

EBU Project Group

P/Meta

— Metadata Exchange Scheme, V1.0

Richard Hopper

Media Data Group, BBC Technology

Since we last reported here on the work of EBU Project Group P/Meta, the Metadata Exchange Scheme has been published in version 0.9 (to encourage user feedback).

This follow-up article describes what is expected to appear in version 1.0 of the Scheme, to be published by Summer 2002 – what it will represent and how it will relate to other metadata schemes, particularly those of the SMPTE and TV-Anytime. The article identifies what should happen next to support implementation and use of the Scheme, and describes some example transactions used in the P/Meta demonstration at IBC 2001.

Context

The case for metadata is now well and truly made. The industry knows its value, led by the EBU/SMPTE Task Force which reported in September 1998 – notably, in this context, the section on “wrappers and metadata”. Since then, the SMPTE in particular has driven standards for the manufacturing industry, and TV-Anytime – for media delivery to the consumer. There is now no doubt that the industry recognizes the significance of metadata. What has been less clear within the industry has been how it goes about capturing and expressing its requirements for specific information transfers between parties. Although we have an abundance of dictionaries and schemes, specific requirements are proving more elusive to capture. To address this gap, the P/Meta project was set up. The P/Meta scheme has been derived from actual business transactions; it is designed to satisfy projected future needs, and it addresses the specific requirements for business-to-business (B2B) metadata exchange. Further, it is closely referenced to the SMPTE metadata standardization work.

In the P/Meta Version 1.0 Scheme, we are confident that we will have a rich resource from which to satisfy all the B2B metadata exchange needs of EBU members. But we don't want to keep it to ourselves! It will be made available to any interested party, simply by registering [1]. We are confident that it will meet most needs of the professional media community, and we expect it to be reasonably stable. However, its boundaries will not be cast in stone – it is extensible, and undoubtedly will be extended. In order to accommodate the inevitable, we already have a change control mechanism in place to support the inclusion of new metadata attributes and sets, and to communicate them.

The Scheme has been scoped to include both “traditional” and new media, and is code-based where appropriate to support language independence.

Further, in developing the Scheme, we have used methods standard to the information industry – business process analysis, data analysis, semantic and syntactic definition, validation, review and change management. This approach gives added confidence to us, as users.

New work is leading towards support for the “metadata gateway” – from business to consumer (B2C) – through collaboration with TV-Anytime.

Project progress

This article builds on the outline of P/Meta published in EBU Technical Review No. 284 (September 2000) [2]. Since then the project has:

- 1) published the V0.9 scheme [1];
- 2) built a demonstrator and shown examples of P/Meta exchanges at IBC 2001 (*see Appendix A*);
- 3) created two XML Schema from different approaches, and is currently evaluating them in order to determine a harmonized way forward;
- 4) contributed to the MXF metadata “Geneva Scheme” and confirmed the requirement for the plug-in interface for other appropriate metadata collections;
- 5) mapped P/Meta to various other schemes, particularly those of TV-Anytime and the SMPTE;
- 6) tested the ability of the scheme to support a range of business-to-business transactions in order to validate it;
- 7) made a proposal to the EBU International Statistics Network to develop extensions for their ESCORT 2.4 genre scheme [3] in close collaboration with the TV-Anytime Forum in order to meet the requirements for programme classification in both the professional and consumer domains.

The project is now focussing its work on final validation, publication and communication, and developing ongoing support and maintenance arrangements for the Scheme.

Scope of the P/Meta Scheme

To recap, the EBU P/Meta Metadata Scheme V1.0 is a scheme for metadata exchange, comprising:

- a flat list of attributes complete with semantic definitions;
- a list of transaction sets, each of which is built from attributes and other sets; each set has its own definition of purpose and content;
- a list of reference data (also known as “enumerated values”, “code values” or “controlled value sets”) for appropriate attributes;
- a syntax and notation for set construction which supports members’ requirements for the assembly of a logical set.

The Scheme deals with the logical content and the meaning of information, as distinct from the technicalities of coding and protocols.

This article describes only the logical scheme – other workpackages in the project are addressing the requirements for unique identifiers (for material and editorial content) and technical exchange.

