

## オーディオビジュアルコンテンツのためのメタデータ国際標準

早稲田大学 大学院国際情報通信研究科

亀山 渉

## 1 はじめに

21世紀のキラアアプリケーションはコンテンツであるという認識の下、コンテンツ流通が様々な角度から検討されている。通信の分野ではコンテンツ配信ネットワーク (CDN: Contents Distribution/Delivery Network) の検討が盛んであり、実用的な見地からの検討もかなり本格化している [1]。また、放送の分野ではデジタル放送と蓄積メディアの進展により、PDR (Personal Digital/Data Recorder) を利用した蓄積型放送方式 (日本では「サーバ型放送」と呼ばれる) も実用化が間近に迫っている状況にある [2]。

このようにコンテンツが流通するための技術やインフラが整備されれば、氾濫するコンテンツの中から自分が興味のあるコンテンツのみを検索して利用するためのユーザインターフェース構築が重要になることは当然で、この観点から、オーディオビジュアルコンテンツのメタデータは主として検討されていると言って良い。しかも、サービスの性質上、様々な機器間やサービス間でのインターオペラビリティが重要であり、この意味で、国際標準化が1990年代の終りから活発に進められてきた。

国際標準化は、主として、ISO/IEC JTC1/SC29/WG11[3] (以下 MPEG と呼ぶ) と TV-Anytime フォーラム [4] によって進められおり、前者は MPEG-7 と俗称される国際標準、後者は TV-Anytime メタデータとして有名である。本稿では、これらの標準の概要について述べる。

## 2 メタデータの役割

オーディオビジュアルコンテンツにおけるメタデータの役割は、他のマルチメディアデータに対するメタデータの役割と基本的に変わることはない。しかしながら、人に視聴されることが最終目的であるコンテンツに付随するメタデータであるため、多少特筆すべき点がある。以下にその役割をまとめる。

## 2.1 検索情報記述としてのメタデータ

一般に、メタデータに主として求められるのはこの役割である。コンテンツをコンパクトに記述し、様々な利用側面から検索が可能となるように、細分化された構造を持ったメタデータが各種の国際標準では規定されている。

## 2.2 権利情報記述としてのメタデータ

コンテンツ利用時における権利処理のための情報記述が、とりわけオーディオビジュアルコンテンツでは重要である。これには買い取りや課金のための情報だけでなく、どのような利用が許されるのかといった情報まで記述できる必要がある。

例えば、蓄積型放送においては、あるコンテンツは蓄積しても視聴が許される、あるコンテンツはリアルタイムの放送時にのみ視聴が許され蓄積は許されない、といった利用条件の区別がコンテンツ毎に必要な。この他にも、ある番組を見た場合にのみ別のある番組の視聴が許される、といったような特定のビジネスモデルに直結したような利用条件も記述できることが望ましい。

このような権利情報や利用条件に関する情報記述を TV-Anytime フォーラムでは RMPI (Rights

Management and Protection Information) と呼び、いわゆる DRM システムが様々な利用許諾処理を行うための情報を提供するものと定義している。RMPI の詳細はまだ定まっていないが、tvax[6] と呼ばれる XrML (Extensible Rights Markup Language[5]) の語彙データが TV-Anytime フォーラムで検討されている。また同様の語彙の検討は MPEG-21[7, 8] として MPEG でも検討されており、こちらの語彙拡張は mx と呼ばれている。

### 2.3 誘引情報記述としてのメタデータ

コンテンツの側面を記述する上述した 2 種類のメタデータがあれば利用に関して特に問題がなさそうに思えるが、実際にはこれだけでは不十分で、コンテンツ流通は期待できない。

一般にコンテンツを流通させるためにはメタデータこそ流通させる必要があり、コンテンツ流通のための技術やシステムの研究開発も重要であるが、メタデータを流通させるための技術やシステム、更にはビジネスモデルの検討も重要であると言える。このようなメタデータ流通環境では、コンテンツの魅力や余すところなく記述できるメタデータを定義し、それを流通させることがビジネス上重要となってくる。

このことから、TV-Anytime フォーラムでは、メタデータをアトラクタ (Attractor) とも呼んでおり、コンテンツがエンドユーザに利用されるための最重要情報であると位置付けている。この意味から、ユーザを誘引するための様々な付加的な情報も有用なメタデータとなり得る。つまり、コンテンツそのものの情報ではないが、コンテンツの運用や利用に関わる情報もメタデータとして利用価値が非常に高いと言える。

