TV-Anytime metadata

A preliminary specification on schedule!

Jean-Pierre Evain
EBU Technical Department

TV-Anytime will allow consumers to watch programmes in the way they want and when they want. Consumers will be able to access content from a variety of sources, including traditional broadcast and new on-line interactive services – for presentation at any time of the day. New tools (including hard-disk storage, set-top boxes, and high-speed Internet access) will allow us to achieve that goal, i.e. to search, select, access, personalize, capture, and manage a wide and diverse range of attractive content.

This article describes the current work to define a metadata solution, within the metadata technical framework approved by the TV-Anytime Forum. The article is based on the preliminary TV-Anytime specification, which was approved for publication on 28 September 2000 at Marina del Rey, California.

NOTE: the different TV-Anytime documents referenced in this article can be retrieved from the TV-Anytime website: http://www.tv-anytime.org.

Introduction

With more than 110 signatories, TV-Anytime aims to specify an end-to-end system which will allow the consumer to select and acquire the content of interest – for consumption at his/her own time of preference.

Content / service providers will need to transform this perceived threat into new opportunities, by embracing new business models in order to attract and sustain consumers. TV-Anytime will provide opportunities – both for traditional and innovative service providers – to increase revenue by expanding their consumer base, and to offer a wider range of content to consumers who will benefit from an extended framework for new transactional services.

Products demonstrated at IBC in September 2000 showed that TV-Anytime must react rapidly in order to prevent gateways to the home and local digital storage from penetrating the market using proprietary solutions. TV-Anytime will propose complementary open solutions for a horizontal consumer market, by exploiting the development of platforms using open APIs.

TV-Anytime has defined a framework based around a data model and a common metadata representation format. This will allow us to seek metadata tools from other groups (e.g. MPEG-7, the SMPTE, EBU P/META [1] and W3C) in addition to those being defined within the Consortium.

TV-Anytime in a nutshell

TV-Anytime will enable applications to exploit local persistent storage in consumer electronics platforms. The TV-Anytime solution will allow us to browse, select and acquire content independently of the means of content delivery, including enhanced / interactive TV (e.g. ATSC, DVB, DBS and others) and the Internet. The TV-Anytime Forum is developing specifications for open interoperable and integrated secure systems, from content creators/providers, through service providers, to the consumers.

The main functions of a simple system, as described by the TV-Anytime business scenarios, can be represented as shown in *Fig. 1*.

The TV-Anytime system has identified three main sub-systems to be specified:

- **Metadata** (generated by a content or service provider, or by a third party in the delivery chain) to search, select and navigate (within appropriately segmented content), either manually or automatically (e.g. using intelligent agents);
- **Content Referencing** to identify, locate and acquire content (also often referred to as the “fulfilment mechanisms”);
- **Rights Management and Protection** to ensure that the rights of content owners and consumers are protected.

Two additional groups complement the structure of the TV-Anytime Forum by setting out (i) to identify the relevant business scenarios and (ii) to assemble an end-to-end system.

Technical solutions are expected to cover – to the widest possible extent – the requirements identified for each technology. Document TVA037R3 [2][3] contains the metadata requirements. A link is also established in document TVA036R6 [2][3] between the business scenarios, a series of TV-Anytime market introduction profiles and the enabling technology.

