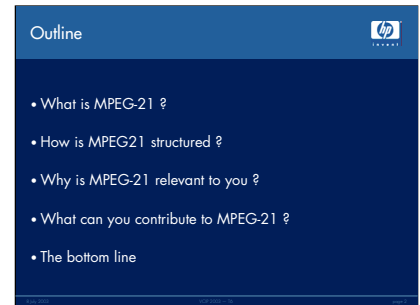


- This picture is an typical view of Lugano as you do not see it on post cards
- Contemporary Lugano is dominated by the architecture of Mario Botta and his disciples
- Palo Alto, where HP Labs is located, has several buildings by Tony Carrasco, who builds in Botta's style
 - And of course in San Francisco we have the SFMOMA built by Botta himself



- MPEG-21 is a large standard that covers a lot of terrain. In this presentation we give a high-level overview, sort of MPEG-21 as seen from outer space. The material comes from Leonardo Chiariglione's web site at






- Today is the 15th anniversary of the establishment of MPEG
- MPEG being so old, we do not need to spend many words on it and can focus on MPEG-21

What is MPEG-21 ? 



- An open framework for multimedia delivery and consumption
- Focal points:
 - Content creators
 - Content consumers

- Today, many elements exist to build an infrastructure for the delivery and consumption of multimedia content. There is, however, no 'big picture' to describe how these elements, either in existence or under development, relate to each other. The aim for MPEG-21 is to describe how these various elements fit together. Where gaps exist, MPEG-21 will recommend which new standards are required. ISO/IEC JTC 1/SC 29/WG 11 (MPEG) will then develop new standards as appropriate while other relevant standards may be developed by other bodies. These specifications will be integrated into the multimedia framework through collaboration between MPEG and these bodies.
- The result is an open framework for multimedia delivery and consumption, with both the content creator and content consumer as focal points. This open framework provides content creators and service providers with equal opportunities in the MPEG-21 enabled open market. This will also be to the benefit of the content consumer providing them access to a large variety of content in an interoperable manner.
- The vision for MPEG-21 is to define a multimedia framework to enable transparent and augmented use of multimedia resources across a wide range of networks and devices used by different communities. For a detailed examination and description of the requirements for the MPEG-21 multimedia framework readers are advised to refer to the MPEG-21 Technical Report, "Vision, Technologies and Strategy," [ISO/IEC TR 21000-1:2001(E) Part 1: Vision, Technologies and Strategy].

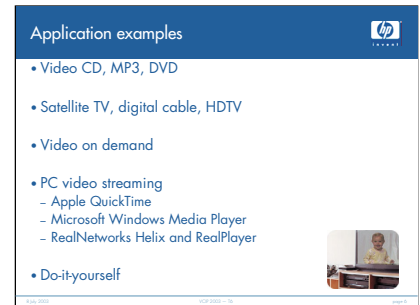


- MPEG is a working group of ISO
 - ISO/IEC JTC 1/SC 29/WG 11
 - Coding of moving pictures and audio
- Development of international standards for compression, decompression, processing and coded representation of moving pictures, audio, and their combination, in order to satisfy a wide variety of applications


Much of the material on these slides is from <http://mpeg.his.comtech.ltd.uk/>

June 2003 VCIP 2003 - T6 page 11

- MPEG, for Moving Picture Experts Group, is an informal name, the formal name being




- These are just some examples of where MPEG standards are deployed commercially
- One class of applications are individual contents packages, much like music records or video cassettes. MPEG's digital technologies allow much better quality or much more contents per retail package
 - One issue is protecting the content and managing contracts, because digital content is so easy to distribute
 - Video CD is popular mostly in Asia and virtually unknown in the US. HP's Memories Disc Creator is an example of Video CD applications
 - Compare DVD to Laser Disk
- The advanced compression of MPEG makes satellite consumer TV commercially viable
- Video on demand is taking on only very slowly and mostly in sports TV. Advanced codecs makes the technology commercially viable in the consumer market and the indexing supported by MPEG-7 allows consumer find the content they want
- PC video streaming is only new, because it requires a fast Internet connection. Being able to watch non-local news can have a big impact on government's control of the media
 - Software products are listed in alphabetical order, which according to Sam Liu is also the quality order in terms of standards conformance, implementation completeness, and bugs
 - Note that the technology is so complex that only an international collaboration can support successful commercial applications. Attempts for proprietary solutions have failed; such as for example Intel's Indeo and Philips' DivX
- Video transcoding requires about 1G Hz of computing power. Motorola/IBM/Apple's G4 technology and Intel's Pentium 4 technology have made it feasible for consumers to rip audio, manage digital photographs, and edit their own digital movies
- Picture is HP's image for the *hp memories disc creator* product


The MPEG family 

- MPEG-1 and MPEG-2 provide interoperable ways of representing audiovisual content, commonly used on digital media and on the air
- MPEG-4 defines how to *represent* content
- MPEG-7 specifies how to *describe* content
- MPEG-21 provides a truly interoperable multimedia framework

June 2000 VCIP 2003 - 10 page 11


- MPEG-1 is the standard on which such products as Video CD and MP3 are based
- MPEG-2 is the standard on which such products as Digital Television set top boxes and DVD are based
- MPEG-4 is the standard for multimedia for the fixed and mobile web
- MPEG-7 is the standard for description and search of audio and visual content
- Work on the new standard MPEG-21 "Multimedia Framework" has started in June 2000
 - Work on MPEG-1 started May 1988 and was approved as a standard in November 1992
 - MPEG-21 may become a standard in December 2004, but 2005 is more realistic
 - According to Sam Liu, historically it takes 10 years for an MPEG standard to be fully commercially deployed

MPEG-1 — ISO/IEC 11172 


- Coding of moving pictures and associated audio for digital storage media
- Video and audio at 1.5M bit/s for CD-ROM
- Five parts:
 - Part 1 (systems): multiplexing & synchronization
 - Part 2 (video): ~VHS quality at 1.15M bit/s
 - Part 3 (audio): stereo at 384K, 256K, 192K bit/s
 - Part 4 (conformance testing): references for decoder
 - Part 5 (reference software): C implementation
- Applications: Video CD, MP3 

ISO/IEC 11172-1:1993 ISO/IEC 11172-1:1993 part 1


- MPEG-1 had compact disks in mind for the medium and VHS for the quality. Note that the data rate is not a limit: MPEG-1 works well at this low data rate. However, it scales to high quality video such as HDTV. In fact, MPEG-2 is a superset of MPEG-1 and many free DVD authoring tools generate MPEG-1 files. In general, the vertical and horizontal sizes of MPEG1 bit streams can be as high as 4096x4096, and the bit rate as high as 15M bits/sec.
- Video CD is a full application of MPEG-1 that can encode a movie on two CDs
 - As of 2001, over 60 million hardware Video CD decodes have been sold worldwide
- MP3 is an abbreviation for “MPEG-1 Audio Layer III”
 - There are numerous software packages to rip tracks from a CD Audio and compress it in MP3
 - Compile your own CDs
 - Swap services like Napster and Kazaa have totally changed the music industry
 - Very profitable business of MP3 players, like iPod

MPEG-2 — ISO/IEC 13818 


- Generic coding of moving pictures and associated audio
- Digital Storage Media Command and Control (DSM-CC) for session set up and remote control of a server, used in set top boxes for satellite and cable TV
- Advanced Audio Coding (AAC) for multi-channel audio
- 4:2:2 profile for TV production studios
- Provisions for Intellectual Property Management and Protection (IPMP)
- Applications: digital TV set top boxes, DVD
- Transport Stream version
- Patent issues



- MPEG-2 is the standard on which such products as Digital Television set top boxes and DVD are based: DVD quality
- Work started in July 1990 and the first 3 parts were standardized in November 1994
- *Transport Stream* version of Part 1 supports efficient transmission over error-prone delivery systems
- *Program Stream* version is similar to MPEG-1, for digital storage media
- QuickTime is the first commercial application supporting AAC, which has much better quality at a given bit rate than MP3
- Many commercial applications, such as the *hp memories disc creator* product use MPEG-1 instead of MPEG-2 because of the patent issues. MPEG-2 decoders can decode MPEG-1, which may not require licensing patents, allowing low-cost applications to create DVDs that can be played in most DVD players
- HP product in this space: *HP Media Cache Server Appliances*, see http://www.hp.com/execcomm/inview/august02/aug02_icde4.html


MPEG-4 — ISO/IEC 14496 

- Coding of audiovisual objects
- MPEG-4 defines how to *represent* content
 - ancestry: VRML
 - interoperability of content structure
 - AFX — Animation Framework eXtension
 - XMT — textual XML format for SMIL, Web3D, etc.
 - adapt transparently to device capabilities
 - FSG — Fine Granularity Scalability
- Extensions of AAC and IPMP, Studio Profile
- MP4 and AVC file formats, multi-user environment
- Patent issues