例えば、あるコンテンツを 1 回視聴する料金は 50 円であるというようなものは、従来の考えでは先に述べた RMPI の一部であり、権利処理機器の中で利用時に内部的に処理できる情報であれば良いと考えられていたが、このような情報はエ

ンドユーザを誘引するための情報として、コンテンツ利用時だけでなく、エンドユーザを視聴に誘うために常時有益な情報である。このようなエンドユーザを誘引するためのメタデータ情報を TV-Anytime フォーラムでは E-Fluer (電子チラシといったような意味) と呼び、メタデータの新しい部分情報として標準化の作業を始めたばかりである [9, 10]。

### 2.4 本稿の範囲

以上のように、一口にメタデータといっても、オーディオビジュアルコンテンツでは様々な側面があり、2.2 や 2.3 で述べたような特有のメタデータ記述も標準化が進んでいる。そのどれもが興味深く、技術的にも面白い内容を含んでいるが、紙面の都合と標準化途中であるという理由から、本稿では取り上げない。以下、本稿では 2.1 の観点から、国際標準方式について述べて行く。

## 3 MPEG-7

### 3.1 概要

MPEG が標準化した初めての非圧縮符号化方式として注目されたのが、俗に MPEG-7 と呼ばれるマルチメディア用メタデータ記述のための国際標準である。正式には “Multimedia Content Description Interface” と呼ばれ、2003 年 9 月末現在、下記に示す複数のパートから成り立っている (いくつかの Amendment 文書や Corrigendum 文書が発行されているがここでは省略する)。

1. ISO/IEC 15938-1:2002, Part 1: System
2. ISO/IEC 15938-2:2002, Part 2: Description definition language
3. ISO/IEC 15938-3:2002, Part 3: Visual
4. ISO/IEC 15938-4:2002, Part 4: Audio
5. ISO/IEC 15938-5:2003, Part 5: Multimedia description schemes
6. ISO/IEC 15938-6:2003, Part 6: Reference software
7. ISO/IEC FDIS 15938-7, Part 7: Conformance testing

8. ISO/IEC TR 15938-8:2002, Part 8: Extraction and use of MPEG-7 descriptions
9. ISO/IEC WD 15938-10, Part 10: Schema definition

この内、パート1はMPEG-7で記述されたデータを圧縮、伝送、操作するための規格、パート2はメタデータを記述するためのスキーマ言語定義、パート3からパート5までがメタデータの詳細な記述方法の規格となっており、これらがMPEG-7国際標準の重要な柱を構成している。これらの関係を図1に示す。

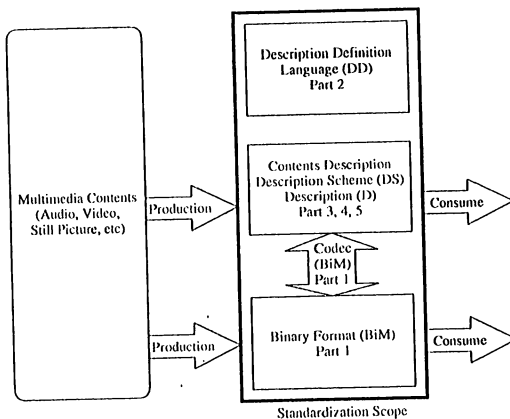


図1: MPEG-7の標準化範囲

### 3.2 システム

後で述べるように、MPEG-7のメタデータはXMLを基にして書かれるが、一般に、XMLで書かれた文書は情報量の観点からは必ずしも小さいものではないという問題点がある。特にMPEG-7では、スキーマ定義だけでも7000行を越すという膨大な規定であるため、詳細にメタデータを記述するとその量は非常に大きなものとなる可能性がある。MPEG-7そのものは汎用性のあるマルチメディア用メタデータ標準であるが、それでも、デジタル放送などでこれを用いる場合、貧弱なメモリしか持たないセットトップボックスでこれを扱わなければならないといったことも十分考えられるため、何らかの対策が要求

条件として上げられていた。

パート1で主として規定するのは、この問題を解決するためのXMLバイナリ化方式であり、BiM (Binary Format for MPEG-7) と呼ばれている。BiMは大変興味深いバイナリフォーマットで、XMLデータを圧縮するだけでなく、バイナリ状態のままXML中の部分木を選択的に復号、伝送、削除、追加することができるような仕組みを備えている。このことから、先に述べたような環境でも十分実用的なXML処理を行えるようになってきている。