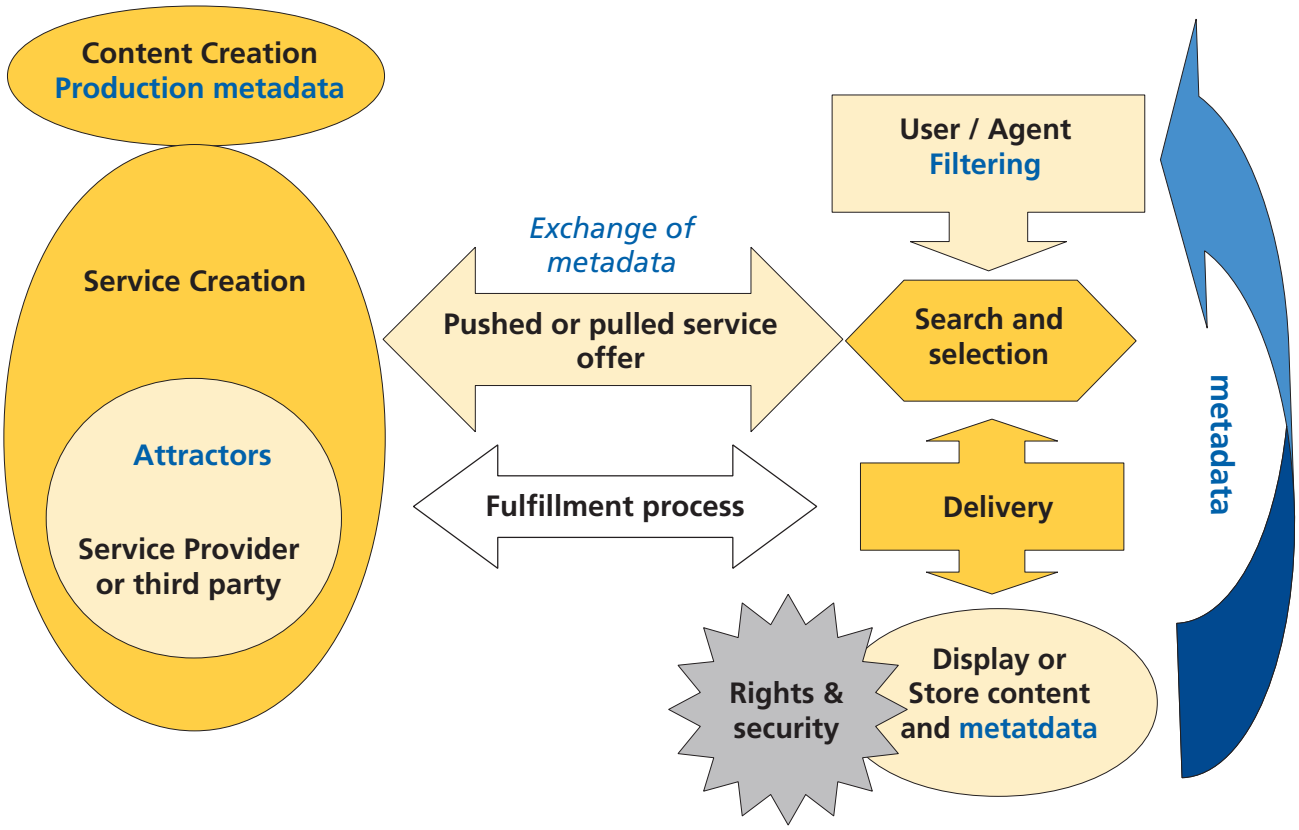

Figure 1
A representation of TV-Anytime content and metadata flows.

Fig. 1 is only one possible representation of the metadata and content flows. However, it is recognized that consumer storage might not be located in the home but in a remote server, which would correspond to a different representation of these flows.

Towards a metadata solution for TV-Anytime

Using the TV-Anytime metadata system, the consumer is presented with descriptive information during the search, selection, navigation and management phases of content consumption. The *attractors* are the information that the user or the user agent will use to decide whether or not to acquire a particular piece of content. The system shall be able to make use of user profiles, preferences, bookmarks; e.g., by using intelligent agents to further assist the user in this task.

Content referencing is a logical process from the output of the selection process to the acquisition of the desired content item. The content referencing mechanisms will ensure that the request is fulfilled through the association of one or more locators (time, physical and logical location) to the actual content identifier. This process will also rely on a common set of basic metadata, associated with the identifiers and locators.

Access to content and metadata will be defined by instance-specific "tacit" or "contractual" usage rules. Content rights management will be essential in protecting content rights. It is expected that these mechanisms will make use of metadata, but to what extent remains to be defined – in particular the control metadata. TV-Anytime will also need to be secure, in order to safeguard metadata such as descriptive, rights management or private information.

The need for an extended metadata framework

Metadata is not new. Giving the generic name of "metadata" to the information attached to content is a first step towards the definition of a framework. This information includes descriptive information to help viewers to identify content in a human-accessible manner.

But the conversion of "notes" – available in a variety of formats (e.g. hand-written, typed, word-processed, voice-recorded) – into "metadata" is not sufficient to justify the development of a metadata framework. The next step in this new digital environment is the use of a common support. Sharing a common format allows the development of a variety of solutions using this information to generate new revenues – business-to-business (B2B) transactions.

The maturity of digital techniques now allows us to think in terms of open platforms in a horizontal market. This will create new opportunities for business-to-consumer (B2C) services. It will require seamless interoperability in a multi-service-provider environment. The solution shall be flexible and extensible, using common representation formats and allowing automatic processing, e.g. using intelligent agents.

A common representation format supporting the interoperable exchange of metadata will benefit from a common data structure based on a common data model, a common definition language (tags, syntax) and a common (still extensible) reference for the semantics and data types.

Such a technical solution needs to be complemented by operational guidelines and procedures – such as metadata (and associated applications) life-cycles, common security measures and common rules for compliance and interoperability testing.

