

マルチメディアとメタデータ

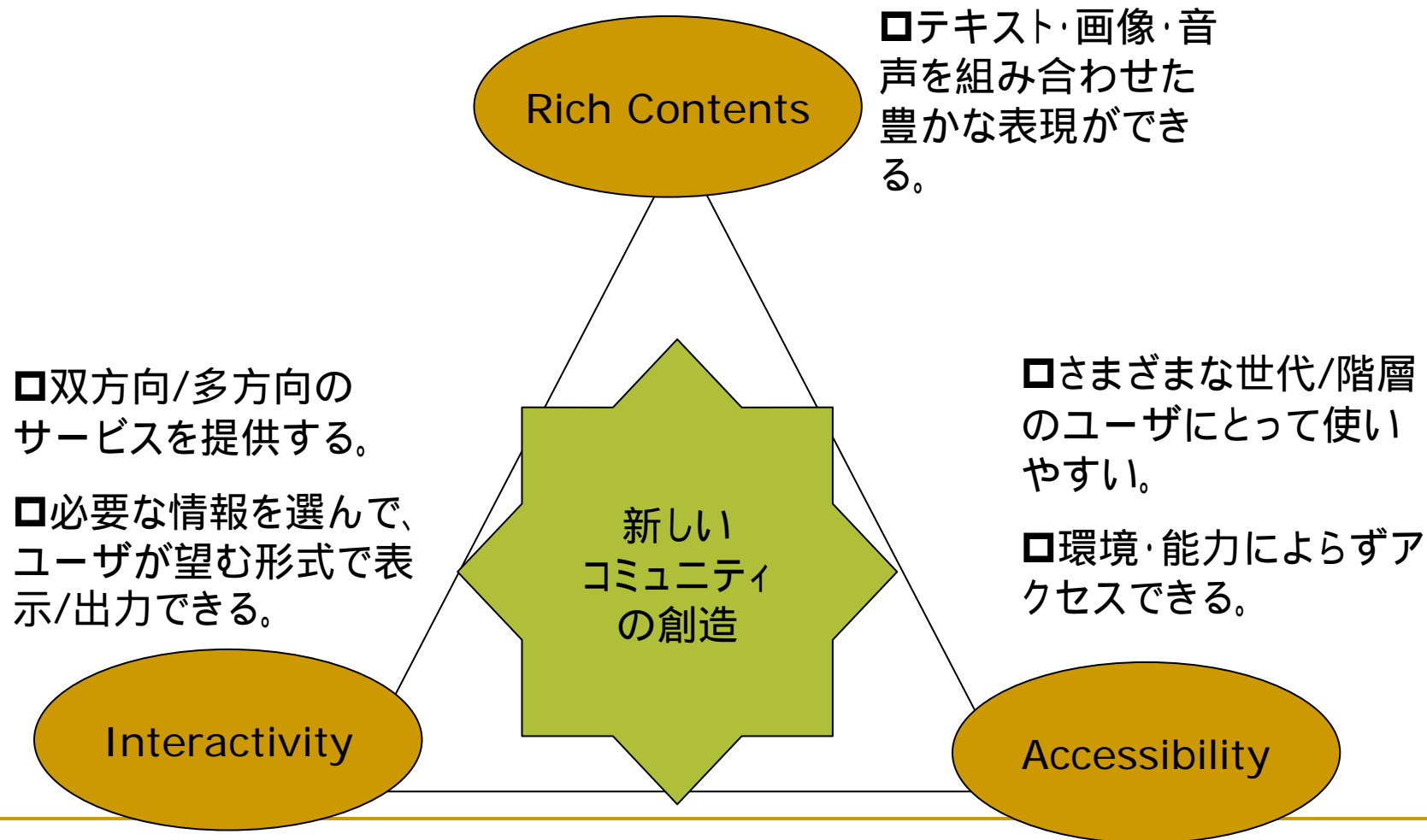
2004年1月23日

第5回XMLコンソーシアムDay

XMLテクノロジー部会 複合コンテンツWG

NTTソフトウェア 奥山信輔

マルチメディアWebがめざすべきもの



マルチメディアWebの課題

- 検索技術の発展
 - 配信・再生のための技術(圧縮、伝送など)はどんどん進歩している。
 - マルチメディアコンテンツの量が増加してくると、その中から必要な情報を抜き出す技術(検索、フィルタリング)が必要。
- 課金モデル
 - 有料のコンテンツを利用する文化が根付いていない。
 - 適切な課金モデルが必要
- 著作権・知的所有権
 - コンテンツを自動的に加工したり、一部を抜粋して組み合わせるような利用形態(ハイライト生成、ニューステロップの自動翻訳など)が考えられるが、これは原則的には著作権の侵害となる。
 - ストリーミングコンテンツの一部をダウンロードして利用する形態の場合、コンテンツに付与されているコピーガードが障害となる可能性がある。
- デジタルデバイド・アクセシビリティ
 - 居住する地域、年齢、その他のハンディキャップにより、Webへのアクセスが制限される人達が存在する。
 - 特に公共のサービス(放送、通信、電子政府など)では、ユニバーサルアクセスを実現する必要がある。