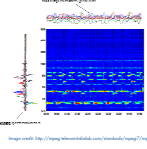


June 2002 VCIP 2003 - 10 page 10

- MPEG-4 is the standard for multimedia for the fixed and mobile web
- Delivers DVD (MPEG-2) quality video at lower data rates and smaller file sizes
- [Advanced Audio Coding \(AAC\) codec](#), provides much more efficient compression than MP3 with a quality rivaling that of uncompressed CD audio
- Scalable quality
- Scalable delivery from cell phones to satellite TV
- Started July 1994, version 1 approved October 1998, major extension approved December 1999
- File format based on QuickTime file format
- First software player supporting MPEG-4: QuickTime in August 2002
- Licensing is very complicated, requiring the mediation of licensing agencies such as MPEG-LA and Via. Licenses can be quite expensive

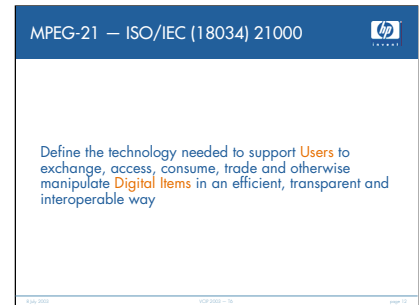
MPEG-7 — ISO/IEC 15938 

- Multimedia content description interface
- MPEG-7 specifies how to describe content
 - describe content way beyond metadata
 - facilitate content management, in particular searching

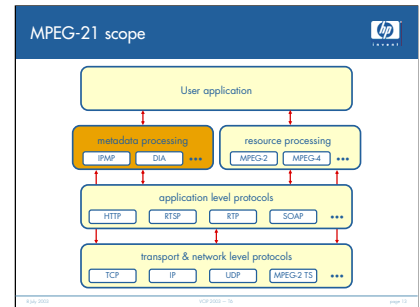


June 2003 VCIP 2003 - 10 page 11


- MPEG-7 is the standard for description and search of audio and visual content
- Unlike previous MPEG standards, it is about metadata, not data
- Work started April 1997
- Multimedia Description Schemes provides standard descriptors and description schemes that are neither visual nor audio



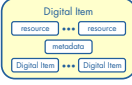
- MPEG-21 aims at defining a normative open framework for multimedia delivery and consumption for use by all the players in the delivery and consumption chain. This open framework will provide content creators, producers, distributors and service providers with equal opportunities in the MPEG-21 enabled open market. This will also be to the benefit of the content consumer providing them access to a large variety of content in an interoperable manner.
- MPEG-21 is based on two essential concepts: the definition of a fundamental unit of distribution and transaction (the Digital Item) and the concept of Users interacting with Digital Items. The Digital Items can be considered the "what" of the Multimedia Framework (e.g., a video collection, a music album) and the Users can be considered the "who" of the Multimedia Framework.
- The goal of MPEG-21 can thus be rephrased to: defining the technology needed to support Users to exchange, access, consume, trade and otherwise manipulate Digital Items in an efficient, transparent and interoperable way.
- During the MPEG-21 standardization process, Calls for Proposals based upon requirements have been and continue to be issued by MPEG. Eventually the responses to the calls result in different parts of the MPEG-21 standard (i.e. ISO/IEC 21000-N) after intensive discussion, consultation and harmonization efforts between MPEG experts, representatives of industry and other standards bodies.
- MPEG-21 identifies and defines the mechanisms and elements needed to support the multimedia delivery chain as described above as well as the relationships between and the operations supported by them. Within the parts of MPEG-21, these elements are elaborated by defining the syntax and semantics of their characteristics, such as interfaces to the elements.



- This is a simplified conceptual diagram; see MPEG-21 Part 8, slide on DI Processing for a more accurate diagram
- Data and metadata are stored, transported, and processed independently, to reduce overhead, because metadata is much smaller than data
- MPEG-21 is a framework to distribute Digital Items. It relies on other technologies where they are available
 - Examples of applications: MediaPlayer, iTunes, RealPlayer
 - Existing MPEG-1, MPEG-2, MPEG-4, etc. codecs are used, which typically are packaged in their own frameworks like QuickTime, DirectX, Helix, etc.
 - The standard Internet protocols are used, such as TCP/IP and HTTP
- The emphasis in MPEG-21 standard development is on technologies for metadata processing


MPEG-21 Digital Item 

- Structured digital objects, including a standard representation and identification, and metadata
- Fundamental unit of distribution and transaction within the MPEG-21 framework
- No further technical meaning



June 2003 VCIP 2003-10 page 10

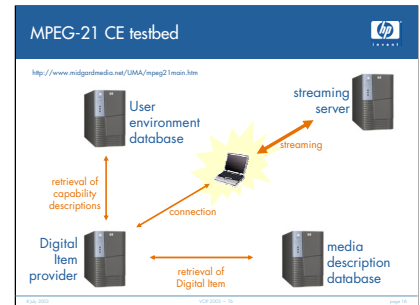
- A Digital Item is the digital representation of “a work”
- Examples:
 - MPEG-2 video stream
 - JPEG2000 image
 - MP3 track
- Note that the definition of a Digital Item is recursive: A Digital Item is a hierarchical container of resources, metadata, and other Digital Items
- The above diagram is a simplification
 - In reality the DI will not contain resources but references to resources. While the resource can be anywhere, and typically will be on a streaming server separate from the site storing the DI. However, at this point we assume the resource is part of the DI, because this is how it will appear to the User
 - The DI also contains an identifier. Because this is in Part 3, we omit it at this point
 - The REL licenses (see Part 5) are wrapped in Digital Items — this is an example of a DI in a DI
- It is not explicit which resources will actually be delivered with a Digital Item, because the metadata may contain a script that computes the particular resource, e.g., the language spoken on the audio track of a movie. A Digital Item Declaration (DID) is required to specify the makeup, structure, and organization of a DI
- It is important to differentiate between a “Digital Item” and an “item.” An item is a narrower technical term than DI and it refers to a grouping of sub-items and/or components that are bound to relevant descriptors

MPEG-21 User 

- A **User** is any entity that interacts in the MPEG-21 environment or makes use of a Digital Item
 - Users include individuals, consumers, communities, organisations, corporations, consortia, governments and other standards bodies and initiatives around the world.
- Users are identified specifically by their relationship to another User for a certain interaction
- MPEG-21 makes no distinction between a “content provider” and a “consumer” — both are Users
 - A single entity may use content in many ways
 - however, a User may assume specific or even unique rights and responsibilities according to their interaction with other Users within MPEG-21

June 2003 VCIP 2003 - 10 page 10

- At its most basic level, MPEG-21 provides a framework in which one User interacts with another User and the object of that interaction is a Digital Item commonly called content
- *Some* such interactions are creating content, providing content, archiving content, rating content, enhancing and delivering content, aggregating content, delivering content, syndicating content, retail selling of content, consuming content, subscribing to content, regulating content, facilitating transactions that occur from any of the above, and regulating transactions that occur from any of the above
- Any of these are “uses” of MPEG-21, and the parties involved are Users.



Universal Multimedia Access testbed developed at Midgard Medialab for **C**ore **E**xperiment use.

- The resource provider (entitled Digital Item Provider) will make a decision as of which variation of the resource is best suited for the user. This decision is based on the user's terminal capabilities, the environment in which the user is operating, and the available resource variations. If the user for example wants to stream a media resource (as depicted in the above figure), this decision will be especially dependent upon the available bandwidth, screen size, audio capabilities and available viewer software in the terminal. Since our testbed is supposed to simulate any User with any terminal under any network conditions, the choices are performed manually by the testbed operator.
- In practice, this would be a (potentially) automated process - the user does not have to actively take part in the decisions that are made. The capabilities and preferences would be automatically extracted by the underlying system. Since all messages and transactions between the different entities in the above diagram are in the form of Digital Items, this follows standardized syntax and semantics.

The parts of MPEG-21




1. Vision, technologies and strategies
2. Digital Item Declaration
3. Digital Item Identification
4. Intellectual Property Management and Protection (IPMP)
5. Rights Expression Language
6. Rights Data Dictionary
7. Digital Item Adaptation
8. Reference Software
9. File Format