All these elements form an overall framework that does not need to be bound to a specific application domain. This will allow an enlargement of the potential development community that can generate metadata content and applications.

A good candidate for an overall metadata framework: the XML family

For the sake of interoperability, TV-Anytime has adopted a common representation format for the exchange of metadata. Interoperability means that a metadata provider using this representation format will be ensured that this information is appropriately interpreted, processed and rendered on different platform implementations. Interoperability also means that different original representation formats can be used – provided that bi-directional transformation into the common format is possible. This TV-Anytime common description definition language is the XML schema. The important features of XML are described in *Panel 1* on the next page.

XML schema has also been selected and superset by MPEG-7. The SMPTE is considering the use of DTDs to allow the bi-directional transformation of SMPTE KLV-encoded metadata into XML. Recognising the importance of XML for exchange, some individual members of EBU P/META are also working on an XML representation of their metadata. Although this is currently out of scope for the Project, its terms of reference are being revised and will bring it into scope. *Fig. 2* shows the relationship between early adopters of XML as a common representation format for metadata.

XML is also used outside a metadata framework, in particular for presentation engines which manipulate declarative content (ARIB, ATSC DASE, DVB MHP). It will be important to study the degree of compatibility of the parsers developed for presentation engines and metadata.

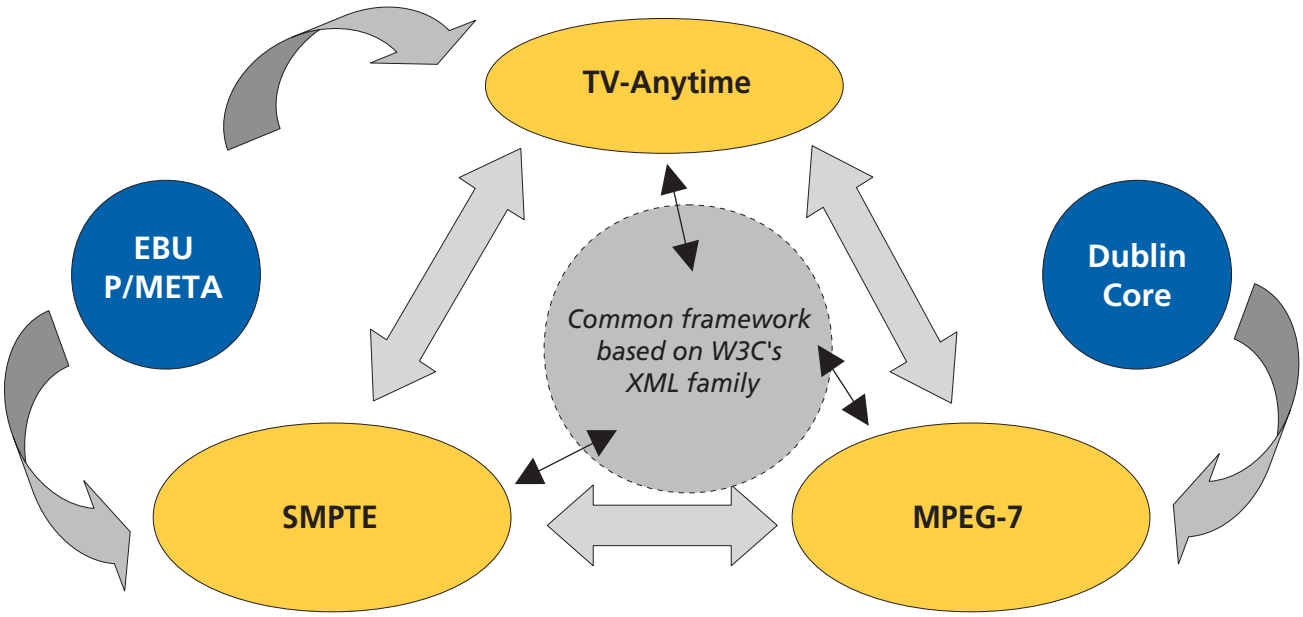

Figure 2
Relationship between the early adopters of XML as a common representation format for metadata.

Dictionaries are used as a common reference for defining the semantics of new XML schema elements (complex types), attributes (simple types) and associated values (e.g. in the form of an enumeration list of controlled terms).

The use of a common representation format allows us to define common libraries as a means of saving bandwidth and maximizing the use and sharing of common resources.

Panel 1
Some important features of XML

XML was designed to describe data in physical and logical structured documents, for interchange on the Web. It is rapidly penetrating the enhanced broadcast and interactive TV domains.