### 3.3 DDL

パート2で規定されるMPEG-7 DDL (Description Definition Language) は、XMLスキーマに若干の拡張を施して定義されたスキーマ言語であり、その拡張は以下のようなものである。

#### 1. 行列を表すデータタイプの付加

次元を示す正の数をアトリビュートとして指定することで、行列表現が可能になるようにするスキーマ定義。例えば、図2のようなインスタンスの作成が可能。オーディオビジュアルの信号特徴は行列で記述される必要があることから追加された。

#### 2. 時間を示す単位の追加

ISO 8601に準拠するフォーマットとして、時刻は“YYYY-MM-DDThh:mm:ss:nnn.ff FNNN±hh:mm”、時間長は“(–)PnDTnHnMnSnNfnF±hh:mmZ”と表現。“nnn.ff NNN”は一秒当たりNNNフレーム存在しそのnnn.ff番目という意味であり、“nNnfnF”は同様に一秒当たりn(3番目のn)フレーム存在しそのn.n(1番目と2番目のn)フレーム番目という意味になる(“5N1f30F”で、一秒当たり30フレームで5.1フレーム目の意味)。映像ではフレーム数で時間を指定することが多く、その表現が可能となるようにした。

```
<IntegerMatrix mpeg7:dim="2 4">
  1 2 3 4
  5 6 7 8
</IntegerMatrix>
```

図 2: 2×4 の行列の定義

### 3.4 低レベルメタデータ

MPEG-7のメタデータは大きく分けて、低レベルメタデータと高レベルメタデータとに分類することができる。ここでいう高低は、人間の意味レベルにおいての高低を示している。つまり、低レベルメタデータはオーディオビジュアル情報から自動的にあるアルゴリズムによって抽出できる原信号の特徴を直接的に表す情報であり、高レベルメタデータはこのような自動抽出ができないオーディオビジュアル情報の意味内容を表す情報である。この様子を図3に示す。

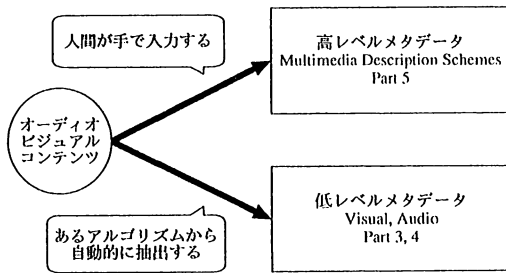


図 3: 低レベルメタデータと高レベルメタデータ

低レベルメタデータは、パート3に映像情報に関するものが、パート4に音声情報に関するものが定義されている。これらのメタデータは、それぞれ、ビジュアル記述子、オーディオ記述子とも呼ばれている。

#### 3.4.1 ビジュアル記述子

映像信号の特徴を効率良く表すものとして、以下のものが規定されている。

1. 色特徴記述子  
画面中の支配色、色の使用頻度、画面上での色配置といった特徴量を示す。
2. テクスチャ記述子

一様な模様が存在、エッジの含まれる割合といった特徴量を示す。

#### 3. 形状記述子

映像中の形を持つオブジェクトのシルエットの特徴量を示す。

#### 4. 動き記述子

被写体の動きパラメータ、被写体の重心の軌跡、動きの激しさや方向、カメラモーションといった特徴量を示す。

#### 5. 位置指定記述子

映像の時空間的な位置を指定し、指定した位置に他の情報を付与する。

#### 6. 顔特徴記述子

人間の顔を表現する記述子で、顔認証などに使用される。

#### 7. コンテナツール

様々な方法で映像情報の部分集合を指定し、それに対してビジュアル記述子を付与する。

先にも述べたように、このような特徴量を抽出するアルゴリズムが標準文書中に定められており、ある特徴量を記述すれば、誰でも矛盾のない均一な検索結果が得られるのがビジュアル記述子の特徴である。

#### 3.4.2 オーディオ記述子

オーディオの低レベルメタデータでは、少し意味内容の高いものも含まれており、ビジュアル記述子とは少々異なった内容のメタデータが定義されている。単なる記述子だけではないことから、オーディオ記述スキームと呼ばれている。

1. オーディオシグニチャ  
オーディオ信号の統計的な特徴量からオーディオ固有の値を指定する。
2. ティンブラ  
楽器の音色の印象を音の信号特徴と組み合わせで記述する。
3. サウンド認識とインデキシング  
サウンドのクラス分類やインデキシングのためのメタデータを記述する。