Emphasize that MPEG-21 covers a very broad scope

- **MPEG-21 Part 1 – Vision**
 - The title "Vision, Technologies and Strategy" has been chosen to reflect the fundamental purpose of the Technical Report.
- **MPEG-21 Part 2 – Digital Item Declaration**
 - The purpose of the Digital Item Declaration (DID) specification is to describe a set of abstract terms and concepts to form a useful model for defining Digital Items. Within this model, a Digital Item is the digital representation of "a work", and as such, it is the thing that is acted upon (managed, described, exchanged, collected, etc.) within the model. The goal of this model is to be as flexible and general as possible, while providing for the "hooks" that enable higher level functionality. This, in turn, will allow the model to serve as a key foundation in the building of higher level models in other MPEG-21 elements (such as Identification & Description or IPMP). This model specifically does not define a language in and of itself. Instead, the model helps to provide a common set of abstract concepts and terms that can be used to define such a scheme, or to perform mappings between existing schemes capable of Digital Item Declaration, for comparison purposes.
- **MPEG-21 Part 3 – Digital Item Identification**
 - The scope of the Digital Item Identification (DII) specification includes:
 - > How to uniquely identify Digital Items and parts thereof (including resources);
 - > How to uniquely identify IP related to the Digital Items (and parts thereof), for example abstractions;
 - > How to uniquely identify Description Schemes;
 - > How to use identifiers to link Digital Items with related information such as descriptive metadata.
 - > How to identify different types of Digital Items.
- **MPEG-21 Part 4 – Intellectual Property Management and Protection (IPMP)**
 - The 4th part of MPEG-21 will define an interoperable framework for Intellectual Property Management and Protection (IPMP). Fairly soon after MPEG-4, with its IPMP hooks, became an International Standard, concerns were voiced within MPEG that many similar devices and players might be built by different manufacturers, all MPEG-4, but many of them not interworking. This is why MPEG decided to start a new project on more interoperable IPMP systems and tools. The project includes standardized ways of retrieving IPMP tools from remote locations, exchanging messages between IPMP tools and between these tools and the terminal. It also addresses authentication of IPMP tools, and has provisions for integrating Rights Expressions according to the Rights Data Dictionary and the Rights Expression Language.
- **MPEG-21 Part 5 – Rights Expression Language**
 - Following an extensive requirements gathering process, which started in January 2001, MPEG issued a Call for Proposals during its July meeting in Sydney for a Rights Data Dictionary and a Rights Expression Language. Responses to this Call were processed during the December meeting in Pattaya and the evaluation process established an approach for going forward with the development of a specification, expected to be an International Standard in late 2003.
- **MPEG-21 Part 6 – Rights Data Dictionary**
 - Following the evaluation of submissions in response to a Call for Proposals the specification of a Rights Data Dictionary (RDD) began in December 2001. The working draft was refined at the following three meetings and a Committee Draft published in July 2002.
- **MPEG-21 Part 7 – Digital Item Adaptation**
 - The goal of the Terminals and Networks key element is to achieve interoperable transparent access to (distributed) advanced multimedia content by shielding users from network and terminal installation, management and implementation issues. This will enable the provision of network and terminal resources on demand to form user communities where multimedia content can be created and shared, always with the agreed/contracted quality, reliability and flexibility, allowing the multimedia applications to connect diverse sets of Users, such that the quality of the user experience will be guaranteed.
- **MPEG-21 Part 8 – Reference Software**
 - The part of MPEG-21 that has most recently been identified as a candidate for specification is Reference Software. Reference software will form the first of what is envisaged to be a number of systems-related specifications in MPEG-21. Other candidates for specification are likely to include a binary representation of the Digital Item Declaration and an MPEG-21 file format.
- **MPEG-21 Part 9 – File Format**
 - An MPEG-21 Digital Item can be a complex collection of information. Both still and dynamic media (e.g. images and movies) can be included, as well as Digital Item information, meta-data, layout information, and so on. It can include both textual data (e.g. XML) and binary data (e.g. an MPEG-4 presentation or a still picture). For this reason, the MPEG-21 file format will inherit several concepts from MP4, in order to make 'multi-purpose' files possible. A dual-purpose MP4 and MP21 file, for example, would play just the MPEG-4 data on an MP4 player, and would play the MPEG-21 data on an MP21 player.

MPEG-21 Part 1 — Vision 

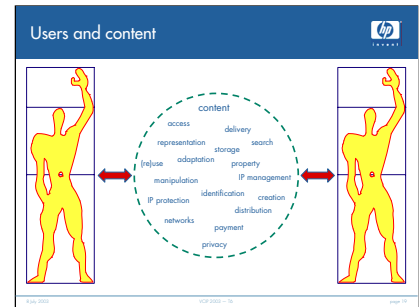
- Define a multimedia framework to enable transparent and augmented use of multimedia resources across a wide range of networks and devices
 1. Provide a vision
 2. Facilitate integration and harmonization of technologies
 3. Provide a strategy for achieving a framework through collaboration




June 2003 VCIP 2003 - 10 page 10

MPEG-21 Part 1 – Vision

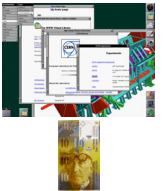
- The title “Vision, Technologies and Strategy” has been chosen to reflect the fundamental purpose of the Technical Report. This is to:
 - Define a ‘vision’ for a multimedia framework to enable transparent and augmented use of multimedia resources across a wide range of networks and devices to meet the needs of all users
 - Achieve the integration of components and standards to facilitate harmonisation of ‘technologies’ for the creation, management, transport, manipulation, distribution, and consumption of digital items.
 - Define a ‘strategy’ for achieving a multimedia framework by the development of specifications and standards based on well-defined functional requirements through collaboration with other bodies.



- MPEG-21 Part 1 sets out the User requirements in the multimedia framework
- A user is any entity that interacts with the MPEG-21 environment, regardless of the intent and function. The dual is that MPEG-21 mediates between Users to allow transaction related to Digital Items
- The granularity spans domains of several orders of magnitude, from terminals on a backbone to devices on slow wireless links, from digital cinema theatre screens to mobile phone displays, from valuable boxing matches to pragmatic MRI scans, etc.
- Ease of use is a *conditio sine qua non* for adoption of the general public; this is a requirement for interoperability and transparent interfaces, indicated by the dashed line

Digital assets 

- World Wide Web's phases
 - 1990 (info.cern.ch) — scientific exchange
 - 1995 (tidal wave) — free content
 - 2000 (dot bomb) — ubiquitous fast network
- Users are starting to recognize the value of their digital asset resources
- Markets must be efficient

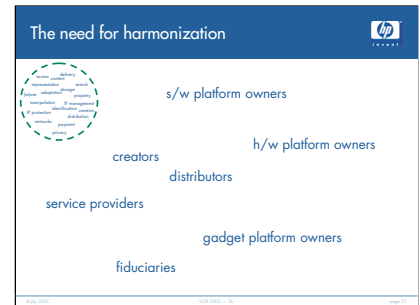


June 2002 VCIP 2003 - 10 page 10

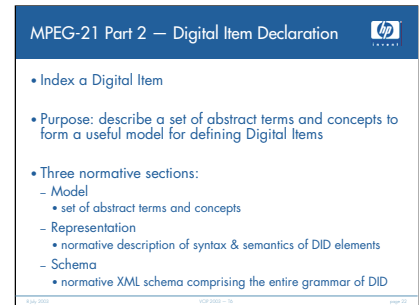
MPEG-21 Part 1 - Vision

[Illustrations. Top: the first WWW browser, by Tim Berners-Lee of CERN, implemented on a NeXT cube September 1990. Bottom: 10 CHF, on its verso you can see the modular]

- The World Wide Web started out in its current image intensive form in 1990 with info.cern.ch and the first graphical browser running on a NeXT. Contents was freely available to other scientists because they all had grants that obliged them to share and publish. The scientists were paid from the grants
- In 1995 Bill Gates declared that the WWW is a tidal wave and forced Internet interoperability on all Microsoft products. This eventually led to the dot-com bubble. The public at large adopted the Web as a source of free contents. The entrepreneurs were paid by small investors
- The Nasdaq crashed in April 2000, ending the cash inflow in the dot-com industry. However, the Internet as the infrastructure survives and large media and telecom conglomerates pick up the pieces at liquidation prices, resulting in available high-speed access at home. The dot-com companies who are able to sell their content for a profit survive.
- Using Søren Kierkegaard's terminology coined at the birth of the tabloid press in the 1850s, in the dot-com phase we saw a leveling of information, where everybody is interested in everything and nothing is too trivial or too important. People just accumulate information—postpone decisions indefinitely; nobody takes action, nobody is responsible for truth; there is no risk in action: there is no mastery, just gossip
- Kierkegaard called this the æsthetic sphere of existence, where the inability to distinguish between trivial and important leads to boredom
- His solution is to provide tools for people to evolve from the æsthetical to the ethical sphere, to empower individuals to emancipate from a leveled world to strong identities, to help people to emancipate from accumulating information to taking action and making commitments. The call is to provide tools to distill information into knowledge
- When this distillation process occurs, free content evolves into valuable digital asset resources. Users are recognizing this and MPEG-21's vision is to provide the framework to realize a new market place with low entry barriers
- The barriers are low when the markets are efficient. This is achieved by allowing providers to compete for Users



- Unlike previous MPEG-1 and 2 standards, which were mostly about codecs, a framework spans a very wide area, which no single player can own, control, or even just design
- Each player has intellectual assets or knowledge in a small number of fields — therefore MPEG-21 is not about who has the best technology, rather it is about “owners” of a field negotiating on interfaces that allow interoperability transparent to the User but leaving sufficient room for “owners” to deploy differentiating technologies that allow them to realize a profit. This process is called harmonization



MPEG-21 Part 2 — Digital Item Declaration

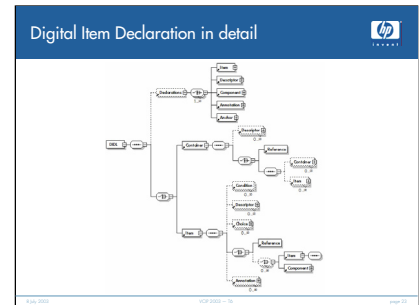
- Index a Digital Item
- Purpose: describe a set of abstract terms and concepts to form a useful model for defining Digital Items
- Three normative sections:
 - Model
 - set of abstract terms and concepts
 - Representation
 - normative description of syntax & semantics of DID elements
 - Schema
 - normative XML schema comprising the entire grammar of DID

