EBU Technical Information I36-2003
Metadata Implementation considerations for Broadcasters

EBU Committee	*First Issued*	*Revised*	*Re-issued*
PMC	2003		

Keywords: **Metadata Implementation**

METADATA IMPLEMENTATION CONSIDERATIONS FOR BROADCASTERS

This document aims to provide some background information on Metadata issues in general, and within the context of MXF-based production environments in particular. It also provides some practical guidelines for broadcasters in terms of what they need to do for themselves and to what extent standards can provide solutions today.

Context

Experience to date shows that broadcasting organisations are adopting a range of philosophies and practical tactics in approaching the introduction of electronic Metadata systems. Nevertheless some common features can be identified which help to categorise the scale and style of their activities. These common features can be described in the two dimensions of:

→ which processes does the Metadata need to support (requirements)

→ how the Metadata will be implemented in systems, databases or files (solutions)

Requirements: Which processes does the Metadata need to support

The approach to analysing the requirements will be defined by the scope of the problem to be solved. The scope can be broadly categorised as Content- or Information-driven. These categories are not mutually exclusive and can coexist.

Content-driven: Requirements arising from the management, processing and exchange of Content between devices in the production, post-production, play-out and archive environments;

Information-driven: Requirements arising from the management of information used in business processes including programme commissioning, planning, editorial work, promotion and delivery.

Solutions: How the Metadata will be implemented in systems, databases or files

The solution approach can be categorised with respect to the technical storage and exchange architecture applied to the Metadata and Essence needed to fulfil the requirements. Solutions fall into two main categories which may also coexist:

Joined: Intimately linked Metadata and Essence for example within the same repository or within the same file. This tends to be used in the Content-driven applications domain.

Distributed: Metadata and Essence distributed across separate repositories and separate files for the purpose of storage or exchange and associated via linkage mechanisms. This tends to be used in the Information-driven applications domain.

The requirements will dictate what Metadata is to be used, as they are process-specific. The solution architecture may consist of a mixture of both the joined and distributed categories.

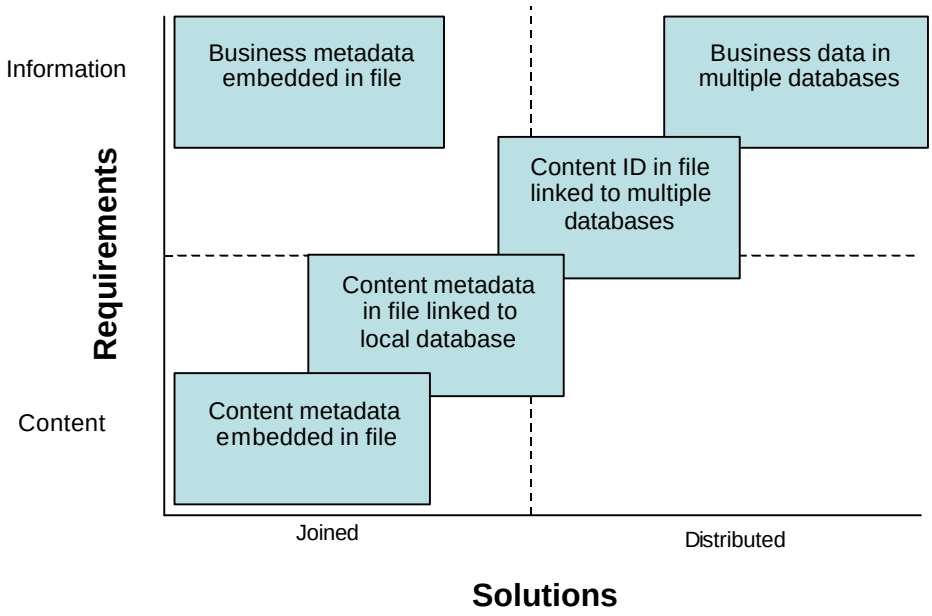

Figure 1: Positioning Metadata Implementations

The scope of the Metadata implementation and the solution architecture (including practical implementation) will largely depend on the rationale for (and sponsorship of) the Metadata initiative. Typically, the organisation will be working to solve one or more of the following problems:

The necessary replacement of ageing stand-alone production systems, (perhaps with Metadata-assisted networked file-based media systems);

To streamline the programme production and archive processes;

To improve information management and integration of electronic media and information systems on a wider basis, either within a broadcast company or between companies.

Activity in this field may be initiated either by engineering or business interests, but should proceed as a partnership between both.

For engineering to drive effectively, existing systems must be becoming obsolete, forcing an upgrade to digital solutions. When networked digital media systems and file-based operations are considered, the need for Metadata to identify and manage the Essence becomes unavoidable.

The primary business drivers for the move away from stand-alone systems, human connectivity and “sneaker net”-based exchange are financial, either to reduce cost or increase output and, in commercial environments, revenue. This may be achieved by streamlining the Content production and distribution processes, or by making existing Content more easily accessible, shareable and re-usable. Solutions for these issues are going to fulfil Content-driven requirements, focussing on the management of Content and its lifecycle.