A particularly important feature of XML is the separation of data from the application. It allows us to manage data separately in files or databases. As an example, an HTML document currently contains embedded data. XML can be used to import data into an HTML document. It can also separate the function of presentation, by using an XSL style sheet.

XML schema adds the notion of data types that was not available in the document type definitions (DTDs).

XML schemas are easily extensible, e.g. through the definition of custom semantics and data types.

XML schema also supports the definition of namespaces. This is important in relation to compliance and interoperability. It is also generally admitted that TV-Anytime will make use of a combination of descriptions issued under different namespaces.

XML is also a gateway to a family of tools including:

- The parser. A parser analyzes an XML document. It can check whether this document is well formed according to the XML syntax. It can also validate an XML description by controlling the document structure, semantics and data types. The parser declares an XML document as a valid XML description (XML schema instance) if all these conditions are met. Validation is highly recommended on either the server or client side, or both.
- XSL (Extensible Stylesheet) to represent metadata as declarative content.
- XSLT (XSL Transformation) to transform an XML document into another XML document.
- SQL (Search and Query Language) to search XML content in a database.
- DOM API to access, retrieve and manipulate metadata from an XML document as an interface to presentation and/or procedural applications. With XML DOM, a programmer can create an XML document, navigate its structure and add, modify or delete its elements.

Further information on XML is available from the W3C website: **http://www.w3.org**.

Finally, the use of XML as a common representation format will allow us to benefit from a wide authoring community.

TV-Anytime: a preliminary metadata specification since 28 September 2000

Requirements and business scenarios

A series of requirements (TVA037R3) has been defined to cover the TV-Anytime business scenarios (TV035R6). As a result, the TV-Anytime metadata system shall allow the development of competitive or complementary applications and services which support, for example, the following features:

1) The metadata system shall allow us to search and select content for enhanced broadcast (local interactivity), interactive TV (with a return path) and Internet access.

2) The metadata system shall support a flexible and adaptable parental guidance system.

3) The metadata system shall support the description of consumer preferences and profiles. It shall also support the exchange of such information across platforms. Multiple users shall be supported. Consumer profile metadata can be used by the system to automatically suggest content.

4) The metadata system shall support the management of multi-lingual metadata. It should provide a solution to deal with multi-lingual data and metadata keys.

5) The metadata system shall allow the definition of metadata that enables the ability to navigate and access different views, versions or editions of a particular piece of content. The presentation of summaries shall be supported.

6) The metadata system shall include the appropriate system elements (e.g. indexes) to support non-linear viewing, such as segment jumping (if authorized), and trick modes such as pause and fast-forward viewing.

7) The metadata system shall support the identification and differentiation of multiple instances of the same content.

8) The metadata system shall support the storage of different related content, delivered through different delivery media at different times (clustering, synchronization), e.g. a television series.

9) The metadata system shall support a rich variety of electronic commerce models and transactional services (pay-per-view, view-once, rights to re-publish, free vs. paid television, etc.).

10) It should be possible for the end-user to enter personal metadata such as annotations, in addition to metadata provided with the content, which may or may not be used to navigate.

11) Metadata descriptions shall support links to other programmes. It shall allow the insertion of links to another media, e.g. a website.

12) The metadata system shall support the collection of viewing history information.

13) The metadata system shall be compatible with the local storage management systems for content and metadata. The metadata system shall support content management features (lifetime, life-cycle, expiration, etc.).

14) Updating mechanisms will have to be specified, possibly specifically for each delivery media.

15) Synchronization between content and metadata shall be maintained when required.

16) It shall be possible to protect part or all of the metadata information in order to maintain its integrity, to grant or refuse access to the information (e.g. about rights). Also, to grant or refuse permission to its alteration.

17) The metadata system shall allow the protection of personal data and privacy by providing restricted access according to contractual, tacit or implicit agreement, or specific authorization.

The metadata specification shall eventually cover all the metadata flows represented in *Fig. 3*.