Scheme objectives

The key objectives of the Scheme are:

- to support identification and recognition of material;

Abbreviations

AAF	Advanced Authoring Format	KLV	(SMPTE) Key Length Value
B2B	Business-to-Business	MDD	(SMPTE) Managed Data Dictionary
B2C	Business-to-Consumer	MXF	(Pro-MPEG) Material Exchange Format
IBC	International Broadcasting Convention	SMPTE	Society of Motion Picture and Television Engineers (USA)
ISO	International Organization for Standardization	TVA	TV-Anytime

- to provide editorial and descriptive information;
- to establish the right to use material;
- to inform the user how to properly open and/or play back the material.

It should be possible to satisfy these objectives:

- independently of the system platform of the implementing organization;
- independently of the user's human language – at least for the exchange of all coded values.

Who's exchanging with whom? – the dataflow

The inter-business dataflow model initially developed by the P/Meta project had three trading entities: content creator, content distributor and content archive. The “consumer” entity was acknowledged from the outset, but declared to be out of scope (*see Fig. 1*).

However, an extension to support the interface to the consumer domain is now being developed in collaboration with TV-Anytime (which delightfully terms the professional domain as the TV-Anytime “back office”!).

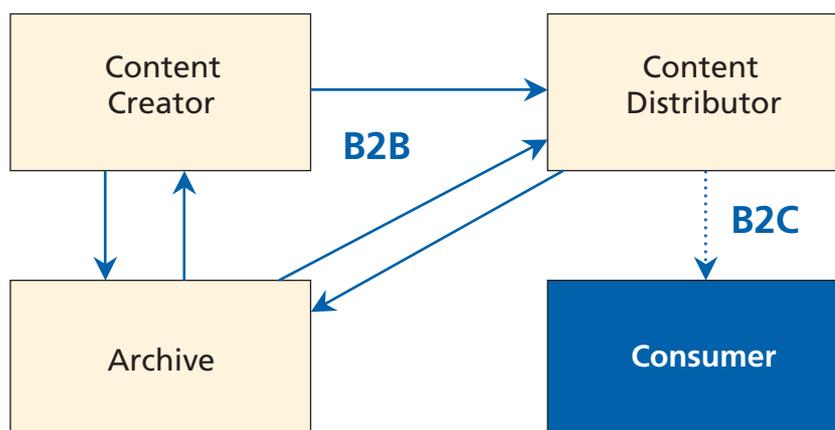


Figure 1
P/Meta business-to-business (B2B) dataflow model

P/Meta Scheme – what it is and what it is not

P/Meta is ...

- a universal standard for metadata exchanges between professional media organizations;
- a definition of common meaning to the data fields and values that most broadcasters use in order to enable exchanges;
- designed to be flexible and suitable for use in a wide range of broadcasting activities;
- both language and system independent;
- a joint development by EBU members on a not-for-profit basis;
- a scheme that makes use of other standards where possible, e.g. ISO country codes.

P/Meta supports ...

- defined metadata for identification, description, discovery and use of content in B2B transactions;
- the exchange of metadata – which will typically be separate from, but may be embedded in, a programme stream or file;
- implementation on any appropriate platform, using XML or KLV (or any other appropriate coding) “on the wire”. It is expected that P/Meta will be implemented using both XML and KLV.

P/Meta is being extended to include the data required by the consumer domain from the professional domain – as specified by TV-Anytime.

P/Meta is not ...

- concerned with the transfer of the media itself;
- a definition for the storage of metadata;
- a definition of the exchange file format or mechanism.

P/Meta does not support ...

- storage of information – it is *not* a database scheme.

This is an important distinction. P/Meta attributes can be assembled into sets in any arrangement – and the sets themselves arranged in any logical way that makes sense to the business. P/Meta attributes and sets are not named, defined or determined in their mutual relationships in a way that can ever satisfy requirements for a storage scheme.

What is the relationship between P/Meta and other information schemes?

P/Meta has developed beneficial relationships with a number of metadata schemes:

- 1) The **SMPTE Metadata Dictionary** [4] and **Sets Registry**, and the closely associated **Material eXchange Format** (MXF) [5] with its Geneva Scheme for descriptive metadata;
- 2) **TV-Anytime** [6] – the mapping to its metadata requirements, and work to develop the Programme Classification Scheme based on **ESCORT 2.4** [3];
- 3) The **EBU Project, P/FRA** (Future of Radio Archives) [7]. A unilateral mapping has been provided from the qualified **Dublin Core** [8] radio archive set to the P/Meta Scheme.