MPEG-21 Part 2 – Digital Item Declaration

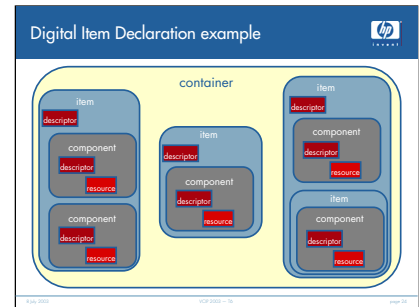
- It is not explicit which resources will actually be delivered with a Digital Item, because the metadata may contain a script that computes the particular resource, e.g., the language spoken on the audio track of a movie. A Digital Item Declaration (DID) is required to specify the makeup, structure, and organization of a DI

A Digital Item Declaration serves to index a Digital Item


- The purpose of the Digital Item Declaration (DID) specification is to describe a set of abstract terms and concepts to form a useful model for defining Digital Items. Within this model, a Digital Item is the digital representation of “a work”, and as such, it is the thing that is acted upon (managed, described, exchanged, collected, etc.) within the model. The goal of this model is to be as flexible and general as possible, while providing for the “hooks” that enable higher level functionality. This, in turn, will allow the model to serve as a key foundation in the building of higher level models in other MPEG-21 elements (such as Identification & Description or IPMP). This model specifically does not define a language in and of itself. Instead, the model helps to provide a common set of abstract concepts and terms that can be used to define such a scheme, or to perform mappings between existing schemes capable of Digital Item Declaration, for comparison purposes.
- The DID technology is described in three normative sections:
 - Model
 - Representation
 - Schema
- The goal of this model is to be as flexible and general as possible, while providing for the “hooks” that enable higher level functionality.
- This, in turn, will allow the model to serve as a key foundation in the building of higher level models in other MPEG-21 elements (such as Identification & Description or IPMP).
- This model specifically does not define a language in and of itself.
- Instead, the model helps to provide a common set of abstract concepts and terms that can be used to define such a scheme, or to perform mappings between existing schemes capable of Digital Item Declaration, for comparison purposes.
- Model:
 - The Digital Item Declaration Model describes a set of abstract terms and concepts to form a useful model for defining Digital Items.
 - Within this model, a Digital Item is the digital representation of “a work”, and as such, it is the thing that is acted upon (managed, described, exchanged, collected, etc.) within the model.
- Representation:
 - Normative description of the syntax and semantics of each of the Digital Item Declaration elements, as represented in XML.
 - This section also contains some non-normative examples for illustrative purposes.
- Schema:
 - Normative XML schema comprising the entire grammar of the Digital Item Declaration representation in XML.



- Remember that a Digital Item (DI) is the digital representation of a work (œuvre); it is the thing that is acted upon (managed, described, exchanged, collected, etc.) within the DID model.
- DIDL: Digital Item Declaration Language
- Abstract model:
 - A container is a grouping of items and containers to form logical packages; it can be labeled with a descriptor.
 - An item is a grouping of sub-items and or components bound to relevant descriptors. Items may be conditional and may contain choices. An item without sub-items is an entity, one with sub-items is a compilation. Items are declarative representations of digital items.
 - A component is a binding of a resource to a set of descriptors. The descriptors contain information about control or structure (bit rare, char set, etc.) 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 info may be a component (e.g., thumbnail) or a textual statement.
 - A condition describes the enclosing element as being optional.
 - A choice describes a set of related selections that can affect the configuration of an item.
 - A selection describes a specific decision that will affect 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.
 - 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. Fragment may be resource type specific.
 - A statement is a literal textual value that contains information, but not an asset.
 - A predicate is an unambiguously identifiable declaration that can be true, false or undecided.



- Relationship of the principle elements within the Digital Identification Declaration Model
- Containers can be used to form logical packages (for transport or exchange) or logical shelves (for organization). They are not items, they are groupings
- Items are declarative representations of Digital Items
- A DID serves to index a DI
- Note that nothing is said on the resource; unlike MPEG 1 and 2, the standard does not say anything on how a resource is decoded. The normative behaviour is that
 1. The recipient parses and validates the DID (current experimental software is very bad at this step)
 2. The recipient searches for the Rights Expression (RE)
 1. If there is no RE (or reference to RE), the DI is ungoverned, which could mean that general copyright law holds (a more strict opinion is that absent an RE no actions are permitted and no rights for any use is granted)
 2. If there is a RE (or reference to RE), the user shall process the RE and abide to it

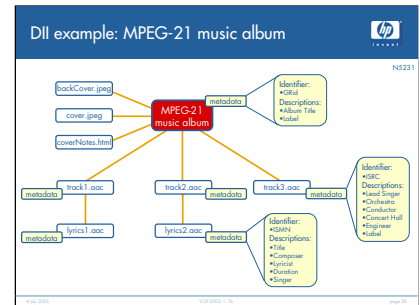
MPEG-21 Part 3 — Digital Item Identification 

The scope of the Digital Item Identification (DII) specification includes:


- How to uniquely identify Digital Items and parts thereof (including resources)
- How to uniquely identify IP related to the Digital Items (and parts thereof), for example abstractions
- How to uniquely identify Description Schemes
- How to use identifiers to link Digital Items with related information such as descriptive metadata
- How to identify different types of Digital Items

June 2003 VCIP 2003 - 10 page 11

- **MPEG-21 Part 3 – Digital Item Identification**
- Digital Items and their parts are identified by encapsulating Uniform Resource Identifiers (URI) into the Identification Description Schemes (DS)
- There is also a provision to register identification systems through the process of a Registration Authority (work in progress)
- The DII specification does not specify new identification systems for the content elements for which identification and description schemes already exist and are in use (e.g., ISO/IEC 21000-3 does not attempt to replace the ISRC (as defined in ISO 3901) for sound recordings but allows ISRCs to be used within MPEG-21)
- Identifiers covered by this specification can be associated with Digital Items by including them in a specific place in the Digital Item Declaration. This place is the STATEMENT element. Examples of likely STATEMENTS include descriptive, control, revision tracking and/or identifying information




- Example of metadata and identifiers within an MPEG-21 music album
- Source: ISO/IEC JTC1/SC29/WG11/N5231 (Shanghai 2002)
- ISMN: International Standard Music Number, ISO 10957:1993
- GRid: Global Release Identifier, see GRid Registration Authority at <http://212.134.114.163/grid/>
- ISRC: International Standard Recording Code, ISO 3901:2001
- Sources:
 - http://www.doi.org/topics/drm_paskin_20030113_b1.pdf
 - <http://www.ifla.org/Vl/3/p1996-1/sect.htm>
 - <http://www.ifpi.org/site-content/press/20030210.html>
- Different Users can have different schemes to define their content. MPEG-21 DII utilizes the XML mechanism of namespaces to support this

MPEG-21 Part 4 — IPMP 

- Improvements over MPEG-4 IPMP:
 - Internetworking
 - IPMP tool retrieval & authentication
 - Integration of Rights Expressions (RDD & REL)
- Intellectual Property Management and Protection involves the enforcement of REL permissions
 - IPMP shall consult REL before any actions are taken in the User's system
- REL: What is protected? What right applies?
IPMP: How is it protected?

June 2003 VCIP 2003 - 10 page 11


- **MPEG-21 Part 4 – Intellectual Property Management and Protection (IPMP)**
- The 4th part of MPEG-21 will define an interoperable framework for Intellectual Property Management and Protection (IPMP)
- Fairly soon after MPEG-4, with its IPMP hooks, became an International Standard, concerns were voiced within MPEG that many similar devices and players might be built by different manufacturers, all MPEG-4, but many of them not interworking
- This is why MPEG decided to start a new project on more interoperable IPMP systems and tools
- The project includes standardized ways of retrieving IPMP tools from remote locations, exchanging messages between IPMP tools and between these tools and the terminal
- It also addresses authentication of IPMP tools, and has provisions for integrating Rights Expressions according to the Rights Data Dictionary and the Rights Expression Language
- Efforts are currently ongoing to define the requirements for the management and protection of intellectual property in the various parts of the MPEG-21 standard currently under development


MPEG-21 Part 5 — REL

- Rights Expression Language
- A machine-readable language
- Can declare rights and permissions
- Uses terms defined in the Rights Data Dictionary

June 2003
MPEG-21-10

- **MPEG-21 Part 5 – Rights Expression Language**
- Following an extensive requirements gathering process, which started in January 2001, MPEG issued a Call for Proposals during its July meeting in Sydney for a Rights Data Dictionary and a Rights Expression Language. Responses to this Call were processed during the December meeting in Pattaya and the evaluation process established an approach for going forward with the development of a specification, expected to be an International Standard in late 2003.
- A Rights Expression Language is seen as a machine-readable language that can declare rights and permissions using the terms as defined in the Rights Data Dictionary.
- The REL is intended to provide flexible, interoperable mechanisms to support transparent and augmented 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 honours the rights, conditions, and fees specified for digital contents. It is also intended to support specification of access and use controls for digital content in cases where financial exchange is not part of the terms of use, and to support exchange of sensitive or private digital content.
- The Rights Expression Language is also intended to provide a flexible interoperable mechanism to ensure personal data is processed in accordance with individual rights and to meet the requirement for Users to be able to express their rights and interests in a way that addresses issues of privacy and use of personal data.
- A standard Rights Expression Language should be able to support guaranteed end-to-end interoperability, consistency and reliability between different systems and services. To do so, it must offer 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.