質のよいコンテンツを、効率よく探し出し、有効に利用するための仕組みが必要

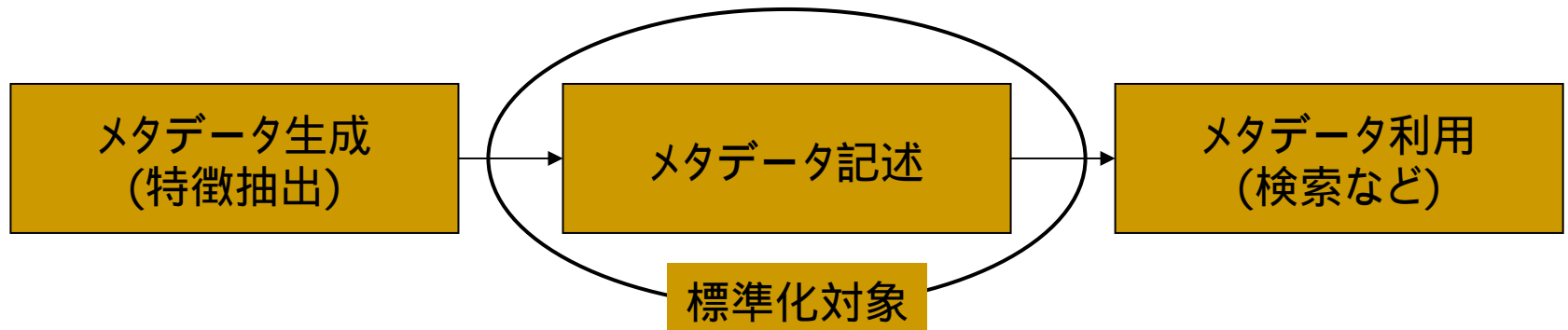
マルチメディアとメタデータ

- 検索が難しい
 - テキストに比べ、データ量・種類が膨大である。
 - 音声認識やテロップ認識などの技術があるが、認識精度・処理コスト(時間)の問題がある。
 - 作者や出演者、演奏者などの情報が明示的に含まれるとは限らない。
 - テキストによらない検索のニーズもある。
 - 鼻歌検索
～ 曲を探したいけどタイトルもアーティストもわからない。。
 - 色や形状を指定した検索
～ もう少し明るい色のカーテンがほしい。。
- 再生時間が長いことが多い。
 - コンテンツ全体が検索されるだけでは不十分で、特定のシーン・フレーズなど一部分を検索する必要がある。

コンテンツを直接検索するのではなく、コンテンツの全体または一部の特徴を記述したメタデータが必要になる。
検索だけでなく、著作権管理や利用形態の記述にもメタデータが重要である。

MPEG-7

- マルチメディアのメタデータの表記方法を定める規格
- マルチメディア特徴を記述することにより、検索やフィルタリングなどの応用を可能とする。
- MPEG(Moving Picture Experts Group)により制定(中)。
- 標準化の対象は特徴(メタデータ)を記述するための表記方法で、メタデータの抽出方法やメタデータの利用方法は対象外。



MPEG-7の構成文書

- 以下の8文書からなる。
 - Part 1: Systems
 - Part 2: Description Definition Language (DDL)
 - Part 3: Visual
 - Part 4: Audio
 - Part 5: Multimedia Description Schemes
 - Part 6: Reference Software
 - Part 7: Conformance
 - Part 8: Extraction and Use of MPEG-7 Descriptions
- Part 1 ~ 6と8は出版済み。Part 7 は2003年度中に出版予定。
 - ISO/IEC 15938-n: nは番号