4. 発話内容

発話認識を行うための補助データの記述を可能とする。

5. メロディ検索

メロディの推移外形や詳細を記述する。

ビジュアル記述子と違い、ある程度 of アプリケーションを想定したようなメタデータが利用できるのがオーディオ記述スキームの特徴である。

3.5 高レベルメタデータ

MPEG-7 の高レベルメタデータは、MDS (Multimedia Description Schemes) とも呼ばれ、パート 5 に詳細が規定されている。規定されている MDS の種類を図 4 に示し、以下に説明を加える。

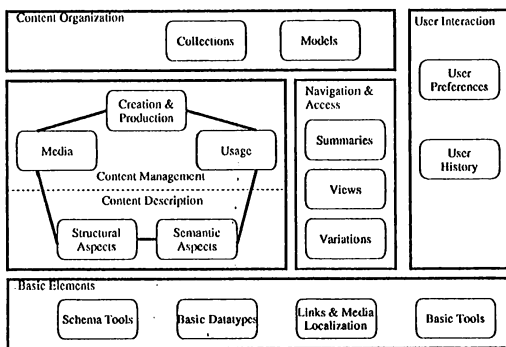


図 4: MPEG-7 MDS

1. 基本要素 (Basic Element)

MPEG-7 メタデータの基本的な構造や要素型の規定の集合を含む。

2. コンテンツ管理 (Content Management)

コンテンツの外部情報として、制作関連、メディア関連、利用形態関連の規定を含む。

3. コンテンツ記述 (Content Description)

コンテンツの構造とその意味内容について記述する規定を含む。

4. その他

ユーザの嗜好やコンテンツ使用履歴などを記述する規定などがある。

MPEG-7 MDS で記述できるメタデータは非常に範囲が広く、スキーマ定義も膨大な量である。これは、MPEG-7 MDS が汎用的に作られているためである。例えば、B2B で交換されるメタデータと B2C で交換されるメタデータでは、記述すべき内容が異なり、様々なメタデータ要素の重要性も異なっていると考えられる。しかしながら、先にも述べたように、MPEG-7 MDS は汎用的であるため、あるアプリケーションにとって必ずしも重要でないメタデータ要素を含んでしまっている。そこで、MPEG-7 MDS を実用的に使用するためには、アプリケーション毎のプロファイル作りが重要であると考えられるが、MPEG での議論はそれほど進んではない状況にある。

4 TV-Anytime メタデータ

4.1 概要

1999 年に設立された TV-Anytime フォーラムでは、近未来に PDR が放送の視聴形態を根本的に変えてしまうであろうことを念頭に、その環境における重要な技術をインターオペラビリティが可能なように規定することを目標に設立された業界標準化団体である。2003 年 9 月末現在、業界横断的に約 70 の団体がメンバとなり、活発な活動を続けている。

TV-Anytime フォーラムでは、技術仕様をフェーズ 1 とフェーズ 2 と呼ばれる 2 つのフェーズに分けて標準化活動を進めており、フェーズ 1 は先に述べた PDR を中心とした蓄積型放送アプリケーションを前提としている。フェーズ 2 の詳細は本稿の範囲を越えるので、文献 [9, 10] を参照して頂きたい。

2003 年 3 月にフェーズ 1 仕様は検討を終了し、規格文書としての安定性と受け入れを考え、ETSI (European Telecommunication Standards Institute) の技術文書として発行することをメンバー一致で選択した。このため、ETSI のサイトから無償で規格書をダウンロードすることができるようになっている [11]。

TV-Anytime フォーラムがフェーズ1で規定する技術要素は、大きく分けて以下の4つに分けることができる。

1. メタデータ
2. コンテンツを参照するための参照ID 規定とIDを同定する方法
3. 権利処理と保護のための技術規定
4. メタデータ交換プロトコルとその保護手法

メタデータは、MPEG-7とは異なり、PDR 環境におけるB2Cでのメタデータ利用を明確なアプリケーションとして想定しており、実用性の非常に高い規定となっている。事実、このメタデータ仕様は、ヨーロッパのデジタル放送規格を推進するDVB (Digital Video Broadcasting)、北米のデジタル放送規格を推進するATSC (Advanced Television Systems Committee)、及び日本の電波産業会で採用され、放送アプリケーションでは広く受け入れられるメタデータ仕様となっている。以下、このメタデータの構造の概要とメタデータ交換プロトコルについて述べる。