These three relationships are now discussed.

P/Meta, SMPTE and MXF

This section aims to clarify the relevant SMPTE, P/Meta and MXF scopes and offer a view on where the respective approaches converge, and how they should be viewed together. The relevant roles and objectives for the three activity groups are outlined here:

SMPTE – Metadata Dictionary & Sets Registry

- The SMPTE is a standards organization that has created the Metadata Dictionary (MDD) [4]. The scope of the MDD dynamic document encompasses all the data elements considered relevant by the industry (as determined by due process).
- The SMPTE has started to create a Sets Registry, defining recognized groups of data elements for storage or exchange. The Sets Registry will describe the business purpose and the structure of the sets that can be used to support any transaction sets required for the exchange of metadata.
- The SMPTE uses Key-Length-Value (KLV) as the structure for its data elements and sets.
- The SMPTE intends to create an XML representation of all its data elements and sets.

P/Meta

- The P/Meta Scheme defines the attributes and transaction sets that have been identified to support the exchange of media assets between EBU members as content creators, distributors and archives. The

Meta project has defined and published this specification which comprises a list of attributes, particular coded values and meanings, and particular sets based on business knowledge and specific analysis.

- The Attributes and Sets within P/Meta are expected to be further developed and validated to support the required range of business transactions. Reference data sets are also being defined where necessary.
- Once validated, Attributes and Sets not provided by the SMPTE will be registered for inclusion within the SMPTE MDD and Sets Registry. For clarity, mappings will be shown between P/Meta attributes and SMPTE elements and sets already included by the SMPTE.
- The P/Meta Attributes and Sets will also be represented in XML. This work is current. If P/Meta is required to be implemented using KLV representation, codes will be provided by the SMPTE MDD and Sets Registry when the P/Meta Scheme has been registered.



MXF – Descriptive Metadata

- The MXF format will allow users to take advantage of non-real-time transfers, and to package together essence and metadata for effective interchange between servers and between businesses. MXF files use KLV coding throughout, for flexibility and extensibility as defined by the SMPTE.
- The MXF team has gathered its metadata requirements from other projects working on metadata transfers (including MyTV, G-FORS, P/Meta, etc) and from the SMPTE Metadata Dictionary. This has been defined as the default scheme (Geneva Scheme) within the MXF format specification suite.
- If another metadata scheme (other than the default Geneva Scheme) is to be used with MXF, this must be fully defined, and must be linked to the structural metadata within the MXF file, i.e. time tracks, etc. The MXF file format permits any suitable additional scheme to be used in place of the default Geneva Scheme – MXF can support multiple such schemes. (The Geneva Scheme has been submitted as a Dynamic Standard within SMPTE provisions, such that the Scheme can be expanded with new elements, sets and even collections of sets, as needed in the future.)
- The MXF specification is a native KLV implementation. However, its descriptive metadata can be translated and represented in XML using an XML Schema and dictionary to be provided in support of exchange with computer systems.

P/Meta, MXF and SMPTE relationship

In summary, the relationship between P/Meta and MXF is expected to be only via SMPTE specifications and dynamic documents. The P/Meta scheme will be fully incorporated within these SMPTE documents. The MXF file format is conformant with SMPTE specifications; anything not within the SMPTE registries will be registered before MXF is published as a standard.

The difference in scope between MXF and P/Meta can be summarized as follows. The MXF format is intended to allow the interchange of finished or “almost finished” material. It is not intended to be an authoring format, capable of supporting the full requirements of storage – that is the role of AAF [9]. However, P/Meta has a broader scope of metadata exchange between Content Producer, Distributor and Archive. It supports *all* business transactions concerning material – even those required before any material exists!

P/Meta and TV-Anytime

TV-Anytime (TVA) has developed an enhanced dataflow model for the “content distributor” trading entity – shown in *Fig. 2*. This shows the additional discrete functions now needed to support new media, as well as traditional delivery, and the position of the B2C interface.

