Digital video broadcasting module

METADATE MANAGEMENT

h.shirazi@qmul.ac.uk

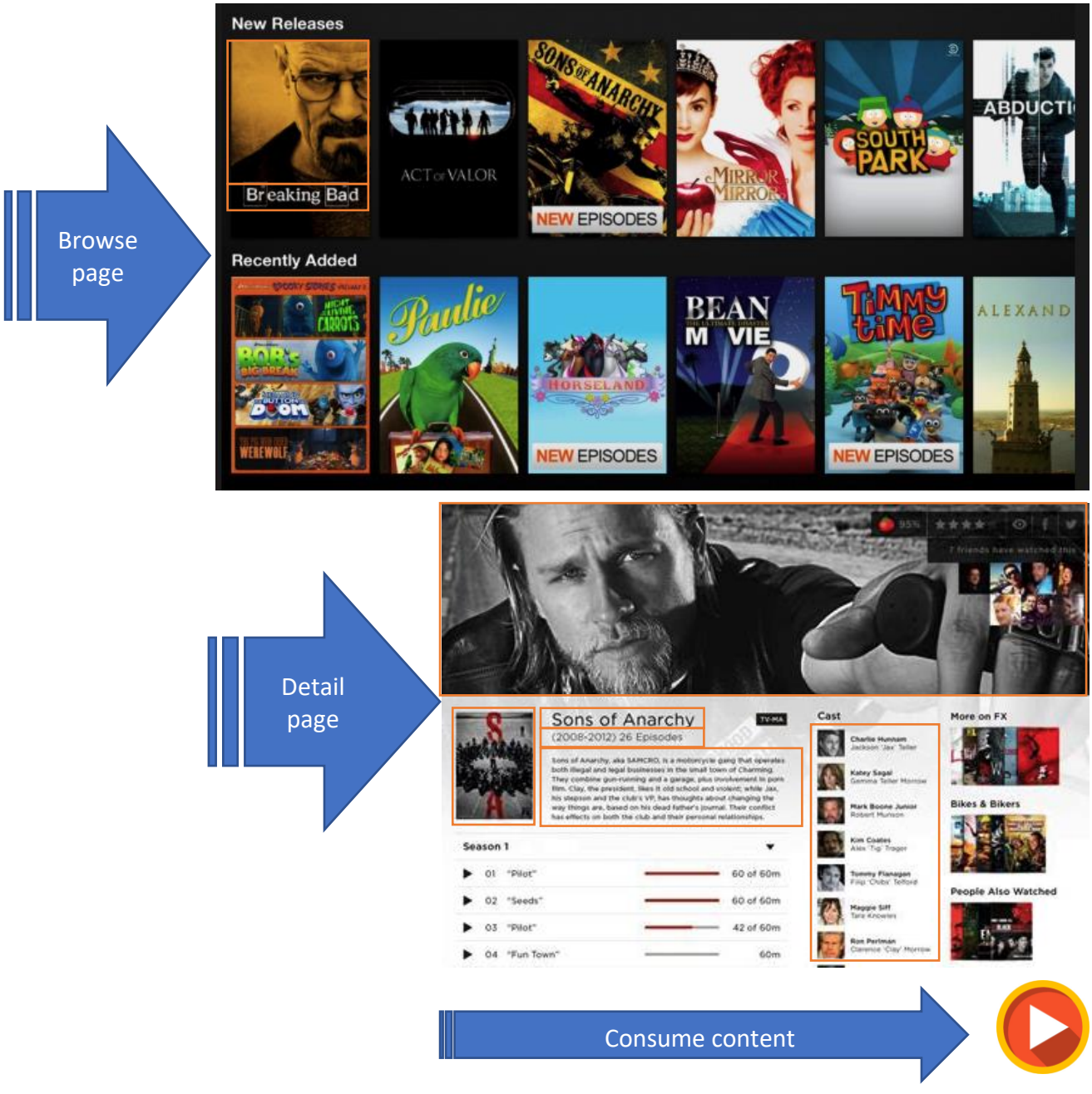
Dr Hamid Shirazi

Table of Contents

Week3: Day-2

Metadata for enriched experience

1. Configurations for IRDs (e.g. geographical information)
2. Catalog browsing, basic (e.g. text and imagery) and enriched information (e.g. IMDB ratings) about content(s)
3. Search, content discovery and recommendations
4. Rights management (business rules, studio rules, content rights, and user entitlements)