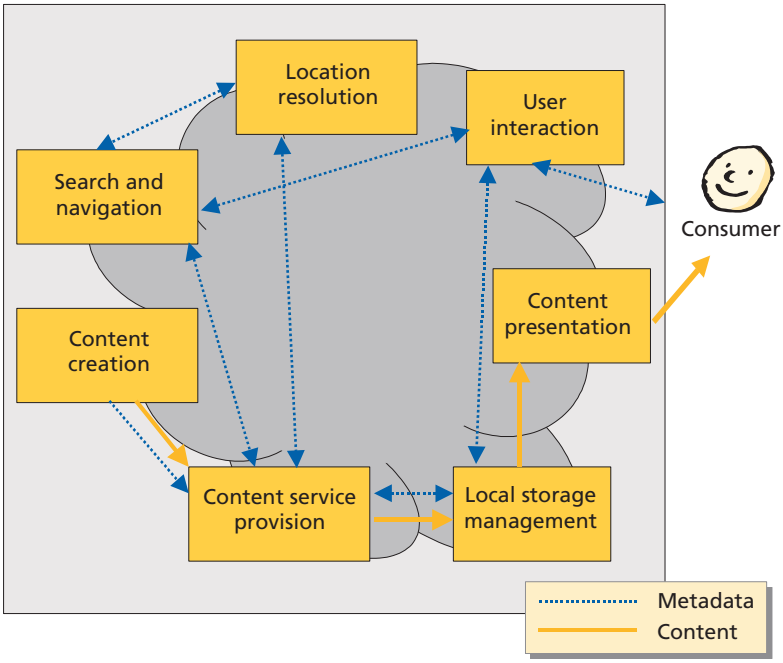

Figure 3
Content and metadata flows in TV-Anytime.

The use of removable media (e.g. smart cards) to exchange user preference and service provider information metadata is not represented but is taken into account in the scenarios.

The technical specification (TVA049)

The preliminary TV-Anytime specification for metadata was adopted on 28 September 2000, at the end of the Marina-del-Rey meeting in California. It addresses both broadcast and Internet delivery. The preliminary specification can be downloaded from the TV-Anytime ftp [3] – document TVA049.

The TV-Anytime metadata solution is a data model described using UML, a modelling language which is independent of any representation format. It is complemented by a representation format (XML and XML-Schema) used to edit the preliminary UML modellized tools of TV-Anytime.

A TV-Anytime system is articulated around the Content Referencing Identifier (CRID). A CRID relates to a specific content / programme instance, or to other CRIDs themselves referring to other different instances. TV-Anytime is considering the use of a global identifier scheme (associated to the CRID) to help in resolving ambiguities such as when pointing at the same piece of content.

The CRID is also the link between a content instance and its associated metadata.

Fig. 4 shows the different clusters of metadata:

⇨ programme description;

⇨ instance description;