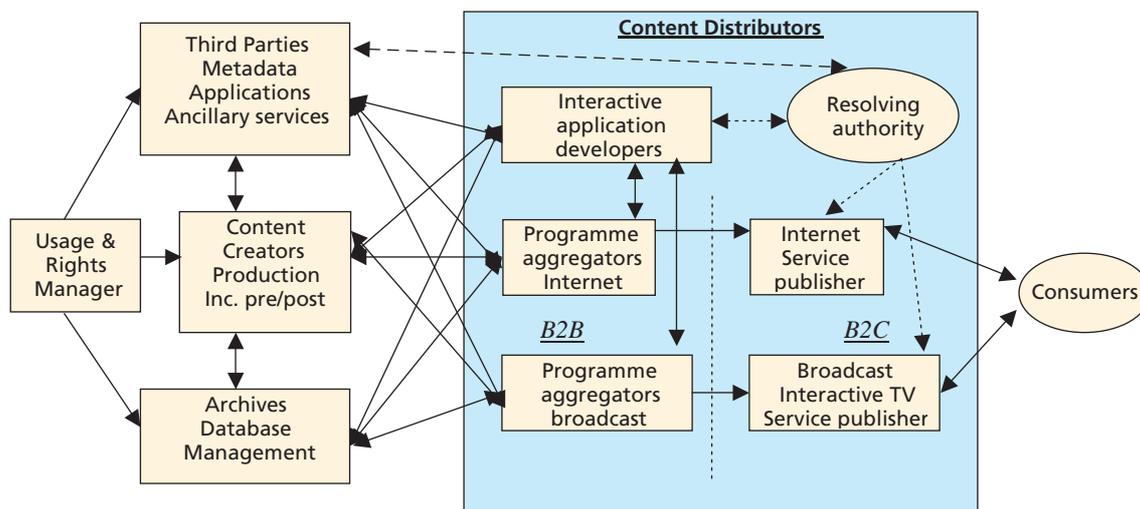


Figure 2
P/Meta dataflow model, extended to show the TVA Content Distributor entity

There are two areas of collaboration between P/Meta and TV-Anytime: the mapping between the schemes at the B2C interface, and development of a “programme classification scheme”, often known as the “genre” scheme.

○ Mapping

The P/Meta view of the TV-Anytime future is that its metadata requirements create end-user demand to “pull/push” data items through the interface between the professional and consumer domains. P/Meta needs to know how many of these data items are also used to support the consumer, and how many of those needed by the consumer are currently missing from the business-to-business scheme – and thus potentially need to be added as part of content creation and delivery.

A draft mapping from P/Meta has been created for a selected TVA metadata set. This work indicates that about 85% of TVA attributes can be supported by just 25% of the 200 or so P/Meta attributes. This gives confidence that there is a specific purpose, internal to the professional domain, being served by P/Meta; further, that its professional metadata offers substantial added value to the B2C metadata gateway.

Further work is expected in continued collaboration with the TV-Anytime Forum to complete the definition of the mappings between the professional and consumer domains.

○ Genre Scheme

Mention has been made of the work done in identifying reference data (standard and coded values) for many of its Attributes. A clear area of commonality is in the area of programme classification by attributes such as “genre”. For consumers to find the programmes they want to watch, and for distributors and gatekeepers to monitor audience behaviour, it would be helpful to have a standardized set of descriptors in use across all B2B and B2C transaction interfaces.

Work on the programme classification reference data scheme, developed from the EBU’s ESCORT 2.4 scheme [3], is now being progressed through the TV-Anytime Forum [6].

P/FRA – radio archive set

At the request of the EBU P/FRA Project [7], the archive metadata set for audio material (which was based on qualified Dublin Core) was mapped to the P/Meta Scheme. This mapping was limited by its one-way relationship. Further work is in hand to develop this un-directional mapping to support a full bi-directional semantic exchange. When agreed, this work will be published by the EBU within the P/Meta V1.0 Scheme. The main work of P/FRA has been published separately as EBU Tech 3293 [10].

Working context of P/Meta

As a flat list of attributes, P/Meta has value in that it helps to define a common set of definitions for metadata items required for B2B exchange, and a common format for the types of data field that would hold that metadata. However, the real strength of P/Meta is the way that it makes it possible to build P/Meta attributes into powerful sets whereby much more information can be meaningfully conveyed. This allows a fairly limited number of attributes to be used in numerous combinations in order to provide a very flexible and accurate means of putting together data about media items.