References

- Architectures:
    - benchmark business models against which the TV-Anytime system architecture is evaluated to ensure that the specification enable key business applications (TS 102 822-1[1]).
    - TV-Anytime System Architecture (TS 102 822-2[2].
- Features:
    - Metadata (TS 102 822-3-1, TS 102 822-3-2[3], TS 102 822-3-3[4], TS 102 822-3-4[5]),
    - Content Referencing (TS 102 822-4[6])
    - Rights Management (TS 102 822-5-1[7], TS 102 822-5-2[8]),
    - Bi-directional Metadata Delivery (TS 102 822-6-1[9], TS 102 822-6-2[10], TS 102 822-6-3[11])
    - Protection (TS 102 822-7[12]),
    - Interchange Data Format (TS 102 822-8[13])
    - Remote Programming (TS 102 822-914)

---

[1] ETSI TS 102 822-1: "Broadcast and On-line Services: Search, select, and rightful use of content on personal storage systems ("TV-Anytime"); Part 1: Benchmark Features".
[2] ETSI TS 102 822-2: "Broadcast and On-line Services: Search, select, and rightful use of content on personal storage systems ("TV-Anytime"); Part 2: Phase 1 - System description".
[3] ETSI TS 102 822-3-2: "Broadcast and On-line Services: Search, select, and rightful use of content on personal storage systems ("TV-Anytime"); Part 3: Metadata; Sub-part 2: System aspects in a uni-directional environment".
[4] ETSI TS 102 822-3-3: "Broadcast and On-line Services: Search, select, and rightful use of content on personal storage systems ("TV-Anytime"); Part 3: Metadata; Sub-part 3: Phase 2 - Extended Metadata Schema"
[5] ETSI TS 102 822-3-4: "Broadcast and On-line Services: Search, select, and rightful use of content on personal storage systems ("TV-Anytime"); Part 3: Metadata; Sub-part 4: Phase 2 - Interstitial metadata".
[6] ETSI TS 102 822-4: "Broadcast and On-line Services: Search, select, and rightful use of content on personal storage systems ("TV-Anytime"); Part 4: Phase 1 - Content Referencing".
[7] ETSI TS 102 822-5-1: "Broadcast and On-line Services: Search, select, and rightful use of content on personal storage systems ("TV-Anytime"); Part 5: Rights Management and Protection (RMP) Sub-part 1: Information for Broadcast Applications".
[8] ETSI TS 102 822-5-2: "Broadcast and On-line Services: Search, select, and rightful use of content on personal storage systems ("TV-Anytime"); Part 5: Rights Management and Protection (RMP) Sub-part 2: RMPI binding".
[9] ETSI TS 102 822-6-1: "Broadcast and On-line Services: Search, select, and rightful use of content on personal storage systems ("TV-Anytime"); Part 6: Delivery of metadata over a bi-directional network; Sub-part 1: Service and transport".
[10] ETSI TS 102 822-6-2: "Broadcast and On-line Services: Search, select, and rightful use of content on personal storage systems ("TV-Anytime"); Part 6: Delivery of metadata over a bi-directional network; Sub-part 2: Phase 1 - Service discovery".
[11] ETSI TS 102 822-6-3: "Broadcast and On-line Services: Search, select, and rightful use of content on personal storage systems ("TV-Anytime"); Part 6: Delivery of metadata over a bi-directional network; Sub-part 3: Phase 2 - Exchange of Personal Profile".
[12] ETSI TS 102 822-7: "Broadcast and On-line Services: Search, select, and rightful use of content on personal storage systems ("TV-Anytime Phase 1"); Part 7: Bi-directional metadata delivery protection".
[13] ETSI TS 102 822-8: "Broadcast and On-line Services: Search, select, and rightful use of content on personal storage systems ("TV-Anytime"); Part 8: Phase 2 - Interchange data format".
[14] ETSI TS 102 822-9: "Broadcast and On-line Services: Search, select, and rightful use of content on personal storage systems ("TV-Anytime"); Part 9: Phase 2 - Remote Programming".