#### 4.2 メタデータの構造

MPEG-7を基にして作られているが、具体的なアプリケーションをあらかじめ設定しているため、不要なものは削り、記述が足りない場合には新たに要素を定義し直し、欠けているものは新たに追加を行っている。TV-Anytime メタデータは、<TVAMain>をXMLのトップ要素として、いくつかのテーブルを中に包含する構造となっている。以下にそれらのテーブルが何を情報として含むことができるかを示す。

1. ProgramInformationTable  
コンテンツに対する情報を保持しており、いわゆる「番組」の情報を持っている。
2. GroupInformationTable  
複数のコンテンツに対する情報を保持しており、シリーズ番組等を表現できる。関連性のある番組を一括して検索する場合等に有効な情報を持っている。

#### 3. ProgramReviewTable

コンテンツに対する批評情報を保持し、評価の結果やレーティング情報を付加することができる。

#### 4. ProgramLocationTable

コンテンツがどこにあり、どのようなプロトコルを使用してアクセス可能かについての情報をもっている。TV-Anytimeのコンテンツモデルでは、同一タイトルのコンテンツに対して複数のサービスがあることを許しているため、このような情報が必要となる。放送として放映スケジュールが組まれたコンテンツなのか、オンデマンドでアクセスできるコンテンツなのか等が記述できる。

#### 5. SegmentInformationTable

コンテンツのハイライト視聴やブックマークの張り付けを許すため、コンテンツの部分(これをセグメントと呼ぶ)を指定する情報を保持する。MPEG-7にも同様の構造があるが、新たに定義し直している。

#### 6. CreditsInformationTable

コンテンツのクレジット情報を保持しており、誰がどこで制作したかやコピーライトの情報等を保持している。

#### 7. ServiceInformationTable

コンテンツサービスに関する情報を保持しており、どのようなサービスが、どの期間、誰によって行われるか等の情報を保持している。

#### 8. ClassificationSchemeTable

メタデータの様々な記述のクラス分けや、別名の付与等に使われる情報を保持している。

この他、MPEG-7から視聴嗜好メタデータ (mpeg7:UserPreferencesType)、視聴履歴メタデータ (mpeg7:UsageHistory) を一部流用しており、これらの情報はユーザメタデータと総称されている。

4.3 メタデータ交換プロトコル

以上のようなメタデータは、例えば MPEG-2 TS の中に含まれてコンテンツと併せて送られたり、オフラインでセットトップボックスが何らかの媒体から読み込んだりすることが想定されている。基本的に、TV-Anytime フォーラムはこのようなメタデータの送り方に対して何ら規定を設けてはいない。というのも、放送は地域的な規定によって運用されていることから一つの方式に絞ることが困難であることや、方式を決めることによって限定的なビジネスモデルしか与えられないことを避けるためである。

一方で、近未来の放送とインターネットの融合を考えると、インターネットは先に述べた地域的な限定性がなく、また、放送を補完する重要なサービスを提供するものと考えられることから、TV-Anytime フォーラムはインターネット上でのメタデータ交換方式については規格を定めた。これが “Metadata Service over a Bi-directional Network” と呼ばれる規定で、パート 6 として標準化された [11]。ここで規定されている方式は基本的に Web サービスのフレームワークを使用するものであり、先に述べた TV-Anytime メタデータを様々な方法で検索できるインターフェースを備えている。概要を図 5 に示す。

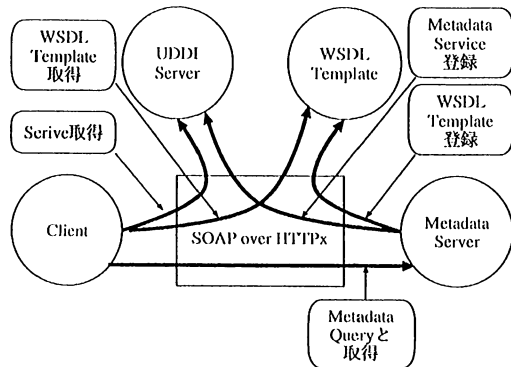


図 5: Web サービスを利用したメタデータ交換

また、このサービスでは、メタデータサーバからの一方向的な情報の流れを仮定しているだけ

でなく、エンドユーザからサーバに対する情報のアップロードも考えられている。例えば、あらかじめ自分のプロフィール情報などをサーバに送っておけば、より自分に適した検索をサーバで行える可能性がある。このため、一部のデータは安全に送る必要があり、図 6 に示すようなセキュアプロトコルスタックを使用する方式が規格化されている。