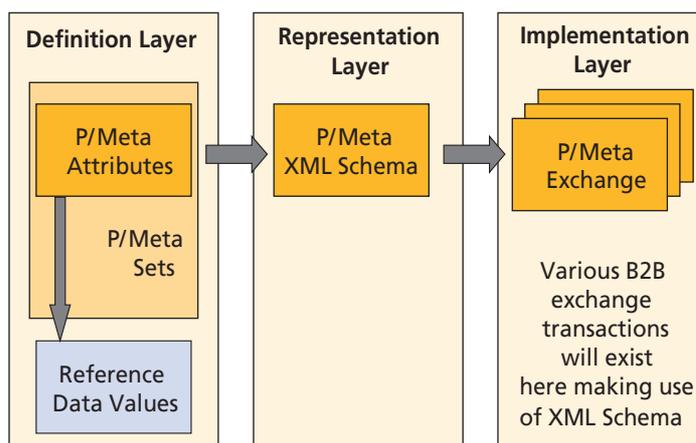


Figure 3
How P/Meta will work

This facility is made even more powerful by defining common reference data values for the P/Meta attributes, wherever this is practical to do, so that users not only have a common understanding of the structure of the data they are dealing with, but crucially its content as well.

Having defined the structure of the P/Meta data in a logical sense, there is a need to realize it physically. There are various ways in which this could be done, dependent on the technology to be used. But our current concentration is on producing an XML Schema (*see Fig. 3*). This, in turn, would provide a variety of XML elements that could be used to power real-life transaction interfaces similar to those envisaged in the IBC 2001 demonstration of P/Meta (*see Appendix A*).

What happens next with P/Meta?

The logical scheme has been developed in Workpackage 1. The mapping to related schemes and subsequent validation thereby completes this workpackage. To be completed soon are tests with transactions from real business use, feedback from the creation of XML Schema, and from broadcast technology tests (with embedding in media streams). Version 1.0 of the logical scheme will then be made available by Summer 2002.

As mentioned already, there is a requirement for “reference data” – standard values, represented in a standard way, so that they can be shared across schemes with a consistent meaning. In some cases, reference data are already available from existing internationally-agreed schemes such as the ISO language, country and currency codes. These show the clear benefit to be won if a similar approach is adopted for other reference data sets.



Richard Hopper is Principal Technical Strategist with the Media Data Group, BBC Technology. He is responsible for technical integration strategy for metadata, engagement with relevant standards activities, and for defining the procurement specifications to ensure conformance of systems with metadata requirements.

Mr Hopper graduated from Imperial College of Science & Technology, London, in 1966 and joined the BBC as a graduate trainee. He moved at an early stage into television systems engineering for production and delivery. He led the Technology Group in BBC Technology Consulting & Projects for several years, before moving into the Media Data Group at its inception in 1997.

The Media Data Group's commission is to provide coherent information support for analysis and data modelling throughout the business life-cycle of programme production, archive and delivery – from commission through to home consumption.

The work on XML Schema will also be published as “P/Meta XML”.

Members will then be able to (i) define business-to-business exchange transactions that are appropriate to their requirements, (ii) deploy P/Meta XML and (iii) implement applications.

The project as a whole is wider. Workpackage 2 has addressed requirements for unique identifiers for material and editorial content, and Workpackage 3 is addressing requirements for the practical use of unique identifiers and for technical exchange.

Proposals for management systems by which to register users and to provide support and maintenance, are being developed in consultation with the EBU.

Acknowledgements

The author acknowledges his debt to colleagues in the EBU P/Meta project – Laurent Boch (RAI), Andreas Ebner (ARD/ZDF/IRT), Peter Mulder (NOB), Dave Seditas, John Jordan and Andy Carter (BBC Technology) – who together have done most of the work underpinning this article – and, particularly, to the Project Manager, Carol Owens (BBC Technology), for her leadership and unstinting commitment to the fulfilment of the vision in which structured information delivers real value to the business; and to the Chief Technology Officer, BBC (John Varney) for the commission to engage in this strategic work.