TV-Anytime (TVA)

- TV-Anytime (TVA) is defined by ETSI[15] and only defines the metadata format for metadata that may be exchanged between various entities such as between the content provider and consumer, among consumers, or between a third-party metadata provider and the consumer.
- It provides specifications for features enable the search, selection, acquisition and rightful use of content on local and/or remote personal storage systems from both broadcast and online services.
- Metadata schema and system specifications should be both used to proper end to end metadata management:
    - XML is the representation format used to define metadata schema
        - XML Schema is used to define how metadata is represented in XML (also for non-XML representations)
    - TV-Anytime metadata specification addresses the formating of the metadata
        - Including recommended binary format, fragmentation, encapsulation of fragments and indexing of metadata descriptions.
- Transportation of metadata is out of the scope of TV-Anytime and it is speficied by DVB, ATSC and ARIB
- The manner in which metadata is stored, accessed and used on the PDR[16] (IRD) is also out of the scope of TV-Any time specifications.

[15] Broadcast and On-line Services: Search, select, and rightful use of content on personal storage systems ("TV-Anytime"); Part 3: Metadata; Sub-part 1: Phase 1 - Metadata schemas

[16] Personal digital recorder

Definitions

Application	specific set of functions running on the PDR - Some applications use metadata, either automatically or under consumer control
Attractor	metadata element that is accessible by the consumer in order to aid in the content selection process, thus attracting the consumer. E.g. title and name of an actor in a television programme
Content creator	Producer of the content
Content provider	Entity who is the prime exploiter of the content
Content reference	Pointer to a specific content item
Content reference identifier (CRID)	Identifier for the content independent of its location
Description scheme	Format defined by MPEG-7 description definition language (DLL) for metadata schema[17]
Descriptor	metadata element, such as an attractor or other information about content such as the key frame index of a piece of video
location resolution	process of establishing the address (location and time) of a specific content instance from its CRID
Metadata	data about content – data about data e.g. title, genre and summary of a television programme. In the context of TV-Anytime, metadata also includes consumer profile and history data
metadata schema	identifier associated with a set of XML schemas that globally identifies those schemas so that they can be referenced externally. A globally unique namespace ensures that the names of types defined by schemas in that namespace do not conflict with types of the same name defined elsewhere.
Metadata system	set of rules describing the syntax and semantics of metadata
MPEG-7	Ongoing effort by the Motion Pictures Expert Group to specify a standard set of content-related metadata applicable to a broad range of applications
Programme	editorially coherent piece of content. Typically, a programme is acquired by the PDR as a whole.
Programme group	One or more programmes that are grouped together TV-Anytime defines several types of programme groups such as "series" and "programme compilation".
Segment	continuous portion of a piece of content, for example a single topic in a news programme
segmentation	process of creating segments from a piece of content

[17] ISO/IEC 15938-2 (2002): "Information technology - Multimedia content description interface - Part 2: Description definition language".

TV-Anytime metadata process model

- Most visible part of metadata: attractors/descriptors or hyperlinks used in electronic programme guides, or in Web pages to acquire particular piece of content
- TV-Anytime system allows consumers to find, navigate and manage content from a variety of intenal and external sources e.g. broadcasters, interactive TV, Internet and local storage
- Also provides information about consumers (consumer profile) and their preferences
- Facilitates automated search processes using descriptive elements and attractors incl. those in EPG
- Main processes as illustrated in Figure 1 are metadata creation, pubishing, content discovery and presentation.