TVA XML (Metadata)
SOAP
HTTP
TLS/SSL
TCP
IP

図 6: メタデータ交換プロトコルスタック

いずれにせよ、このようなメタデータ交換プロトコルは TV-Anytime フォーラムが初めて規格化したものであり、将来、様々な応用や拡張によって、メタデータ交換とその利用に基づいた新しいサービスが期待できると思われる。

5 その他のメタデータ標準

その他のメタデータ標準として重要なのは、Pro-MPEG フォーラム [12] と AAF アソシエーション [13] が共同で開発した MXF (Media Exchange Format) である。これは、B2B におけるオーディオビジュアル素材やオーディオビジュアルパッケージの交換を目的に作成されたメタデータであり、いわゆるボスプロなどの処理に大いに利用されるものと期待されている。事実、BBC や CNN では、MXF 対応の放送機器を大量に導入しようとしており、放送業界では急速に普及するメタデータ標準となることが予想される。

一方で、このような B2B メタデータを B2C メタデータ、つまり TV-Anytime メタデータにどう

やって効率的に変換するかという課題もあり、この点に関して、TV-Anytime フォーラムと Pro-MPEG フォーラムは情報交換を始め、その可能性や具体的な手法について検討を開始しようとしている。

## 6 むすび

本稿では、オーディオビジュアルコンテンツのためのメタデータ国際標準として、MPEG-7 と TV-Anytime メタデータについて概説した。

MPEG-7 は汎用的に作られた規格であるため、様々なメタデータを交換するためのピボットとしての利用が考えられるが、現在までにそのような動きは顕著化していない。一方で、映像や音声の特徴量を利用して、映像データベースや音声データベースでの検索方法として利用する方法はいくつか事例が報告されている。

これに対し、TV-Anytime フォーラムは、PDR を利用したコンテンツの利用という明確なアプリケーションを目標とし、MPEG-7 を基にして B2C で使用できる実用的なメタデータ規格を作り上げた。また、メタデータを交換するプロトコルも積極的に検討しており、近未来に予想されるリッチなメタデータサービスの基盤技術を提供していくことと思われる。

また、B2B においても MXF というメタデータ標準が今後広まり、オーディオビジュアルコンテンツが円滑に流通するためのメタデータ基盤が整いつつあると言うことができるだろう。

以上のように、オーディオビジュアルコンテンツ市場では、メタデータを使用したコンテンツ流通とコンテンツ利用が急速に広まることが予想され、ここで使われるメタデータ標準が、他の分野で検討あるいは使用されているメタデータ規格に大きな影響を及ぼすことが考えられる。

インターオペラビリティの困難なメタデータは利用価値が低いため、様々なメディア間やビジネス間でのメタデータ交換が今後大きな問題となってくると考えられる。この課題をどう克服するか

が、今後の焦点の一つとなってくるであろう。

## 参考文献

- [1] 奥村, 青柳, 篠原, “CDN と通信・放送融合”, 電子情報通信学会誌, Vol.86, No.3, pp.193-197 (2003 年 3 月)
- [2] (社) 電波産業会, “サーバ型放送における符号化、伝送及び蓄積制御方式標準規格”, 電波産業会標準規格, ARIB STD-B38 (2003 年 2 月)
- [3] <http://www.chiariglione.org/mpeg/index.htm>
- [4] <http://www.tv-anytime.org/>
- [5] <http://www.xrml.org/>
- [6] TV-Anytime Forum, “RMP Specification Drafting Process, Specification Workbook”, WD550 (2002 年 3 月)
- [7] ISO/IEC FDIS 21000-5, “Multimedia framework (MPEG-21). Part 5: Rights Expression Language”. (2003 年 8 月)
- [8] ISO/IEC FDIS 21000-6, “Multimedia framework (MPEG-21) Part 6: Rights Data Dictionary” (2003 年 8 月)
- [9] TV-Anytime Forum, “Call for Contributions for TV-Anytime Phase 2 Technologies on Sharing, Synchronization, Packaging, Targeting & Interactivity”, TV179r3 (2003 年 8 月)
- [10] TV-Anytime Forum, “TV-Anytime in a Connected World, A Vision Paper for TVA Phase 2”, TV191r2 (2003 年 8 月)
- [11] ETSI, “Broadcast and On-line Services: Search, select and rightful use of content on personal storage systems (“TV-Anytime Phase 1”)", ETSI TS 102 822-1, -2, -3-1, -3-2, -4, -6, -7 (2003 年 9 月予定)
- [12] <http://www.pro-mpeg.org/>
- [13] <http://www.aafassociation.org/>