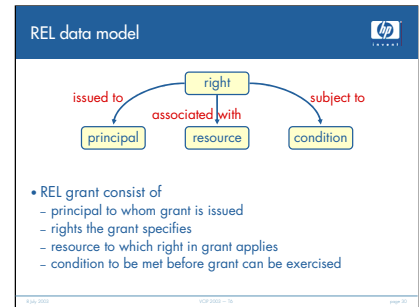
MPEG-21 Part 5 — REL 

- The Rights Expression Language consists of licenses and grants that give specific permissions to Users to perform certain actions on certain resources, given that certain conditions are met
 - Grants can also allow Users to delegate authority to others
- User's system shall parse and validate the RE
- User's system shall check permissions before any further action is done
- DID parser is responsible for discovering and identifying where to gather licenses
- REL licenses are wrapped in Digital Items


June 2003 VCIP 2003 - 10 page 10

- **MPEG-21 Part 5 – Rights Expression Language**

- Delegation of authority enables secondary markets; small commerce is supported
- Not only for commerce, but for any publishing of Digital Items, i.e., all private information can be protected
- User applications carry the burden of handling the RE, i.e., Users are assumed to be honest — how are malicious hackers dealt with? By controlling decryption
- A lot of complexity is put on the DID parser, which becomes a critical component of an MPEG-21 system
- Digital Items are sufficiently general that they can be used as wrappers, allowing for elegant implementations
- Issue: protection of privacy



- **Principal** (e.g., MPEG-21 User or Terminal) is always exactly one party, never a set
 - Identified by information unique to that party and authenticated
 - Example 1: public/private key infrastructure (PKS), certificate
 - Example 2: multiple credentials that must be simultaneously valid
- **Right** is the "verb" that a principal can be granted to exercise against some resource under some condition. Typically, a right specifies an action (or activity) or a class of actions that a principal may perform on or using the associated resource
- **Resource** (e.g., Digital Item) is the "object" to which a principal can be granted a right
 - a digital work (such as an e-book, an audio or video file, or an image)
 - a service (such as an email service, or B2B transaction service)
 - a piece of information that can be owned by a principal (such as a name or an email address)
- **Condition** specifies the terms, conditions and obligations under which rights can be exercised
 - a time interval within which a right can be exercised
 - existence of a valid, prerequisite right that has been issued to some principal
 - conditions appropriate to using digital works (e.g., watermark, destination, and renderer)

MPEG-21 Part 6 — Rights Data Dictionary 


- Set of clear, consistent, structured, integrated and uniquely identified Terms to support REL
- Specification of dictionary structure and methodology to create dictionary
- Dictionary is prescriptive, inclusive, and has audit provisions
- Legal definitions are mapped from other Authorities
- Supports mapping & transformation of metadata from terminology of one namespace (or Authority) into that of another namespace in automated or partially-automated way
- Dictionary is based on a logical model, the Context Model, which is the basis of the dictionary ontology

June 2002 2002-06-10 page 11

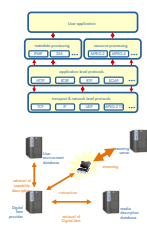
- **MPEG-21 Part 6 – Rights Data Dictionary**

- Following the evaluation of submissions in response to a Call for Proposals the specification of a Rights Data Dictionary (RDD) began in December 2001. The working draft was refined at the following three meetings and a Committee Draft published in July 2002. The following points summarise the scope of this specification:

- The Rights Data Dictionary (RDD) comprises a set of clear, consistent, structured, integrated and uniquely identified Terms to support the MPEG-21 Rights Expression Language
- The structure of the dictionary is specified, along with a methodology for creating the dictionary. The means by which further Terms may be defined is also explained
- The Dictionary 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 is 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 specification has audit provisions so that additions, amendments and deletions to Terms and their attributes can be tracked
- RDD recognises legal definitions as and only as Terms from other Authorities that can be mapped into the RDD. Therefore Terms that are directly authorized by RDD neither define nor prescribe intellectual property rights or other legal entities
- As well as providing definitions of Terms for use in the REL, the RDD specification is designed to support 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 dictionary is based on a logical model, the Context Model, which is the basis of the dictionary ontology. The model is described in detail in the specification. It is based on the use of verbs which are contextualised so that a dictionary created with it can be as extensible and granular as required

MPEG-21 Part 7 — Digital Item Adaptation 

- Goal: achieve transparent interoperable access to distributed multimedia content
- Enable ad hoc formation of User communities in which contents is shared with agreed or contracted
 - Quality
 - Reliability
 - Flexibility
 - Diversity
- Guaranteed user experience



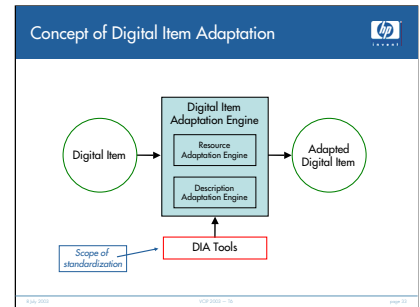
The diagram illustrates the MPEG-21 Part 7 architecture. At the top, 'User applications' are shown interacting with 'Content providers'. Below this, 'User communities' are formed, each containing 'User applications' and 'Content providers'. These communities are connected to a central network of terminals (represented by icons of a laptop, a PDA, and a mobile phone). The network is labeled 'Network' and shows bidirectional communication between terminals and the communities. The bottom of the slide contains the text 'June 2002', 'MPEG-21-7', and 'page 12'.

- **MPEG-21 Part 7 – Digital Item Adaptation**

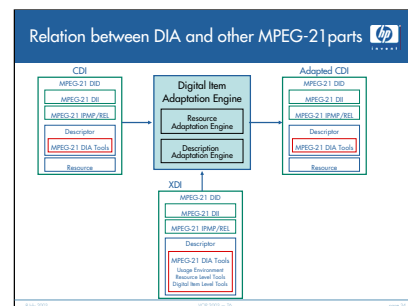
- The goal of the Terminals and Networks key element is to achieve interoperable transparent access to (distributed) advanced multimedia content by shielding users from network and terminal installation, management and implementation issues. This will enable the provision of network and terminal resources on demand to form user communities where multimedia content can be created and shared, always with the agreed/contracted quality, reliability and flexibility, allowing the multimedia applications to connect diverse sets of Users, such that the quality of the user experience will be guaranteed.

- Example

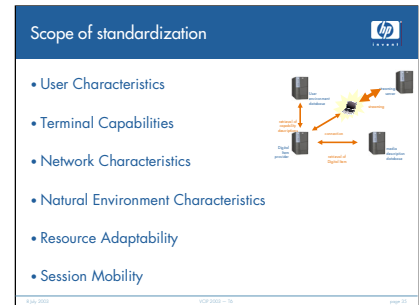
- To advertise movies by posting movie trailers on the Internet, studios typically post three version of the trailers, called Small, Medium, Large, for Users on dial-up modem, broadband modems, or connected directly to the Internet
- With MPEG-21 DIA, only a single stream needs to be stored, and the framework transcodes the bitstream in real time so that it can be rendered at optimal quality of service



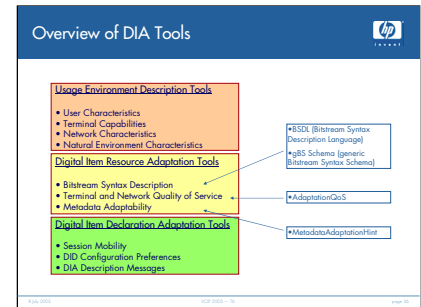
- Towards this goal the adaptation of Digital Items is required. This concept is illustrated in this. As shown in this conceptual architecture, a Digital Item is subject to a resource adaptation engine, as well as a descriptor adaptation engine, which produce together the adapted Digital Item.
- It is important to emphasise that the adaptation engines themselves are non-normative tools of Digital Item Adaptation. However, descriptions and format-independent mechanisms that provide support for Digital Item Adaptation in terms of resource adaptation, descriptor adaptation, and/or Quality of Service management are within the scope of the requirements.
- Separation of resource adaptation from description adaptation
- Note that the DIA engine is a non-normative tool: vendors compete by offering better engines
- The DIA tools that are standardized are the descriptions and format-independent mechanisms that provide support for Digital Item Adaptation in terms of resource adaptation, descriptor adaptation, and/or Quality of Service management




- CDI is a Content Digital Item: it contains a reference to a resource
- XDI is a Context Digital Item: it contains references to DIA tools