When looking at the overall business picture, some organisations may view the Content lifecycle as only one strand of their overall business, and may conclude that better management of information is required across a much broader front. The Metadata required to support the Content lifecycle is often originated and used elsewhere, and may be related to business processes which do not involve using Essence at all. This scope may extend to the enterprise level and gives rise to the Information-driven requirements as defined above.

Therefore the first task for any organisation is to clarify its motivation and objectives, and to decide on the scope of its Metadata initiative.

Within the scope of “Content-driven” requirements the initiative may be applied and justified at a relatively local level, perhaps in one project. In the context of the “Information-driven” requirements the initiative will require sustained executive backing and investment from the corporate centre. In either case, significant attention must be paid to ensuring that the organisation fully understands its requirements, and can effectively apply possible solutions.

1. What users must carry out for themselves?

Broadcasters need to be aware that a decision regarding the use of electronic Metadata systems and file-based operations is a fundamental strategic decision that requires substantial preparatory work by the users themselves. This is a task where standards or specifications can only deliver support. Currently, there is also considerable debate regarding the effective application of the Metadata facilities provided by file-based Content formats such as MXF, which requires detailed user consideration.

The following paragraphs highlight the most important issues for the preparatory work to introduce electronic Metadata systems in general (step 1 to 3, 5 and 6) and for the application of Metadata in file-based implementations (step 4).

1.1. Step 1 - Understand the Rationale for the Introduction of electronic Metadata

The organisation must decide the scope of the initial activity and the direction towards the final integration target. This means that the organisation needs to develop a long-term Metadata strategy to ensure that initial projects are interoperable and future-proof. A robust approach to Content lifecycle management is a common objective for many broadcasters.

There should be a clear financial or creative benefit from the introduction of electronic Metadata, with senior management support for investment and managing the change. The most important thing to acknowledge is that new technology will not deliver benefit unless it enables process improvement, so that is where the analysis should start.

1.2 Step 2 - Analyse the Processes to be changed

Analyse the current processes and workflows,

Identify current and future requirements and define new functionality to be supported

- → Design the improved version of the process and, if possible, take into account future business needs
- → Identify the financial impact – costs and savings
- → Define a step-by-step plan for making the change

This work will not only help with workflow and system design, but will provide the basis for defining the Metadata requirements, and input to the investment case. Widely accepted techniques can be used like the Unified Modelling Language (UML) for defining Use Cases.

Be aware that the proposition may be simple or highly complex, but the key requirement is for clarity of both scope and expression (e.g. documentation).

1.3 Step 3 - Define the Information needed in the Process

Identify current data used and exchanged – in forms, information systems etc.

Define additional future information requirements, considering a potentially broader scope

Create a formalised common view using agreed terms and definitions – a “Metadata scheme” (for example also including data models)

Review this against Metadata standards (published by an accredited standards body) and decide how far any of them are applicable with respect to the requirements. Examples for such standards are SMPTE Metadata Dictionary (RP205), Proposed SMPTE Standard DMS-1 (380M), EBU Tech. 3295 (P_META v1.0), EBU Tech. 3293 (Metadata for Radio Archives).

Create the overall Data Specification for the project, however simple or complex, by using a published standard if appropriate, or by defining an organisation-specific data model including semantic definitions, or through a combination of the two.

All project data specifications, including any data model, should be produced using a recognised methodology (for example Entity Relationship Diagrams). This work requires expert understanding of both information analysis and broadcasting.

1.4. Step 4 - Define the System Context

- → Identify the current systems to be replaced or integrated with any new systems
- → Identify where the data specified in step 3 is currently created, read (and used), updated (and deleted).
- → Define the scope of the new system, in functional terms.

→ Define where file-based Essence or Content will be used, what Metadata might be required, and which part of it will be embedded or externally linked

→ Adopt methods for the encoding of Metadata and Essence based on the data specification, support of standards and potential products. Examples for different encoding technologies are XML or KLV.

→ Agree the need for Information integrity to be maintained across the system and its wider context, in terms of persistence of the embedded Metadata. For example: is it valid for long-term storage or is it transient and only valid for the immediate process? Should some other source be regarded as the "master record" e.g. the archive catalogue?

→ Define the import and export interfaces for Essence and Metadata into and out of the system, in particular with respect to the file and its embedded and externally linked Metadata, e.g. from existing systems, including the technical means of exchange.

1.5 Step 5 - Assess Solutions

- Use the specifications developed in steps 2 to 4 to create an evaluation model, assess supplier solutions and integration approaches.

- Having selected a supplier or system integrator, identify the gaps between the proposed solution and the desired specification, and either request changes or identify workarounds.

1.6 Step 6 - Manage Implementation

If implementing a third party solution, record any differences between the intended data specification and the delivered system (this may only apply at the interface level).