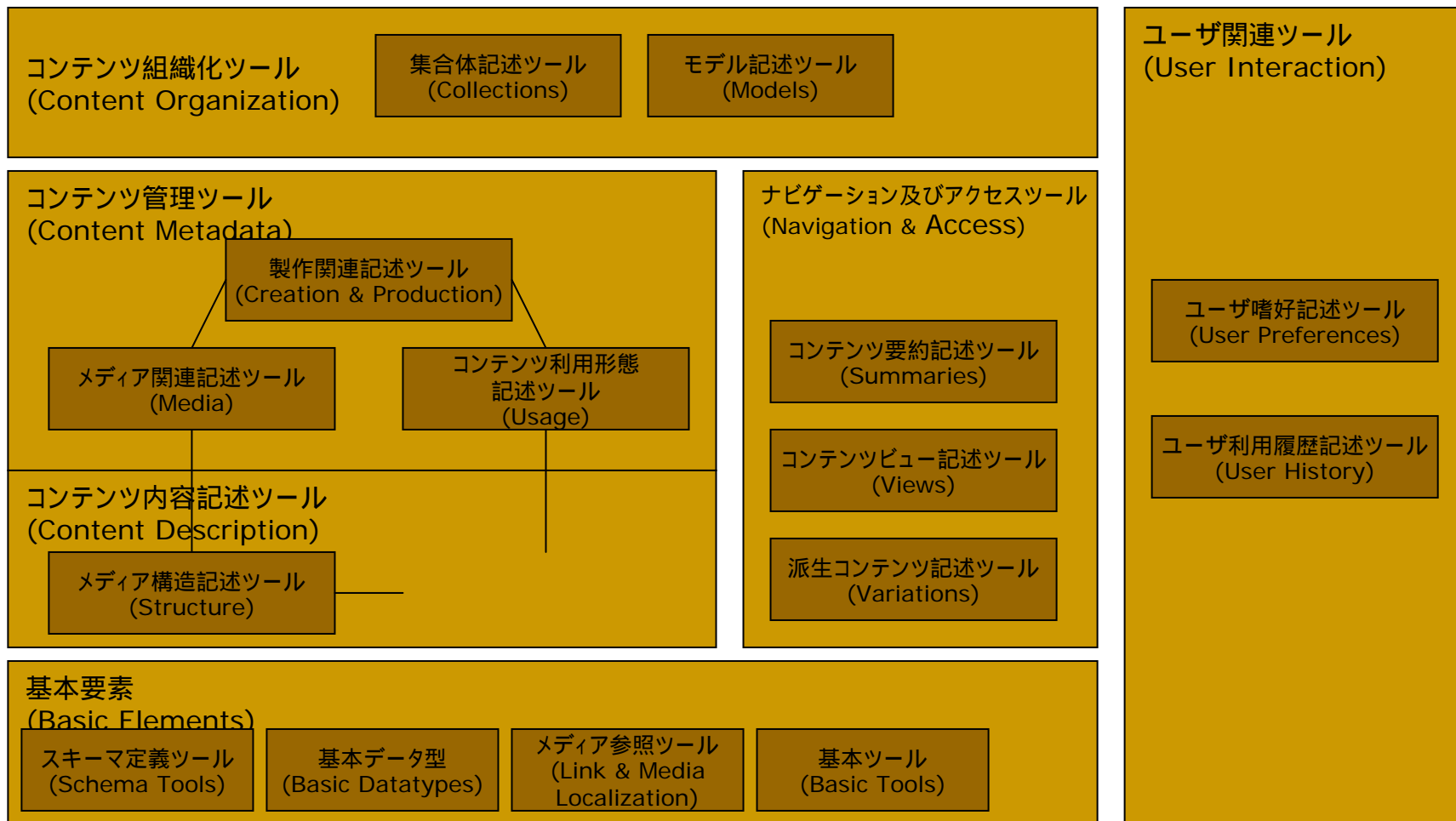
MPEG-7の標準化項目

項目	説明
記述子 (Descriptor; D)	マルチメディアコンテンツにおいて、ある単一の特徴を記述するための基本ツール。MPEG-7では表記方法(Syntax)および意味(Semantics)を規定。
記述スキーム (Description Scheme; DS)	複数の記述ツール間の構造、意味的關係を規定した枠組み。MPEG-7では表記方法(Syntax)および意味(Semantics)を規定。
データ型 (Data-Type)	特にマルチメディアコンテンツの特徴を記述するものではなく、他の記述子または記述スキームの構成部品として使われるツール。
記述定義言語 (Description Definition Language; DDL)	記述子および記述スキームの表記方法を規定するための言語。XML Schemaをベースにマルチメディアコンテンツの記述に必要なデータ型を追加。
システムツール (System Tools)	MPEG-7メタデータを効率的に伝送、蓄積し、コンテンツとの同期をとるためのアーキテクチャおよびツールを規定。

特徴量の概要

- ビジュアル特徴記述ツール
 - 色、テクスチャ(模様)、形状、動きなどを記述する低レベルの記述子
 - 位置、顔特徴、コンテナなどの記述ツール
- オーディオ特徴記述ツール
 - 波形、パワー、スペクトル特徴、基本周波数などを記述する低レベルの記述子
 - 音色、インデックス、発話内容、メロディなどの記述ツール
- マルチメディア・コンテンツ記述スキーム
 - ビジュアル・オーディオ以外の特徴記述を記述するツール。
 - テキストに基づく意味、分類、場所、感性、関係などを記述する基本ツール。
 - コンテンツの製作者、分類などを記述する管理ツール。
 - コンテンツの内容、構造などを記述するツール。
 - コンテンツの要約、ビュー(見え方)を記述するツール。
 - ユーザの嗜好や利用履歴を記述するツール。
 - などなど。