Bibliography

- [1] EBU P/Meta Metadata Exchange Scheme, V0.9
http://www.ebu.ch/pmc_meta.html
- [2] EBU Technical Review No. 284, September 2000
http://www.ebu.ch/trev_284-hopper.pdf
- [3] EBU Escort 2.4 system for the classification of RTV programmes
http://www.ebu.ch/tech_escort2-4.pdf
- [4] SMPTE Metadata Dictionary as specified in SMPTE RP210a
<http://www.smp-te-ra.org/mdd/index.html>
- [5] Material eXchange Format (MXF)
<http://www.pro-mpeg.org>
- [6] TV-Anytime website
<http://www.tv-anytime.org>
- [7] EBU Project Group P/FRA
http://www.ebu.ch/pmc_fra.html
- [8] Dublin Core Metadata Initiative
<http://dublincore.org/>
- [9] AAF Association
<http://www.aafassociation.org/>
- [10] EBU doc. Tech 3293-E: **EBU Core Metadata Set for Radio Archives**
http://www.ebu.ch/tech_t3293.html

Further reading:

- [11] BBC Standard Media Exchange Framework
<http://www.bbc.co.uk/guidelines/smef/>

Appendix A: The P/Meta demonstration at IBC 2001

A good example of how P/Meta can be the basis for the future exchange of content-related information between producers, archives and deliverers was given in the demonstration at IBC 2001. *Fig. A1* shows a high-level system view.

This demonstration simulated the transfer of clips and metadata between four different organizations using a purpose-built application, simulating real-life system functionality using actual screenshots.

Four organizations provided clips, metadata, and screen styles, in their local languages:

- NAA (in Dutch);
- BBC (in English);
- SWR(in German);
- RAI (in Italian).

To show how P/Meta can facilitate the transfer of such information, visitors to the EBU stand at IBC were able to search, select and view the clips and metadata within one organization and simulate the delivery (a “push” type transaction) to another. They were also able to edit some of the metadata for an item to personalise the demonstration.

The demonstration was in three parts, with each part showing a different aspect of the P/Meta exchange. Users could choose between:

- the simple transfer of a clip and its accompanying metadata;