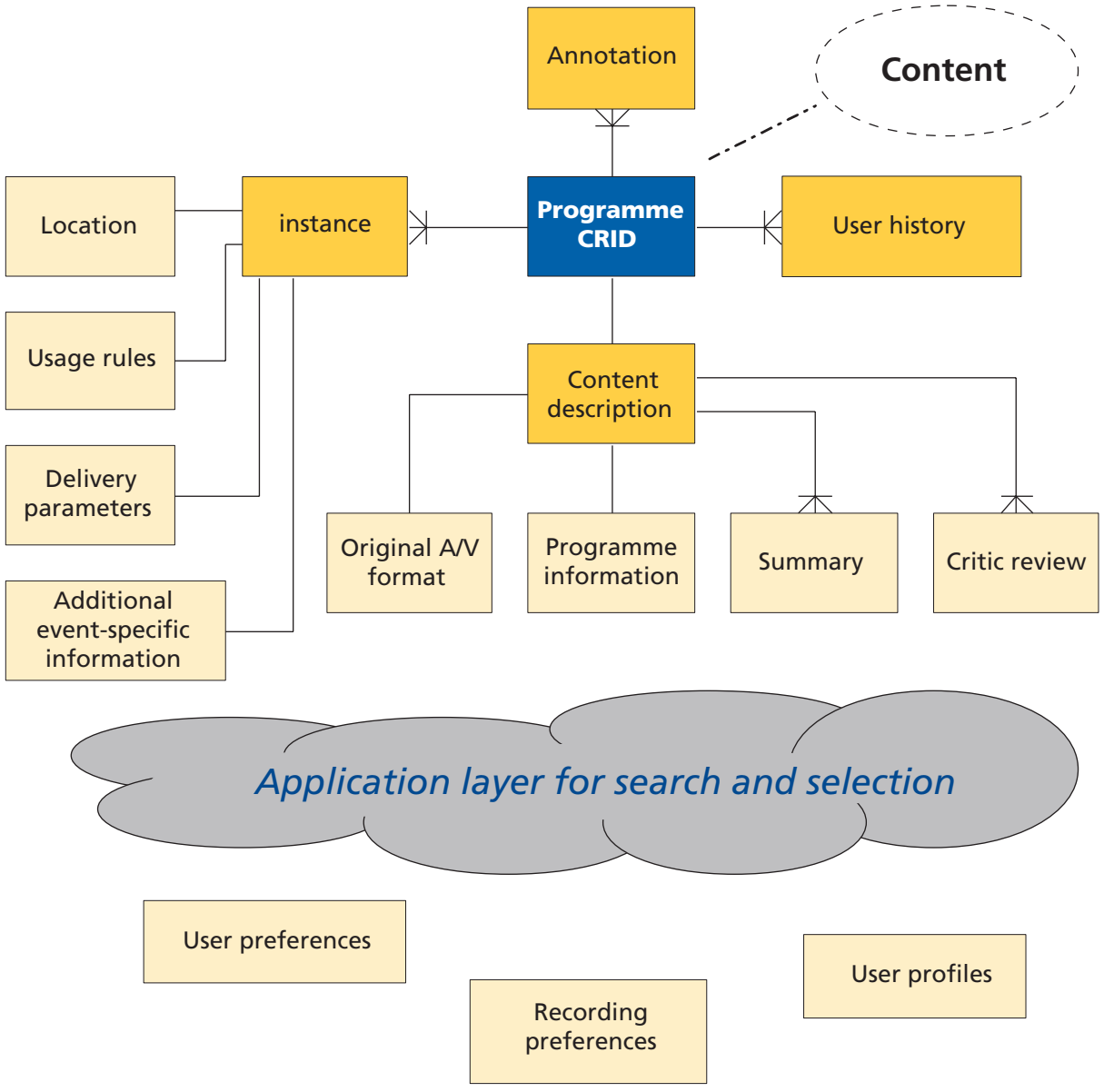

Figure 4
UML representation of the relationship between CRID and metadata.

⇨ other CRID-related metadata;

⇨ consumer metadata.

Several description schema proposals covering the following metadata aspects were received as responses to the Call for Contributions, TVA014R3, published on 17 December 1999. Several of these tools have been adopted or identified for integration into the TV-Anytime specification as follows:

⇨ **Content description metadata:** Group Description DS, Program Description DS, Parental Guidance DS, Summary DS, Media Review DS, Frame Signature DS (to be defined)

⇨ **Instance description metadata:** Internet Instance Description DS, Broadcast Instance Description DS, Service Instance Description DS, Channel Domain Instance Description DS, Channel Instance Description DS, Usage Rules DS (to be defined, related to copy protection, rights management and transactions), Targeting DS (to be defined), Media Format DS (see MPEG-7), Financial Information DS (to be defined).

⇨ **Other CRID-related metadata:** Annotation DS (to be defined), Bookmark DS (to be defined)

⇨ **Consumer metadata:** User Preferences DS, Recording Preferences DS, Environment Characteristics DS (to be defined)

Special attention must be paid to *segmentation* due to the high potential it carries in terms of new applications, services and associated metadata. The first TVA metadata specification has highlighted the issues but most solutions remain to be developed. The review of the business scenarios has led to the identification of two types of content:

⇨ Content which has been edited to allow, for example, segment navigation, insertion, replacement or skipping. Segments are therefore well identified and can be associated metadata. CRID-like mechanisms can be used to locate these segments within a piece of content. However, different pieces of content can also be seen as different segments of a group.

⇨ If content has not been physically segmented when edited, it becomes more difficult to identify segments. But this is still possible using, for example, relative time references even if these techniques are not easy to implement. However, frame-accurate timing is also useful for other specific applications such as near video on demand.

There are many system aspects associated with segmentation. These can be categorized in two main groups depending on whether there is a need for real-time synchronization, or not. In each case, different solutions will be required to attach and deliver metadata information associated with each segment.

An overview of the TV-Anytime metadata system

Fig. 5 presents an overview of the end-to-end system for TV-Anytime metadata creation, delivery and processing from a server to a client.

TV-Anytime is currently focusing on its data model and the definition of a preliminary set of tools:

⇨ As already mentioned, TV-Anytime has opted for an XML framework. Description tools are UML modellized and represented, using the XML-Schema syntax, and structured. XML-Schema allows the extending of a data structure and associated semantics on purpose. If bandwidth and platform resource were infinite, one