マルチメディア・コンテンツ記述スキーム



MPEG-7の応用例

- ハイライト生成ツール(韓国)
 - ニュースやテレビ番組などから自動的にハイライトを生成するツール
 - 映像コンテンツからカメラの動き、テロップ、シーン・カットの境界を検出し、それをMPEG-7でデータベースに蓄積する。これから番組ジャンルごとにハイライト生成する。
- ハミング検索(MIT)
 - ある曲の1フレーズの特徴を元に、音楽データベースを検索する。
- セミナー画像検索 (日本)
 - リコーのMPMeister
 - セミナーなどのプレゼンテーション映像を資料とともにWebコンテンツ化するツール
 - 映像とスライドの同期にSMILを、検索用にMPEG-7を使用している。

MPEG-7の応用例

- 携帯電話向け動画配信 (2001年9月)
 - NTT DoCoMoと日本IBMによる共同開発
 - ユーザが指定した条件(長さ、好み)に合わせて動画シーンを自動編集し、ダウンロードする。
 - たとえば、サッカーの試合からハイライトシーンのみを選んで編集することができる。「コーナーキックの場合はその後が重要」などの編集ルールを指定できる。また、ユーザが優先的に見たいシーンを指定して編集することもできる。
 - 携帯電話のJavaプログラムがMPEG-7を解釈する。番組内容によるMPEG-7の構造の違いもこのJavaプログラムが吸収する。
 - MPEG-7のオーサリングツールもあわせて開発。

MPEG-21

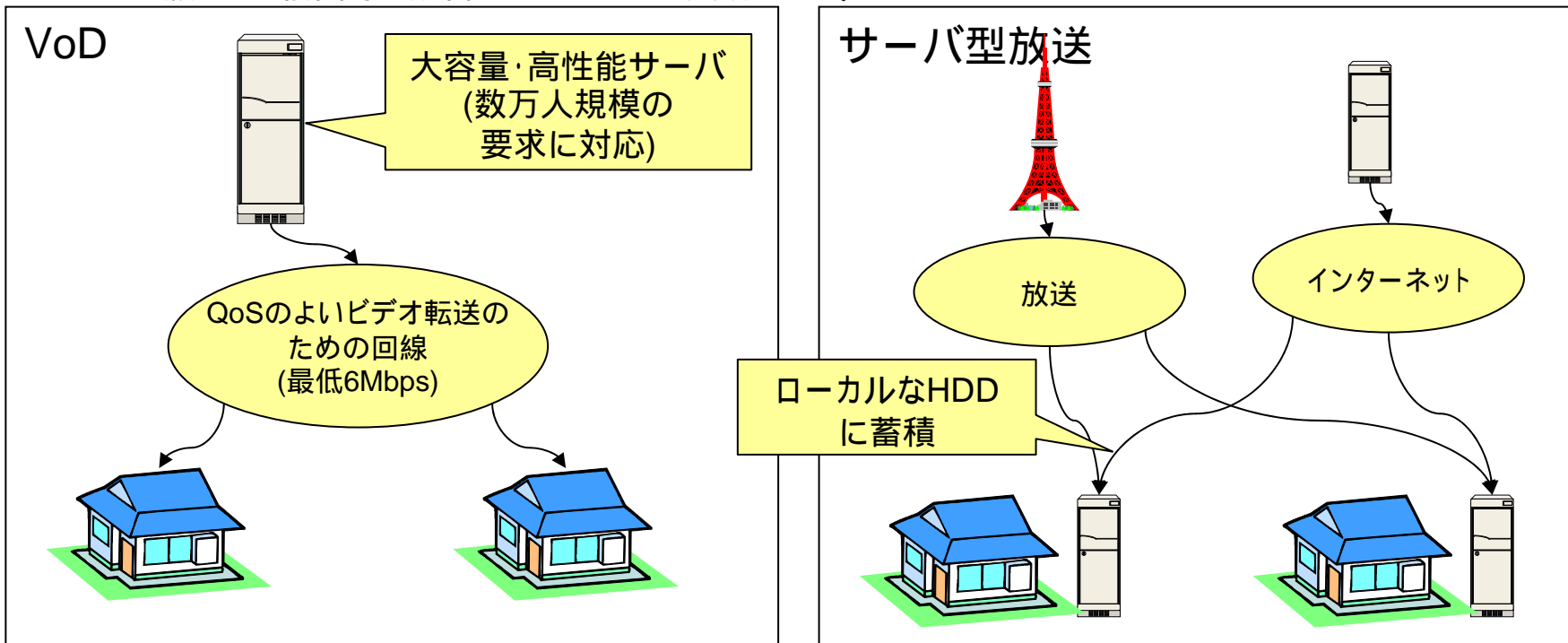
- デジタルコンテンツの幅広い利用を可能とするためのマルチメディア・フレームワーク
- 現在、通信プロトコル、映像符号化、メタデータ定義、著作権管理など、マルチメディアコンテンツ流通のインフラに必要な要素がばらばらに存在している。これらを包括的に結びつけるためのフレームワーク。
- 以下の11のパートからなる。
 - 1 Vision, Technologies and Strategy
 - 2 Digital Item Declaration
 - 3 Digital Item Identification
 - 4 Intellectual Property Management and Protection Architecture
 - 5 Rights Expression Language
 - 6 Rights Data Dictionary
 - 7 Digital Item Adaptation
 - 8 Reference Software
 - 9 File Format
 - 10 Digital Item Processing
 - 11 Evaluation Methods for Persistent Association Technologies