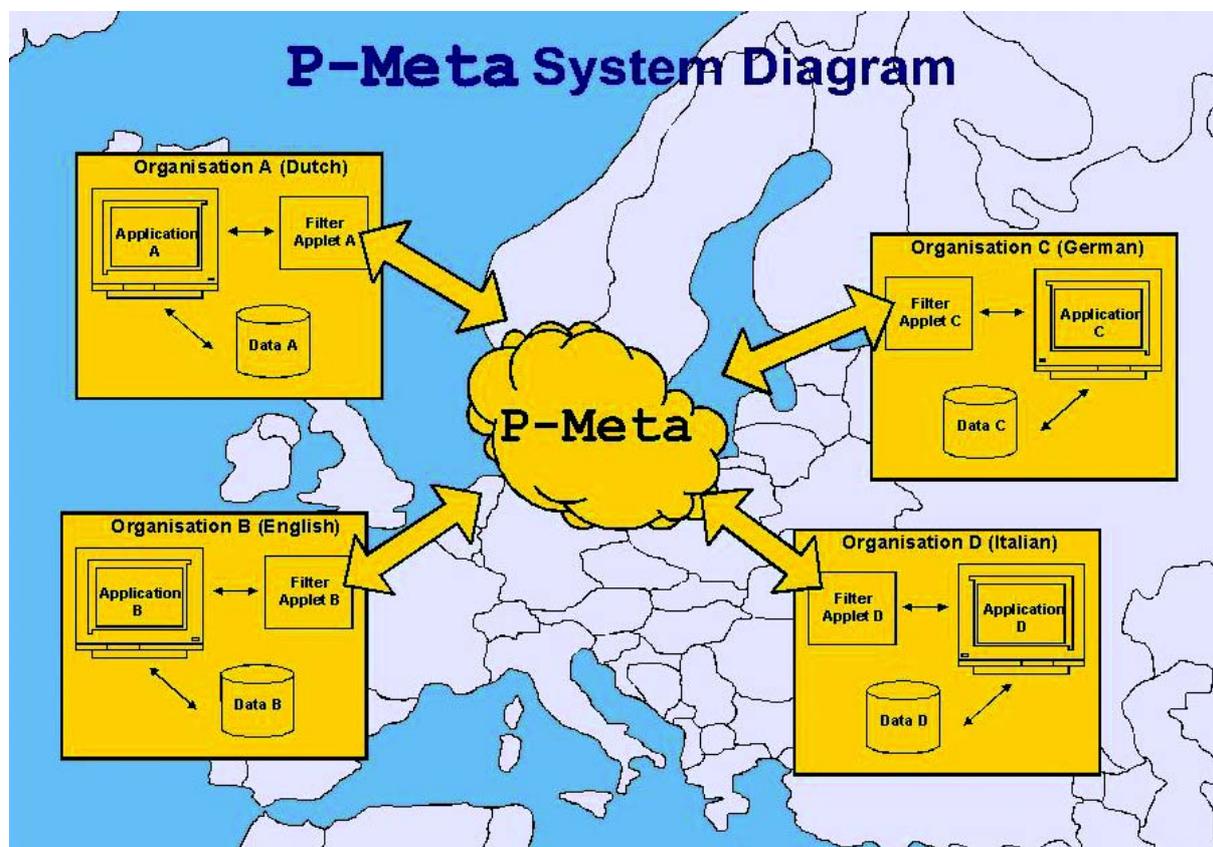


Figure A1
A P/Meta “system” view

- the transfer of usage information (such as feedback to a programme producer, giving details of when a programme sold to another broadcaster actually went to air);
- a more complicated example showing the use of some of the complex data sets supported by P/Meta.

Click on any of these images to download a larger PDF version that can be printed to an A4 printer in landscape format