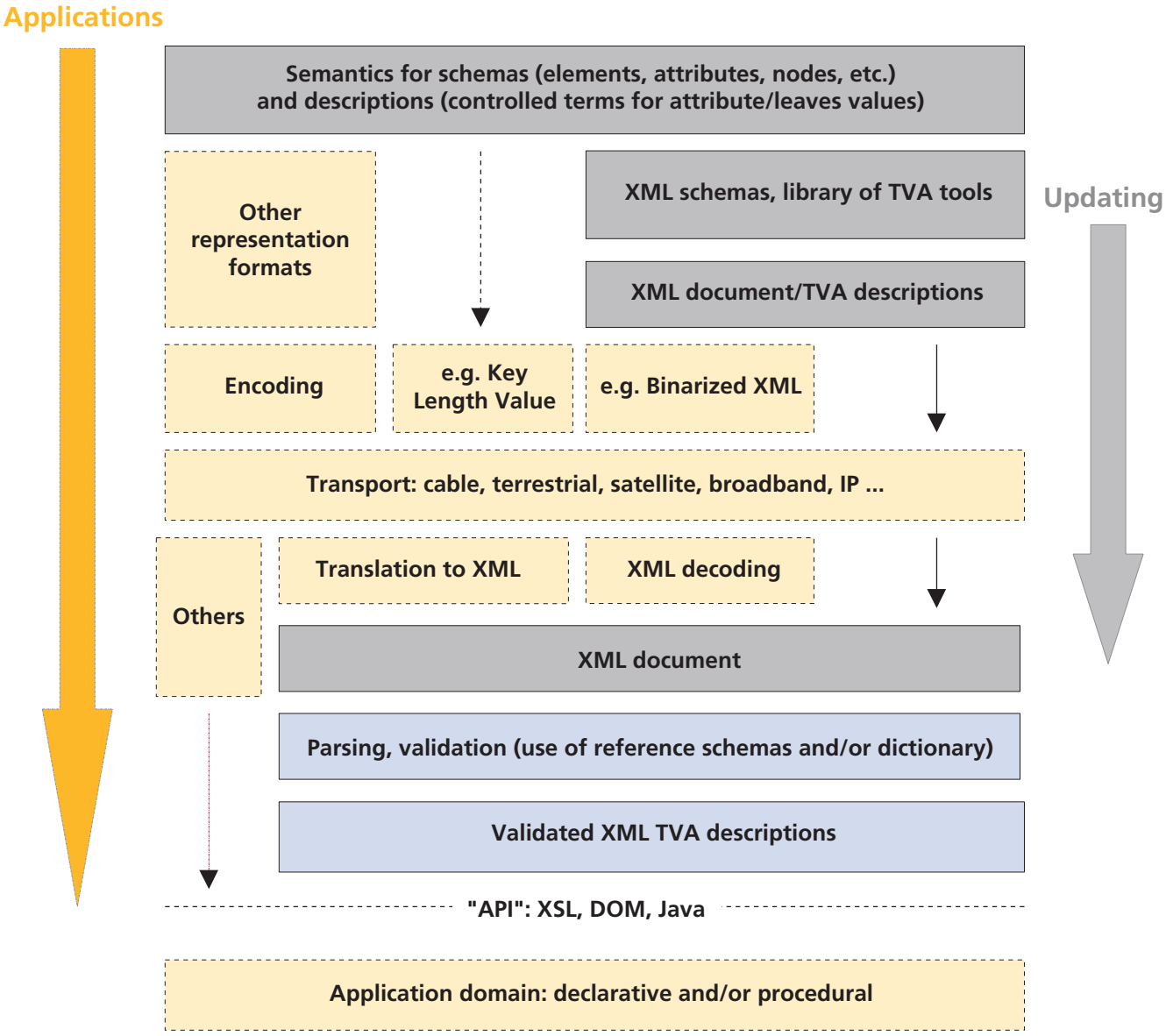

Figure 5
An overview of a TV-Anytime metadata system.

could imagine that each schema and its associated semantics being systematically re-defined and delivered with an application. However, if a common representation format guarantees interoperability, it is useful to have a common library of schemas to maximize the use of the network and client resources. A dictionary can also be used to establish a reference for semantics including controlled terms (enumeration lists) to be used as values when instantiating schemas into descriptions.

⇨ It is not necessary to systematically use XML and XML-Schema. Other formats can be used which can be translated into XML on the client platform in order to benefit from the XML application environment. One example is the use of SMPTE metadata that can be KLV encoded and delivered to the box where KLV-XML translation would occur (work in progress within the SMPTE). The alternative consists of using a different representation format for specific applications.

⇨ An XML description is, above all, an XML document. An XML document is analyzed by a parser. The parser can control whether the description is a well-formed XML document according to the XML syntax. It can also check whether the description is valid according to the XML-Schema structure and semantics. Such a validation can take place on the server or the client side. A basic parser represents 50 Kbytes while a validating parser requires around 300 Kbytes. Both are compatible with current footprints of multimedia home platforms and home computers. However, a resident schema library and dictionary will probably consume more memory space. TV-Anytime will probably need to address these aspects including the definition of a parser normative behaviour.