TV Anytime

- TV Anytime Forumで検討されている、サーバ型放送システムの標準
- DAVICでTV Anytime and TV Anywhere という概念が提唱されたのを受け、1998年から標準化が進められている。
 - DAVIC(Digital Audio-Visual Council): VoDの標準化を目指した標準化団体
- 日本からのメンバー(NHK、NTTなど)の貢献が大きい。
 - TV Anytime Forumの副議長は早稲田大学の亀山教授
 - TV Anytime メタデータ Phase2のリーダーはNTTの川森氏

サーバ型放送

- 放送、インターネットなどさまざまな方法で配信される番組(コンテンツ)を大容量のサーバに蓄積・記録し、利用者がいつでも(Anytime)視聴することのできる放送システム
- 蓄積型放送とも呼ばれる。
- VoD(Video on Demand)との違い:
 - 回線・サーバに要求される性能や品質がVoDに比べ低いので、低コストで実現できる。
 - 放送・通信回線を統合したサービスを実現できる。



サーバ型放送のメリットと課題

■ メリット

□ 利用者にとって

- 利用者の見たい時に好みのコンテンツを視聴できる。
- 好みの番組の自動蓄積、特定シーンの自動抽出など多様な視聴サービスを受けることができる。
- ニュースや天気予報など最新の情報を蓄積・更新し、参照することができる。

□ コンテンツ提供者にとって

- 放送局のコンテンツの二次利用を促進できる。
- 情報の配信時間を分散できるため、電波やネットワークの有効利用を図ることができる。

□ メーカーにとって

- 新しい送信装置、ネットワーク機器、受信装置が必要になる
- デジタル放送の市場拡大

■ 課題

□ 広告モデルとは異なる収益モデル

- コンテンツの有料化? 有料コンテンツは普及するか?
- 新しい広告モデルも必要

□ ユーザニーズに合わせた多様なコンテンツ・サービスの提供

- ただし、既に2002年からepサービスが始まっているがほとんど普及していない。。。なぜか?
- 実は「新しいサービス」はそれほど大切ではなく、今あるような番組をより便利に視聴できることが必要なのかもしれない。

□ コンテンツの有効利用を図るための検索・メタデータ技術の開発

TV Anytime Forum

- DAVICの活動終了に伴い、1999年9月に発足
- ヨーロッパ、アメリカ、アジア、日本の企業を中心とする59社で構成(2003年5月現在)
 - 主な参加企業
 - BBC(英)、BT(英)、Deutch Telecom(独)、France Telecom(仏)、EBU(European Broadcast Union)
 - NHK、NTT、フジテレビ、日本テレビ、TBS、松下、東芝、ソニー
 - マイクロソフト、サムスン、モトローラ、フィリップス

TV Anytime Forumの体制、検討内容

WG	活動内容
WG Business Models (ビジネスモデルWG)	ビジネスモデル、要求仕様、ベンチマーク(実装の基準)などの検討
WG System, Transport Interfaces and Content Referencing (システム、通信・I/F、コンテンツ・リファレンスWG)	TV Anytimeシステムのアーキテクチャ「コンテンツ参照」(コンテンツを一意に参照する仕組み)の検討
WG Metadata (メタデータWG)	TV Anytime メタデータの検討
WG Rights Management and Protection (権利の管理と保護WG)	コンテンツに関する知的所有権の管理、保護

TV Anytime メタデータ

- TV Anytimeシステムにおいて、利用者が自分の好みのコンテンツを検索するためのメタデータ
- 放送・インターネットなどの配信形態を意識せずに検索できる。
 - EPG (Electronic Program Guide; 電子番組案内)
 - デジタル放送で使用される番組案内
 - ECG (Electronic Contents Guide; 電子コンテンツ案内)
 - EPGをインターネット上のコンテンツの検索にも使用できるように拡張したもの
 - ECGを実現するためのメタデータ TV Anytime メタデータ
- 基本的な考え方 / コンテンツ利用の3フェーズ