If developing the system in-house, ensure that the technical database design keeps in step with the data model and its semantic definition by monitoring and recording any changes e.g. for storage or processing optimisation.

Keep a record of both the semantic Metadata model and its implementations (storage data models and encoding, exchange data models and encoding as a basis for any future up-grades or system integration.

Note: The semantic model will be needed for modelling future business requirements; the data models and their encoding will be needed as a basis for future technical changes and maintenance.

Use centralised change management to record changes in the data specification and its implementation.

Monitor new industry standards over time to see if and how they impact the implementations.

If the approach to Metadata implementation is intended to cover a number of projects over time, whether meeting Content-driven or information-driven requirements, it will be necessary to establish and maintain an in-house Metadata design authority.

2. Metadata in the context of MXF - Implementation Considerations

Where the organisation aims to use Metadata embedded in files, a number of further considerations must be taken into account. The following applies specifically to the Material Exchange Format (MXF). For a glossary of terms used in this part, please see the Appendix to this document.

2.1 Requirements, Considerations

The in-house analysis outlined above should result in formal descriptions of processes and application environments. This description will include the data and media flows through the processes and application environment, the definition of a data structure for Metadata exchange and the information elements for specific use cases (this may be an iterative process). In addition, this analysis will help to identify the areas where Metadata is linked to or carried within MXF.

Users should as far as possible consider whether the available published Metadata dictionaries and schemes, such as the proposed SMPTE standard DMS-1 (already with KLV encoding for MXF), EBU Tech.3295 (P_META), and EBU Tech.3293 (Core Metadata Set for Radio Archives) can meet the requirements of their needs. This will help to achieve interoperability.

2.2 Solutions Considerations

MXF uses KLV coding as its general encoding protocol. Descriptive Metadata is supported via a generic plug-in mechanism. In order to use this plug-in mechanism KLV encoding of the descriptive Metadata is required. This KLV encoding can be carried out at a scheme level or at the element level. There can be multiple instances of the same or different schemes in the same MXF file.

- In scheme level encoding, one instance of the scheme is encoded as a single KLV packet using a SMPTE registered Key. The Key determines how the application handles the data that follows the Label for the given Length. Users should be aware that MXF decoders will not be able to decode the Metadata encoded at the scheme level.

Example: An XML instance of a scheme encoded in one KLV packet. The XML instance of the scheme transported via MXF becomes understandable only after exporting from MXF and parsing by an additional application;

Alternatively, it would be possible to encode and register any scheme as KLV at the individual element level. In cases where the scheme depends on structuring of the elements to convey meaning, the strong and weak references mechanism of MXF should be used to encode these structures in KLV.

Example: a draft proposal for mapping EBU Tech. 3295 (P_META) including its set structures and repeating groups as KLV is under development by the EBU project P/TV-File. The KLV encoding for DMS-1 is defined in this way.

If backward compatible extensions are required for a SMPTE DM scheme (following the rules of SMPTE 359M for dynamic documents) the extensions should be registered through the SMPTE in order to maintain interoperability. For example: backward compatible additions to the DMS-1 will lead to a new version of the DMS-1; modifications which are not backwards compatible may lead to a new Descriptive Metadata Scheme (DMS-n);

The linkage mechanism between MXF files and external Metadata should use the Unique Material Identifier (UMID) specified in SMPTE 330M-2002;

Note: The linkage mechanism between files requires further investigation;

If deciding on a proprietary Metadata solution (with no registered Key definition, known in MXF as Dark Metadata), users should be aware that all the file-specific applications, except those parsing Metadata from the MXF file, must be customised.

There will be environments where MXF files are interchanged with AAF applications and where the AAF application is required to be able to process the Descriptive Metadata located in the MXF file. In order to be compatible with the constraints of the AAF object model, a MXF descriptive Metadata scheme needs to use a single inheritance object model and to ensure the uniqueness of each property.

Appendix: Glossary of terms

SEMANTIC	The meaning or relationship of meanings of a term or set of terms.
SCHEME	As in Semantic Metadata Scheme. A formally structured logical method and notation for defining information in terms of the subjects described, their characteristics and the relationships between them.
ELEMENT	As in a descriptive Metadata element. A characteristic of a subject being described, which cannot be divided further into sub-characteristics. Also referred to as an attribute, property or item.
ENCODING	The technical means of expressing semantic information in application readable form.
KLV	Key Length Value binary coding syntax as standardised in SMPTE 330M.
KEY	That part of the KLV syntax which identifies the attribute or element or item to which the value applies (SMPTE Universal Label as used, for example, in the SMPTE dictionary/registries).
METADATA PLUG-IN MECHANISM	Term used to identify the interface provided by the MXF structural Metadata which allows the insertion of descriptive Metadata into an MXF file preserving its integrity and expressing its relationship (time, track) to the essence. Further information can be found in the proposed SMPTE standards MXF Format and the related proposed engineering guidelines.
INSTANCE	An occurrence of information values.
PARSING	The electronic process of reading data sequences and breaking them down into recognised elements for further analysis.