- In May 2002, a number of responses to the Call for Proposals on MPEG-21 Digital Item Adaptation were received. Based on the evaluation of these proposals, a Working Draft has been produced. The specific items targeted for standardization are outlined below
- User Characteristics: Description tools that specify the characteristics of a User, including preferences to particular media resources, preferences regarding the presentation of media resources, and the mobility characteristics of a User. Additionally, description tools to support the accessibility of Digital Items to various users, including those with audio-visual impairments, are being considered
- Terminal Capabilities: Description tools that specify the capability of terminals, including media resource encoding and decoding capability, hardware, software and system-related specifications, as well as communication protocols that are supported by the terminal
- Network Characteristics: Description tools that specify the capabilities and conditions of a network, including bandwidth utilization, delay and error characteristics
- Natural Environment Characteristics: Description tools that specify the location and time of a User in a given environment, as well as audio-visual characteristics of the natural environment, which may include auditory noise levels and illumination properties
- Resource Adaptability: Tools to assist with the adaptation of resources including the adaptation of binary resources in a generic way and metadata adaptation. Additionally, tools that assist in making resource-complexity trade-offs and making associations between descriptions and resource characteristics for Quality of Service are targeted
- Session Mobility: Tools that specify how to transfer the state of Digital Items from one User to another. Specifically, the capture, transfer and reconstruction of state information will be specified



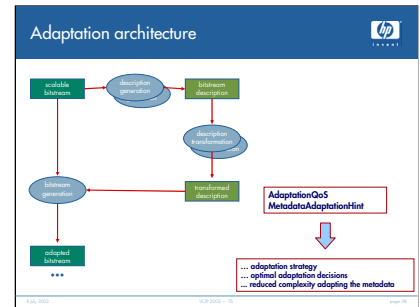
- It is useful to group DIA tools into three groups:
 - Usage environment description tools
 - Tools that assist in the adaptation of the actual resource
 - Tools for the adaptation of declarations
- Here we are interested mainly in the second group, DIRA, which essentially deals with transcoding
 - Need to introduce a structure on the resource so it can be edited: (g)BSD [see next slide]
 - Need to means for deciding on trading off parameters for QoS, i.e., the relationship between
 - > QoS constraints such as bandwidth
 - > Feasible adaptation operations satisfying these constraints
 - > Associated media resource qualities that result from adaptation
 - Need to adapt the metadata, essentially hint information that can be used to reduce the complexity of adapting the metadata contained in a DI

Bitstream Syntax Description 


- A BSD describes the syntax (high level structure) of a binary media resource
- BSDL: XML schema based language to design specific bitstream syntax schemas for particular media formats
- gBS schema: generic schema enabling the construction of resource format independent bitstream syntax descriptions

June 2003 VCIP 2003 - 10 page 10

- The design effort in this standardization effort is to come up with an architecture that allows network nodes to adapt a binary resource without having to know its specific representation format
- This is a key issue for interoperability, scalability, and maintainability
- A resource adaptation transformation is then an editing operation that performs simple modifications (delete, etc.) on a stream that can be performed in the simplest case under the control of an XSLT style sheet
- BSDL (Bitstream Syntax Description Language) schemas can be used by a generic processor to automatically parse a bitstream and generate its description, and vice versa
- The normative gBS schema (generic Bitstream Syntax Schema) enables resource format agnostic adaptation engines to transform bitstreams and their descriptions
- Hierarchical approach, e.g., frames and scenes in a video
- Allows for annotation with semantic handles called markers, e.g., violence rating of scenes
- BSD itself must be adaptable
- Difference between BSDL and gBS schema:
 - BSDL provides means for powerful codec-specific adaptation
 - gBS schema is for generic nodes such as gateways and proxy servers, which do not know anything about codecs and have constrained computational resources

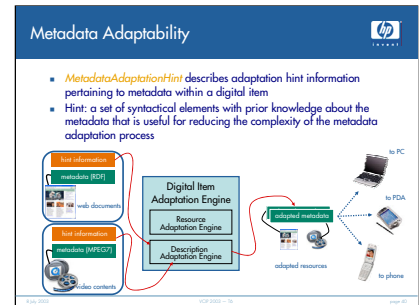


- Description can be gBD or BSD
1. Description generator parses bitstream with a schema and generates its description
 2. Adaptation engine decides optimal adaptation given the constraints
 3. Description is transformed, with references relative to scalable bitstream
 4. Transformed description is used to generate adapted bitstream
 5. Transformed description is updated so that references are relative to adapted bitstream


Terminal and Network Quality of Service 

- *AdaptationQoS* specifies the relationship between constraints and feasible adaptation operations
- Constraints: BandwidthInKbps, ComputationTimeInMillisecs
- Utilities (qualities): PSNRInDB
- Adaptation Methods:
 - frameDroppingAndOrCoefficientDropping, requantization, fineGranularScalability, waveletReduction, spatialSizeReduction
- UtilityFunction:
 - describes possible adaptation operators and associated qualities using a set of constraint points as indexes
 - Linear interpolation is assumed between constraint points
- LookUpTable:
 - additional multi-dimensional sets of data to support more elaborate adaptation scenarios
- StackFunction
 - tool for describing the data in numerical function format

- Terminal and network QoS addresses the problem of media resource adaptation to constraints imposed by terminals and/or networks for QoS management
- The AdaptationQoS descriptor specifies the relationship between constraints, feasible adaptation operations satisfying these constraints, and associated utilities (qualities)
- The AdaptationQoS tool lets an adaptation engine know what adaptation operations are feasible for satisfying the given constraints and the quality resulting from each adaptation
- In this way, terminal and network QoS management is efficiently achieved by adaptation of media resources to constraints
- In general, the AdaptationQoS description is generated in a media resource server and is delivered along with the associated media resource to an adaptation engine located at a network proxy or a terminal
- The generation of the AdaptationQoS description can be done for each media resource stored in a server in advance in the case of on-demand applications
- In the case of streaming of live events, the description could be generated by a prediction-based approach in real-time
- The stack function is one of HP's contributions. HP's big contribution is the AQoS-BSD-Link

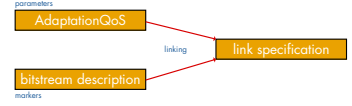


- Hint information is a set of syntactical elements with prior knowledge about the metadata that is useful for reducing the complexity of the metadata adaptation process
- Example: information about the metadata of a video resource
 - InstanceFileSize expresses the actual file size of metadata for the video resource
 - TotalNumOfElements expresses the number of elements that are included in the metadata
 - For the VideoSegment DS used in this example, the following hint information tells us that there are 10 'PointOfView' tags and that these tags exist in the first 3 layers of the metadata

AdaptationQoS - BSD Link 

In some cases it is convenient to specify intrinsic operations based on a universal model for scalable bit-streams

A specified operation can be seen as a **link** that is composed of a mapping condition between identified parameters and an operation, which is conducted if this mapping is fulfilled

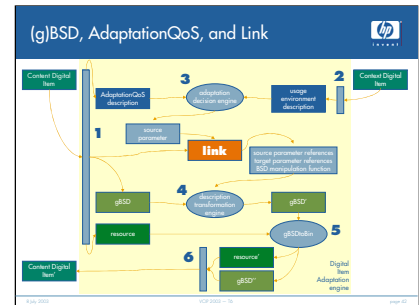


```


graph LR
    subgraph parameters
        A[AdaptationQoS]
    end
    subgraph markers
        B[bitstream description]
    end
    C[link specification]
    A ---|linking| C
    B ---|linking| C
  
```

June 2003 VCIP 2003 - 10 page 11

- Link is necessary when interoperability is desired
- AdaptationQoS is a steering application containing parameters for the adaptation
- The BSD contains markers that identify sections of the resource that can be adapted
- In a closed implementation, the adaptation engine uses the parameters to edit the resource at the markers, whereby the implementations knows how parameters relate to markers
- When interoperability is desired, a third structure called **link** is necessary
- Link specifies how parameters relate to markers, allowing anybody to provide an adaptation service



- **AdaptationQoS - BSDLink** is just called **link** in short.
- This slide is a system view of how link fits in the DIA architecture, based on an illustration by Jörg Heuer
- The adaptation process adapts a CDI (Content Digital Item) based on the Usage Environment Characteristics contained in an XDI (Context Digital Item)
 - gBSD is used to describe the resource's structure
- Walk through the slide as follows:
 1. The CDI is extracted
 - Four items are extracted: the AdaptationQoS description, the link, the gBSD, and the resource itself
 2. The XDI is extracted
 - The XDI usually contains the usage environment, like TerminalCharacteristics, NetworkCharacteristics, etc.
 3. The adaptation decision is made
 - The ADE matches up the environment description against the AQoS description
 - The output is a reference to an I/O pin of the AQoS description, which is the source parameter for the link
 4. The description is transformed
 - The DTE transforms the input gBSD based on
 - the source parameter references
 - the target parameter references
 - the BSD manipulation function specified in the link
 - The output is the transformed gBSD'
 5. The gBSDtoBin process is used to generate the new bitstream, with the address corrections explained in the slide "Adaptation Functions (Update class)"
 6. The adapted CDI is generated
 - The adapted resource and description are combined and the output CDI is generated