選択

コンテンツを検索し、選択する。
メタデータを使用。

位置解決

選択フェーズで特定されたコンテンツを探す。
「コンテンツ参照ID(CRID)」を使用。

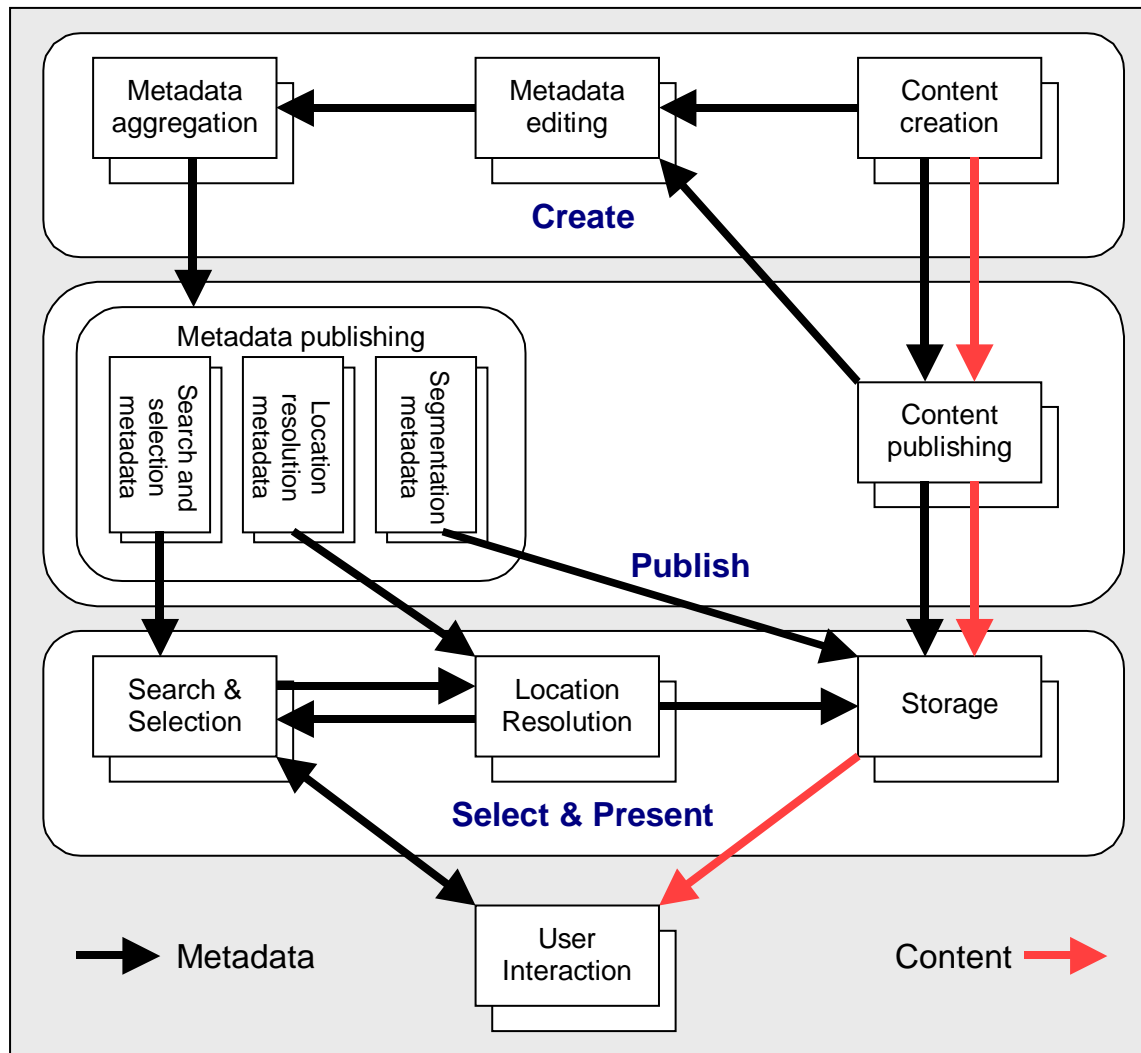
消費(視聴・蓄積)

位置解決フェーズで発見されたコンテンツを再生・録画する。
メタデータを使用。

TV Anytime メタデータの概要

- XMLに基づくフレームワーク
 - XML Schema
 - MPEG-7も参考にしている。
 - 既に存在するEPGの規格を参考にしている。
 - DVB(Digital Video Broadcasting): ヨーロッパ
 - ARIB(Association of Radio Industries and Business; 電波産業会): 日本
 - ATSC(Advanced Television System Committee): 米国
- 既存のEPGに含まれない機能も追加
 - 特定シーンの検索を可能にする「セグメント記述メタデータ」
 - 利用者の視聴動向を記録し、嗜好に合わせた番組の視聴や蓄積を可能にする「視聴履歴メタデータ」「嗜好メタデータ」
- 最新版はV1.3の修正第3版
 - <http://www.tv-anytime.org/ftp/index.html>
 - ftp://tva:tva@ftp.bbc.co.uk/pub/Specifications/COR3_SP003v13.zip
 - 仕様だけでAパート(160ページ)、Bパート(70ページ)のドキュメント
- ARIBや総務省でも採用の方針
 - ARIB STD-B38「サーバー型放送における符号化、伝送及び蓄積制御方式」でTV Anytimeメタデータ、MPEG-7の採用を規定。
 - 総務省「映像コンテンツ等の著作権処理の円滑化に資するメタデータ体系」でもTV Anytimeメタデータとの相互変換を考慮。

