006, *Information technology – Multimedia framework (MPEG-21) – Part 15: Event Reporting*, 2006.
22. ISO/IEC 21000-16:2005, *Information technology – Multimedia framework (MPEG-21) – Part 16: Binary Format*, 2005
23. ISO/IEC 21000-17:2006, *Information technology – Multimedia framework (MPEG-21) – Part 17: Fragment Identification of MPEG Resources*, 2006.
24. ISO/IEC 21000-18:2007, *Information technology – Multimedia framework (MPEG-21) – Part 18: Digital Item Streaming*, 2007.
25. ISO/IEC 14496-12:2005, *Information technology – Coding of audio-visual objects – Part 12: ISO base media file format*, 2005.
26. ISO/IEC FCD 23000-1, *Information technology – MPEG-B – Part 1: Binary MPEG format for XML*, N7251, Busan, Korea, April 2005.

27. T. Wiegand, G. J. Sullivan, G. Bjontegaard, and A. Luthra, "Overview of the H.264/AVC Video Coding Standard", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, no. 7, pp. 560-576, July 2003.
28. T. Wiegand, G. J. Sullivan, J.-R. Ohm, and A. K. Luthra, "Introduction to the Special Issue on Scalable Video Coding—Standardization and Beyond", *IEEE Transactions on Circuits and Systems for Video Technology, Special Issue on Scalable Video Coding*, vol. 17, no. 9, pp. 1099-1102, September 2007.