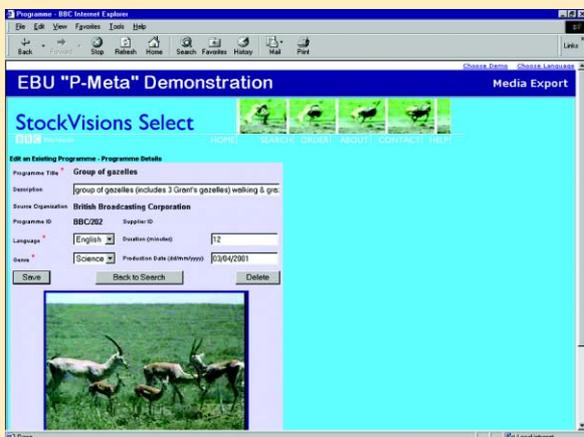


Figure A2
Source selection screen

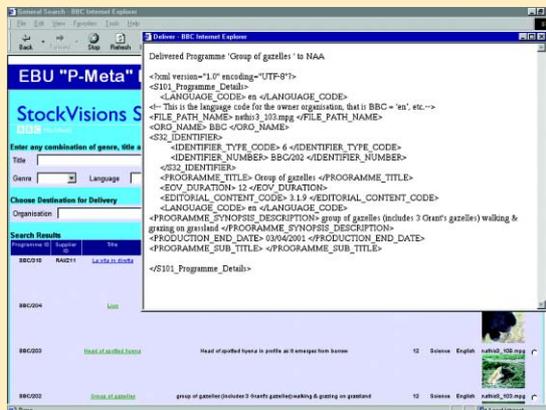


Figure A3
Source metadata in XML exchange document

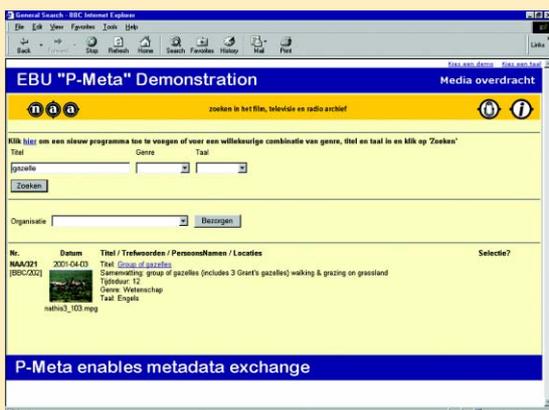


Figure A4
NAA screen showing received clip and metadata from the BBC

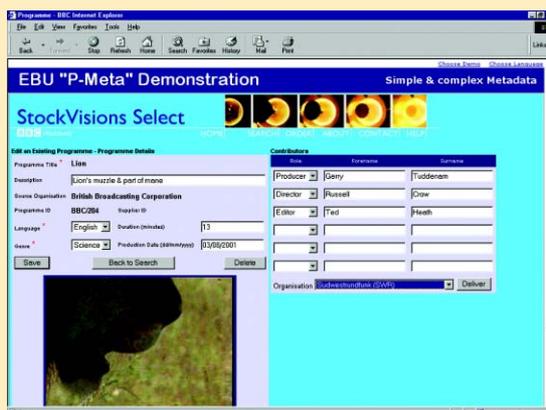


Figure A5
Source screen showing multiple contributor fields

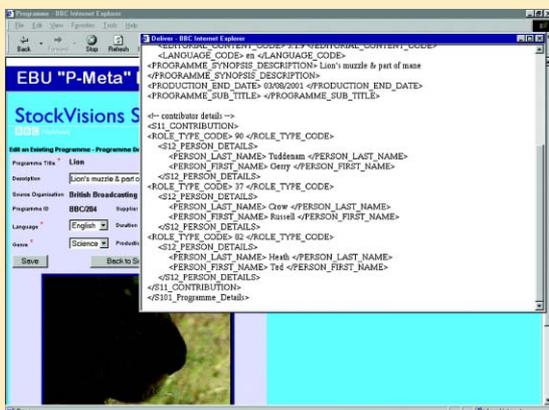


Figure A6
XML document showing the Contribution set

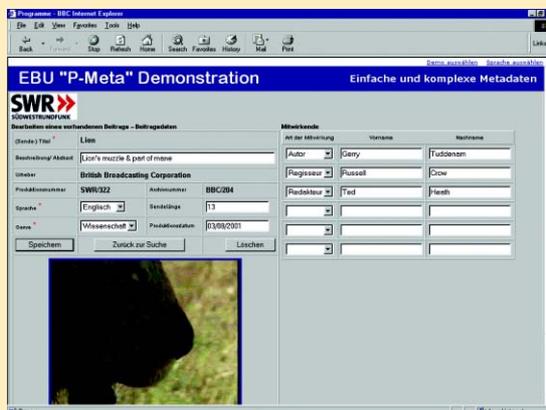


Figure A7
SWR display of Contribution Set

The P/Meta transaction between organizations was shown as a pop-up XML-style display, showing the P/Meta tags and the corresponding metadata values.

In the first example described here, we will see how P/Meta might be used to facilitate the transfer of a clip from the BBC in the UK to the NAA in the Netherlands. The user at the BBC is able to request information from the system about available BBC clips, via a menu (*Fig. A2*). The screen contains references to many P/Meta attributes and, by entering specific values for these, the user is able to select items which meet their choice criteria, in this case – wildlife footage about gazelles.

Once the clip has been selected, it can be sent to NAA. *Fig. A3* shows how the P/Meta transaction might look as an XML transfer. Notice the P/Meta attributes and values. These would be received in the Netherlands by the NAA in a similar format, having been passed via P/Meta. The media assets and accompanying metadata would be transferred successfully using the common template provided by P/Meta. Additionally, those P/Meta data items which are codified values would automatically be displayed in Dutch rather than in English.

Thus we can see (*Fig. A4*) that the field displayed earlier as “Language” with a value of “English” has become “Taal” with a value of “Engles”. The “Genre” field has remained “Genre” because the word is the same in English and Dutch, but the value has changed from “Science” to “Wetenschap”. Furthermore, the field named “Title” in the English version has become “Titel” in the Dutch version, but the content of the field is still “group of gazelles” because this field is not a coded value and thus cannot be automatically interpreted across languages.

In some cases there may be a need to transfer large groups of repeating data. A good example of this is when data concerning contributors to programmes is to be transferred. This may involve passing on the names of the roles involved in a production, and the names of the people filling them. In the example described now, we will pass on the names of the Producer and Director of a wildlife clip about lions, but the same technique is easily applicable to other repeating groups. In fact, there is no reason why the entire end credits of “Lord of the Rings” – with all the acting and supporting roles – should not be passed on in the same way.

In *Fig. A5* we see how a list of contributors to a programme could be filled in by the BBC in the UK, ready for transfer to SWR in Germany. In this case, the roles are production roles selected from a pre-set list. This will enable P/Meta to automatically populate the field names for “Producer” and “Director” into their German counterparts. The names of characters in a drama production could still be entered as free-form text but, of course, these would appear as the original free text. (Note, there is no translation as such: each coded value is referenced in a table, to provide the local preferred term in the local language.)

In the XML code that accompanies the transfer (*Fig. A6*), we can see how the “S11 Contribution” set contains as many repetitions as needed to carry all the data which has been entered on the screen.

Each occurrence of the Contribution set contains all the P/Meta attributes needed to carry details of any kind of contribution to a production. In our example, this allows all the details to arrive for display on the SWR screen in Germany (*Fig. A7*).