TV Anytime メタデータ プロセスモデル



メタデータの種類

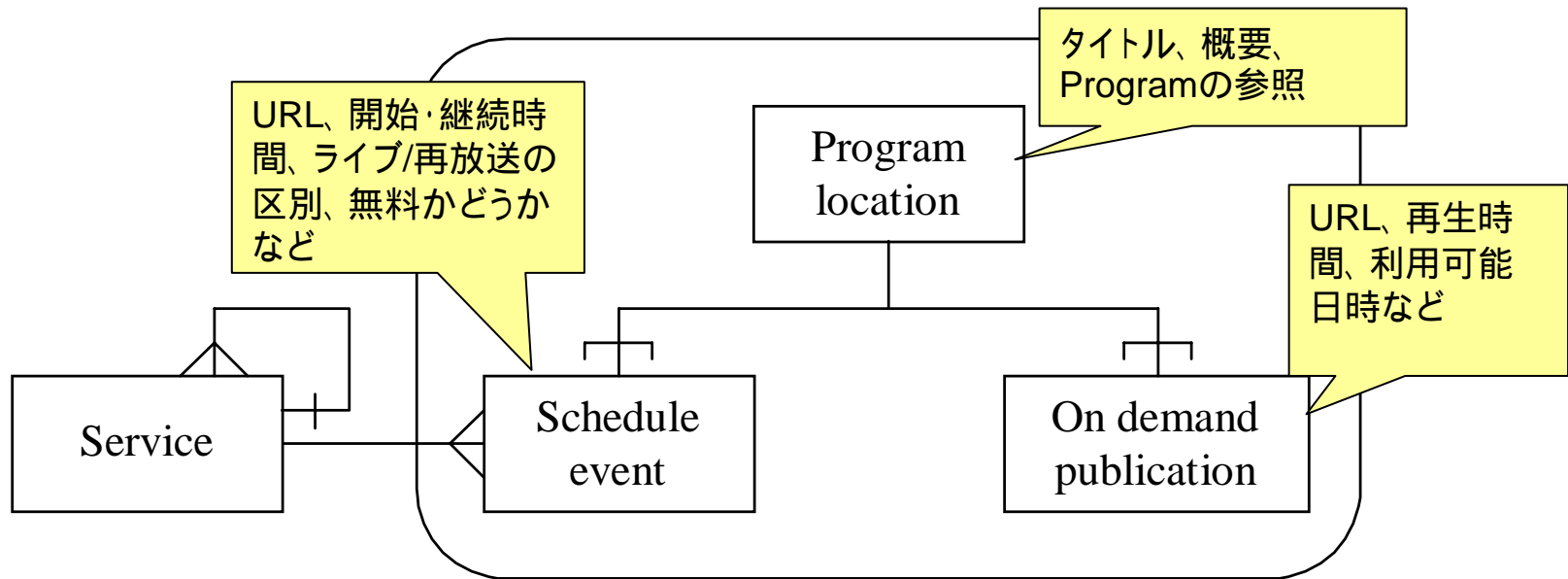
- 以下の4種からなる。
 - コンテンツ記述メタデータ (Content Description Metadata)
 - コンテンツのリリースや放送形態に依存しない一般的な情報
 - タイトル、内容、ジャンルなど
 - インスタンス記述メタデータ (Instance Description Metadata)
 - コンテンツの特定のインスタンスに関する情報
 - インスタンスのロケーション(場所や放送時間)や配信メディアなど
 - セグメント記述メタデータ (Segmentation Metadata)
 - コンテンツに含まれるセグメント(時間で区切られた部分)に関する情報
 - セグメントのタイトル、概要、キーワードなど
 - 視聴者に関するメタデータ (Consumer Metadata)
 - コンテンツの視聴履歴や視聴者の嗜好などの情報

コンテンツ記述メタデータ

- コンテンツとして Program と Program Group を記述できる。
- ProgramはProgramIdで識別される。ProgramIdはコンテンツ参照ID(CRID)
- Program Program、Program Program Group、Program Group Program Groupの関係はそれぞれ多対多で記述できる。
 - つまり、ひとつのProgramが複数のProgram Groupに属することができる。
 - Program自体が複数のProgramから構成されることも許される。(Aggregated Program)
- 基本コンテンツ記述として以下の情報を記述できる。
 - Title, MediaTitle, ShortTitle, Synopsis, PromotionalInformation, Keyword, Genre, ParentalGuidance, Language, CaptionLanguage, SignLanguage, CreditsList, AwardsList, RelatedMaterial, ProductionDate, ProductionLocation, CreationCoordinates, DepictedCoordinates, ReleaseInformation
- Review情報も記述できる。
 - フリーテキストによる評論、レーティング、評価者情報など