MPEG-21 Part 8 — Reference Software 

- Success of a standard depends on the availability of reference software
- Plan to use the software developed in Core Experiments (CE) as a basis
- Platform independence
- Future repository (requires membership)
 - <http://mpeg.nist.gov/cvsweb/MPEG-21/>
- Temporary repository
 - <http://www.tlir.uow.edu.au/cgi-bin/mpeg-refsw.pl>
- Current main issue: parsing DID

Date: 2002-10-23 10:00:00 AM Page: 11

- **MPEG-21 Part 8 – Reference Software**
- The part of MPEG-21 that has most recently been identified as a candidate for specification is Reference Software
 - Reference software will form the first of what is envisaged to be a number of systems-related specifications in MPEG-21
 - Other candidates for specification are likely to include a binary representation of the Digital Item Declaration and an MPEG-21 file format
- The development of the Reference Software will be based on the requirements that have been defined for an architecture for processing Digital Items
- At the October 2002 Shanghai meeting a first list of candidate software from CEs was compiled
- At the December 2002 Awaji Island Meeting the discussion about the reference software started in earnest
 - Until the official repository at NIST will be ready (at this time it is ready but empty), the University of Wollongong is hosting a temporary repository
 - To access the NIST site, you have to be a member of your nation's MPEG delegation
- There are currently two difficulties
 - Not all CE software is platform independent and easily reusable
 - Currently nobody is parsing and validating the DID, because it is hard
 - U of W and the Midgard Media Lab are the most active parties in this effort
 - > <http://www.midgardmedia.net/UMA/mpeg21main.htm>

Digital Item Processing


- Digital Items act as a structure for organizing resources and its descriptions
- Need a mechanism for defining a set of operations by which a terminal can process a DI or DID
- Currently considering to specify a set of operations that can be used to process DIs: Digital Item Method
- A DIM defines an intended method for configuring, manipulating and/or validating a DI

June 2003 VCIP 2003 - 10 page 44

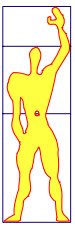
- At the October 2002 Shanghai meeting the action was in the Requirements room
- The issue was: how do we actually implement an MPEG-21 framework
- Specifically, what is a terminal supposed to do when it receives a DI?
- Note the difference between DIM and DIP
 - Digital Item Processing encompasses all aspects of processing a DI from an application perspective

- These slides on DIP reflect the status at the Awaji Island Meeting

Methods vs. processing




- Interoperability of Digital Items means that terminals must handle the DIs in a consistent manner
- Digital Item Methods provide a way to specify a selection of preferred procedures by which the DI should be handled at the DI level
 - a menu of user interaction possibilities
- Digital Item Processing encompasses all aspects of processing a DI from an application perspective
- Applications build DIP environments around a fundamental DIME



June 2003 VCIP 2003 - 10 page 10

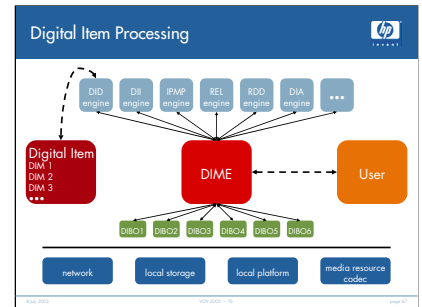
- Emphasis is on the Users: author, publisher, distributor, ...
- DIMs specify what — DIPs specify how
- Example: for a music album DI and “AddTrack” DIM might be provided such that a user can ass a new track in the preferred format of the DI
- DIMs have nothing to do with the processing the media resources; they only deal with the User’s interaction with a DI

Digital Item processing terminology 

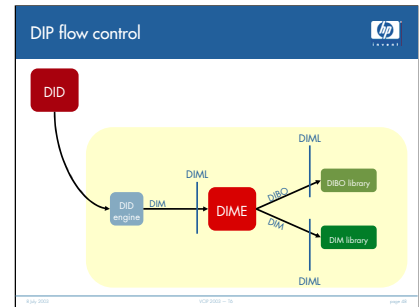
- CDI — Content Digital Item
 - a DID containing the actual content
- DIBO — Digital Item Base Operation
- DIM — Digital Item Method
 - method that can be applied to a DID
 - DIME — DIM Engine
 - part of the terminal responsible for executing the DIM
 - DIML — DIM Language
- DIP — Digital Item Processing
- MI — Method Item
- PI — Processing Item
- XDI — Context Digital Item

June 2003 VCIP 2003-10 page 46


- This is a good point to catch the breath
- We are introducing a number of new acronyms and this is a good point to quickly review them
- The DIBOs describe the basic operations on which DIMs are built
- The operations can be viewed as the vocabulary of the DIMs
- A terminal may already have a library of DIMs available, so in practice the execution of a DIM can be the invocation of a stored DIM, instead of a script calling DIBOs (MPEG-21 parlance: a **list of operations**)
- DIML is the language used to express DIMs
- DIME is the engine on which DIMs are executed
- Object Map: association of the Objects in the CDI with an object type (a type that can be associated with an argument of a DIM)
- Processing Item: item containing the Object Map and the list of DIMs
- MI: item containing the signature and the implementation of the DIMs to which a PI refers



- Dashed arrows: When a DI arrives at a terminal, the two interact. Similarly, the User interacts with his MPEG-21 terminal
- Solid arrows: interactions inside the MPEG-21 terminal
- DIME is the glue that couples the parts of MPEG-21



- Flow control of Digital Item Processing (DIM, DIME and DIML)
- When a DID arrives at a terminal, it is received by the DID engine
- This engine recognizes a DIM and sends it to the DIME
 - The interface between the DID engine and the DIME is DIML, Digital Item Method Language
- The execution of the DIM may rely on DIBOs and/or DIMs that are already available at the terminal
 - The interfaces between the DIME on the one hand and the DIBO and DIM libraries on the other hand are again DIML
- The DIML has not yet been standardized

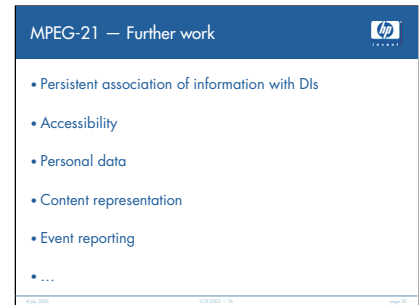
MPEG-21 Part 9 — File Format 

- An MPEG-21 file format shall be capable of storing MPEG-21 Digital Items
 - all components of the DI within a single file
- The MPEG-21 file format will inherit several concepts from MP4, in order to make 'multi-purpose' files possible

June 2003 VCIP 2003 - 10 page 49

- **MPEG-21 Part 9 – File Format**

- An MPEG-21 Digital Item can be a complex collection of information
- Both still and dynamic media (e.g. images and movies) can be included, as well as Digital Item information, meta-data, layout information, and so on
- It can include both textual data (e.g. XML) and binary data (e.g. an MPEG-4 presentation or a still picture)
- For this reason, the MPEG-21 file format will inherit several concepts from MP4, in order to make 'multi-purpose' files possible
- A dual-purpose MP4 and MP21 file, for example, would play just the MPEG-4 data on an MP4 player, and would play the MPEG-21 data on an MP21 player
- Requirements have been established with respect to the file format and work on the WD has been initiated




- These are some Requirements that have been identified and described, but not much work has been performed in the standardization process at the time of this writing
- Persistent Association of Information with Digital Items
 - To categorise all the techniques for managing identification and description with content
 - This will include 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 unauthorised removal and modification
 - PAT: Persistent Association Tool, which is a general technology element that provides the capability of associating information with Digital Items and their Resources
- Accessibility
 - It is assumed that MPEG-21 will provide optimised accessibility to Digital Items for Users with a broad variation in human capabilities that should be supported by all the requirements in this document. Access to Digital Items for such Users may require methods for providing and controlling alternative selectable media presentation forms, structures for storage and transport of logically linked media of different kinds and methods for transformation between media forms.
- Personal Data
 - As a result of the rapid development in computer technology large quantities of information relating to individuals are routinely collected and used by public administrations and in every sector of business, irrespective of national boundaries
 - MPEG-21 should provide tools to enable individuals to protect their personal data and retain their privacy
- Content Representation
 - To provide, adopt or integrate content representation technologies able to efficiently represent MPEG-21 media resources, in a scalable and error resilient way
 - To support content representation of the media resources that are synchronisable, multiplexed and allow interaction
- Event Reporting

Why is MPEG-21 relevant to you ? 