⇨ Compliance and interoperability will also need to be addressed. It is even more important that the TV-Anytime solution for metadata will make use of tools developed under different “namespaces” including not only TV-Anytime itself but also part or all of MPEG-7, the SMPTE, EBU P/META and possibly others not yet identified.

⇨ Another important aspect is the metadata life-cycle. The expertise gathered in describing application life-cycles when using a common open API environment will be useful. The problem is basically the same and consists of defining rules for sharing common platform resources in a multi-application multi-compliant implementation environment. The life-cycle of a set of metadata will be closely related to the life-cycle of the associated content and applications. Mechanisms to efficiently update schemas and descriptions are already under consideration.

⇨ Security too deserves particular attention. It is foreseen that many data elements will need to be protected against access and modification. Consumer private information is one example of metadata that will need to be safeguarded.

It is currently not the TV-Anytime intention to specify all possible metadata transport mechanisms. It is believed that this should be the responsibility of, for example, ARIB, ATSC and DVB. However, the study of segmentation and updating mechanisms may lead to recommendations towards specific solutions, or adaptations of existing delivery systems.

TV-Anytime is also not directly concerned with the API. However, the choice of the TV-Anytime XML parser will have an impact on other XML environments such as ATSC DASE, ARIB BML, DVB MHP-ML and W3C. TV-Anytime will probably at least have to define the application requirements for declarative metadata (to be directly presented to the user using, for example, XSL) and procedural metadata (to be processed, for example, using intelligent agents – e.g. DOM access and Java procedures).

Still a lot of work on the Agenda

The Marina Del Rey meeting has been decisive. The definition of a data model and an overall TV-Anytime metadata framework represents a key milestone in the definition of an evolutionary TV-Anytime specification.

The TV-Anytime metadata AHG can now concentrate on the business scenarios and, in collaboration with MPEG-7, EBU P/META, the SMPTE and others, look for the definition of the tools currently missing.

Seamless interoperability with the solutions for content referencing will be ensured.

The role of metadata in copy protection and rights management will be further investigated.

It will also be time to engage communication towards other groups (e.g. ARIB, ATSC, DVB, W3C) who might import/adapt the TV-Anytime solutions.

Acknowledgements

As Chairman of the TV-Anytime metadata Ad-hoc Group, the author wishes to thank all the TV-Anytime meta-members who have actively contributed in the definition of this preliminary metadata specification. Special thanks must go to Curtis Eubanks (Sony),

Jean-Pierre Evain graduated from ENSEA (Cergy-Pontoise, France) in 1983. After more than six years in CCETT and Deutsche Telekom, he joined the EBU Technical Department's "New Systems and Services" division in 1992 as a Senior Engineer on the co-ordination of broadcasting Research & Development projects.

Mr Evain has participated in several European projects and he led the CEC DG3's ISIS project UNITEL on Set Top Boxes and Application Programming Interface standardization. This work advanced him to chair the Multimedia Home Platform Launching Group. He now acts as Secretary to the DVB MHP.

Jean-Pierre Evain currently represents the EBU in the TV-Anytime Forum, where he chairs a technical group on Metadata. As a consequence, he also occasionally represents the EBU in MPEG-7.

Dave Marsh (Microsoft), Oliver Gardiner (BBC), Frans de Jong (NOB), Peter Mulder (NOB) and Peter van Beek (Sharp Laboratories, USA).

Bibliography

[1] R. Hopper: EBU Project Group P/META – Metadata Exchange Standards EBU Technical Review No. 284, September 2000.

[2] TV-Anytime web-site: http://www.tv-anytime.org

[3] TV-Anytime ftp site: ftp://tva:tva@ftp.bbc.co.uk/

[4] W3C school web-site: http://www.w3schools.com/

Abbreviations

AHG	Ad hoc group
API	Application programming interface
ARIB	Association of Radio Industries and Businesses (Japan)
ATSC	Advanced Television Systems Committee (USA)
CRID	(TV-Anytime) Content Referencing Identifier
DBS	Direct broadcast(ing) by satellite
DOM	Document object model
DS	(TV-Anytime) Description Schema
DTD	(SMPTE) Document Type Definitions
DVB	Digital Video Broadcasting
HTML	Hypertext markup language
IBC	International Broadcasting Convention
KLV	(SMPTE) Key Length Value
MHP	(DVB) Multimedia Home Platform
MPEG	Moving Picture Experts Group
SMPTE	Society of Motion Picture and Television Engineers (USA)
SQL	Structured query language
TVA	TV-Anytime
UML	Universal modelling language
W3C	World Wide Web Consortium
XML	Extensible markup language
XSL	Extensible style sheet language