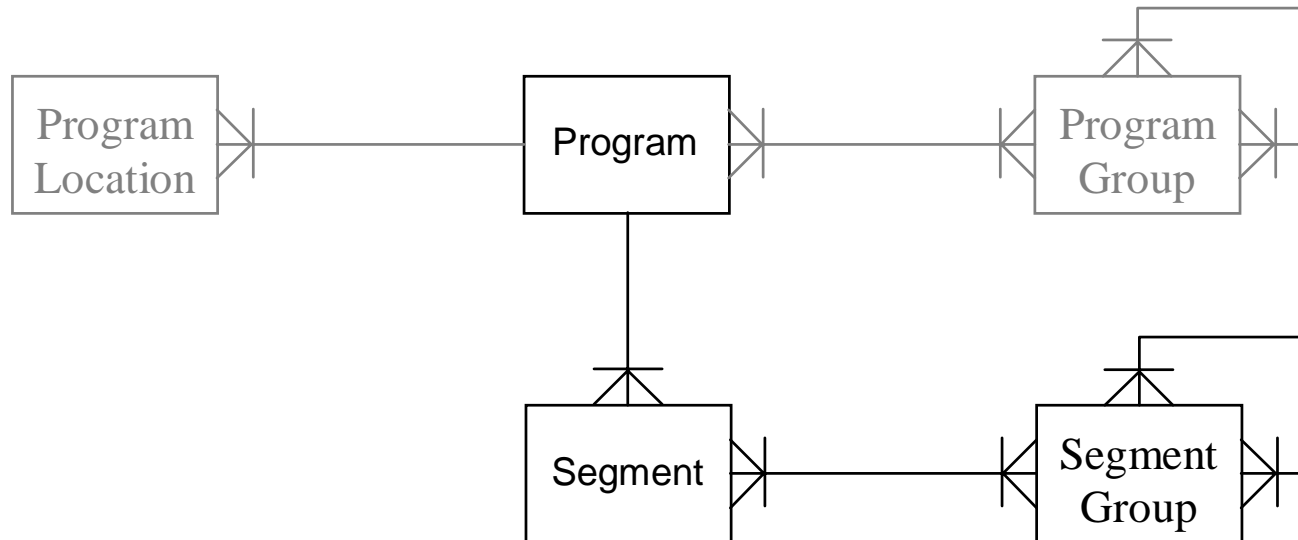
インスタンス記述メタデータ

- ProgramLocationによりProgramの個々のインスタンスを記述する。
- ProgramLocationとProgramはCRIDで対応付けられる。
- 1つのProgramLocationは1つのProgramにしか対応させられない。Program Groupに対応させることもできない。
- ProgramLocationは、主にインターネット配信に対応する On demand publication、主に放送配信に対応するSchedule eventの2種類に分けられる。さらに、Schedule event の場合はServiceにより放送チャンネルの情報を記述できる。



セグメント記述メタデータ

- Segment、Segment Groupを記述できる。
- あるSegmentは特定のProgramに所属する。
- Segment Segment Groupの対応は多対多である。
 - つまり、ひとつのSegmentが複数のSegment Groupに属することができる。
- Segmentは以下の情報を持つ。
 - SegmentId, SegmentLocator(長さ、位置)、KeyFrameLocator(キーフレームの位置)、Title, Synopsis, Keywords, RelatedMaterial



視聴者に関するメタデータ

- Usage History(視聴履歴)、User Preferences(視聴者嗜好)の2種類がある。
- Usage Historyは、視聴したプログラムのIDや視聴時間、操作の履歴などを記述する。
 - 操作履歴により、たとえば、何度も繰り返し見たところとか、スキップしたところなどがわかる。
- User Preferencesは以下のような情報を記述できる。
 - UserIDentifier
 - FilteringAndSearchPreferences (フィルタリングや検索の嗜好)
 - CreationPreferences (制作属性; 制作時期や出演者などの嗜好)
 - ClassificationPreferences (ジャンル、言語などの嗜好)
 - SourcePreferences (配信元に関する嗜好)
 - BrowsingPreferences (視聴方法に関する嗜好)
 - SummaryPreferences (要約方針に関する嗜好)
 - PreferenceCondition (嗜好適用条件)

メタデータ文書の構造

- TV Anytimeメタデータ文書は、ルート要素(TVAMain)の下に、前述の4種類のメタデータを含む構造になる。