- All rich media commerce on the Internet will be driven by MPEG-21
- Only companies that drive the standard will reap substantial profits
 - First to market
 - Licensing fees

- We have seen that MPEG-21 aims at defining a normative open framework for multimedia delivery and consumption for use by all the players in the delivery and consumption chain.
- This open framework will provide content creators, producers, distributors and service providers with equal opportunities in the MPEG-21 enabled open market.
- The price for abstention from the standardization is that later it will be very expensive to adopt the standard, which eats into profits
 - By implementing at least a prototype during the standardization process and by participating in CEs, which will make your implementation interoperable with other organization's implementations, you will be able to implement a commercial product as soon as the proposed standard is in Final Committee Draft form, which can give you a year's advantage over competitors
 - Because you are first to market, you can charge premium prices, because you compete on features instead of competing on price
 - Because during the CE you have cooperated with the other players in the field, you have been able to create a product that is well differentiated and complementary to the products of the other players; you do not incur the risk of launching a me-too product
- If you contribute to the MPEG patent portfolio, you share in licensing revenue
- If you do not contribute, you have to pay licensing fees
- If you also participate in the industry forum, you can control the value of you patent portfolio

The case against participation 

- MPEG standards take 10 years to become profitable
- My competitor is not participating
- There are strong drivers in the standard and we can just leverage on them
 - Licensing fees are lower than the total cost of participating in the standardization effort
 - We do not have deep pockets
 - OS companies control the desktop anyway
- MPEG-21 is too complex for me

June 2003 200306-10 page 10

- This is a list of arguments I have heard
- The story with Microsoft's Web browser should teach a lesson
 - When they realized that Web browsers will be a key control point for e-commerce, they bought the second best product (Mosaic) and used it to kill all other products
 - Now that it has a monopoly, it has jacked up the prices:
 - > On Mac OS, Internet Explorer can be rented for \$80 a year or can be received as part of an MSN subscription costing \$22 per month
 - > On Windows, the rumor is that MS will switch to a yearly rental model for Windows, which will include Internet Explorer
 - > Because the browser requires a Microsoft Passport, it is not possible to just use an old version, which may not be supported by MS web servers anyway
- This is not just a Microsoft tactic, it is a general monopolist tactic
 - Do not expect Adobe, Oracle, Apple, Sony, Samsung or anybody else to behave any differently if they can build a monopoly because everybody else withdraws from the market
- MPEG-21 is complex, but not hermetic
- It is not a low hanging fruit, it is a comprehensive solution, which requires high investments but will also bring high returns
- One argument I have heard many times is that many of the efforts in which Microsoft invests large sums fail, like the xBox, MSN, and many others
- People argue that they can afford it because they are so rich
- Well, they are so rich because they are willing to take big risks


The case for participation 

- Learn and work with the leading experts
- Understand standards from an evolutionary standpoint
- Work efficiently
- National Body protocol prevents power games
- Stay current on trends
- Be familiar with the competitive landscape
- Discover opportunities for your unique contributions




June 2003 VCIP 2003 - T6 page 11

- It is very hard to work in the vacuum of an ivory tower and just reading documents
- It is easier to learn from people who have already learned
- Standards are very dry documents, you mostly just get syntax
- The semantic is heavily debated in the standardization process, but it cannot be recorded
- The MPEG process is very efficient
 - Well structured meetings allow to work fast and systematically
 - The meetings are very flexible, with BoGs (break out groups) and AhGs (ad hoc groups) forming whenever needed
 - Matters are discussed until a consensus is reached
 - Only the best solution is accepted
 - Collaboration continues between meetings when experts collaborate on CEs (core experiments) and discussions continue on e-mail reflectors
- Anywhere there is more than one person there is politics
 - ISO has strict rules on how the standardization process is carried out
 - Because substantial comments are submitted by National Bodies rather than individuals, power games are minimized, as protocol must be followed
- MPEG meetings and e-mail reflectors are where new technologies are first divulged
- Because everybody plays with open cards, you know quite well what your competitors are up to
- MPEG-21 is very broad and therefore offers many opportunities for contributions, no matter what your many expertises may be

The case for participation (cont.) 

- Score card: focus research on areas of competitive advantage, partner for the rest
- MPEG-21 is an extensive and complex standard
 - Only by participating you can maintain competitive advantages
 - Only by participating you learn with whom to partner



June 2003 VCIP 2003 - 10 page 10

- It is often said we should not reinvent the wheel: if a technology exists, we should just buy it when we need it
- This is often misconstrued as we should not do something because another company is already doing it
- As we have seen at the beginning, MPEG is very extensive and covers many technologies
- Only by participating ...
 - You know if your inventions are competitive or not
 - You learn where you have unique technologies we should cultivate and protect for future business advantage

Work with the experts 

- Under the MPEG aegis, the world's top experts collaborate to create the technology for rich media distribution frameworks
- Leverage on synergies, negotiate differences
- Co-invent new emergent properties
- Own intellectual properties instead of licensing them
- Develop technical partnerships that can evolve in fruitful business partnerships
- Learn what products are on the horizon
- Learn who your competitors will be before product development starts

June 2003 VCIP 2003 - 10 page 11

- This is just a third re-iteration
- Rather than going through the slide, ask if there are questions


How to make money 


- Patent your unique contributions & submit patent statements to ISO
- Participate in the industrial forum to build markets
- Share in the licensing revenue stream
- Have products ready before the standard is published
- Build business alliances at an early stage
- Know the competition



June 2003 © 2003 - 10 page 10


- The licensing process has changed quite dramatically in MPEG's 15 years
- At the beginning, there was no licensing, for example you can use anything in MPEG-1
- As research labs had to generate direct royalty streams to justify their existence and as software patents have been allowed, the companies sponsoring the MPEG experts retain the ownership of the artifacts they contribute by owning patents and licensing them
- The industry fora create markets for the technologies standardized by MPEG
- A number of competing commercial licensing agencies assemble specific patent portfolios and license them to specific industry or service segments; the patent assignees share the revenues; companies that have not contributed are faced with stiff fees that makes it hard to compete
 - Note that the fees are not predatory, once you have participated you know how expensive participation is
 - From a business point of view, a participating expert knows his employer's intended business model and makes sure that a competitor cannot jeopardize this business model
- Because the standardization process is based on CEs (core experiments), you have a working prototype in your hand for which you have already confirmed interoperability
 - Developing a product remains a simple matter of programming, documenting, and marketing
- Especially for MPEG-21, complex alliances will form
 - Having worked for years with the other players, you already have alliances that can be formalized through contracts by your executives
- You will not have to guess what your competitors are up to, or spy on them — you learn to know them very well

What can scientists contribute to MPEG-21 ? 




- Researchers perform long-term research and can pursue activities not yet on a product horizon
- Academia has vast pools of inexpensive brain power

- At the time of this writing it is not possible to predict what will be the big money maker in ten years
- Many required technologies have not yet been fleshed out
- While a product organization cannot engage in speculative product development, a research organization is chartered to explore high-risk opportunities
- MPEG-21 is a huge system
 - Even implementing a small portion, e.g., IPMP requires extensive experimental development
 - Academia can draw on a vast reservoir of top intellectuals that are willing to work for a piece of paper

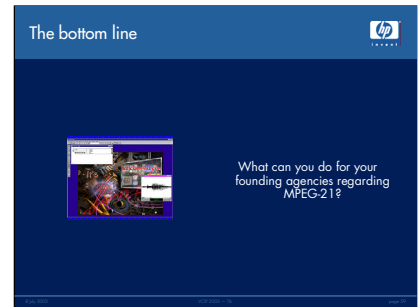
Play to win 

- MPEG groups the world's leaders in rich media
- Leaders win by competing with other leaders



June 2003 VCIP 2003 - 10 page 10

- If you just stay in our sandbox, listening to your business managers without taking initiative, you will atrophy
- Researchers are like athletes that have to constantly compete with the best to keep their lead
- MPEG groups these leaders




- Many national standards bodies, strategic planners, chief technical officers, national science foundations, research consortia, etc. already attend the meetings, either at the board level or in the industrial fora
- Do they need you or can they do it alone?

Time horizon 

- Development time horizon is based on product cycle and vintage chart size
 - Must be able to predict components available for purchase
 - Typically about 4 years
- Research time horizon is based on fundamental breakthroughs that allow paradigm shifts
 - Emergent properties mature over long periods
 - Typically about 10 years
- MPEG-21 is more than 4 years away from general adoption

June 2003 VCIP 2003 - 10 page 40

- There are many technological unknowns that make it very hard to predict how the MPEG-21 framework can actually be implemented
- Recall slide 48 "Digital Item Processing flow control"
 - A request for proposal for the languages to be used for DIBOs, DIME, and DIML is not even out
 - ISO requires platform independence, suggesting C++ and Java for DIBOs & DIME and JavaScript and Perl for DIML
 - However, in practice in many core experiments the implementers use Microsoft Studio and its wizards, which leads to highly non-portable code
 - > For example, when .Net and VisualBasic are used, the implementation will not run on a PC on which .Net has not been installed
 - Similar arguments hold for other elements of MPEG-21, like service discovery, codecs such as AVC, etc.
- All these unknowns make it impossible to make a Gantt chart and do meaningful product scheduling
- Academia, industrial research labs and consortia are better suited for high risk technology bets
 - Not tied to product development cycles
 - Can quickly develop prototype code
 - Researchers fluent in emerging technologies
 - Know how to work with public prototype components and open source

Summary 

- What is MPEG-21 ?
 - An open framework for multimedia delivery & consumption
- Why is MPEG-21 relevant to you ?
 - All rich media commerce will be driven by MPEG-21
- What can you contribute to MPEG-21 ?
 - Your unique expertise
- The bottom line
 - You have the expertise and credibility to successfully represent your employer's interests in MPEG-21

- Leave this slide up without reading it
- Instead make your elevator sales pitch for the specific audience



- Questions and Answers