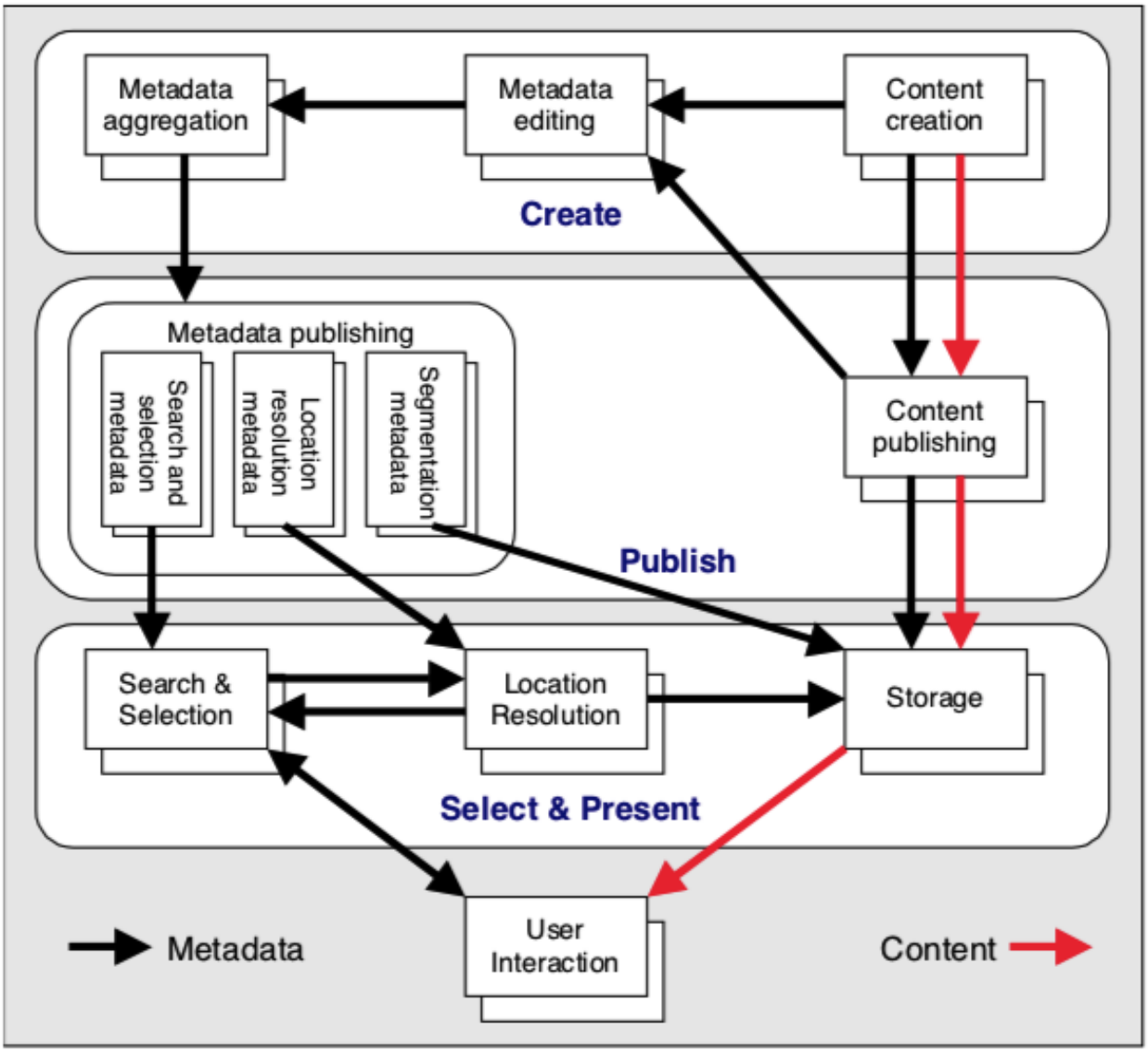


Figure 1: metadata and content flow

Metadata processes

- Content creation process
    - It represents the production of a piece of content or a programme.
        - During the production process, the programme content is created and information about the programme may also be captured.
    - At this stage, however, the metadata is unlikely to be in a form that can be directly exposed to a user - editing may be required.
- Content publishing process
    - Once content has been created, the content is then available for publication by a content publisher.
        - e.g. as part of a broadcast service or as a publication on the Internet.
    - The content publishing process defines **instantiations of programmes** – e.g. information about "where" the programme can be found.
        - In broadcasting, it means a **schedule** for the services that are published.
- Metadata editing process
    - It takes raw information from the content creation and publishing processes and edits this into a form that is suitable for representing the content to the end consumer.
    - The output of this process is edited metadata for the programmes and/or metadata describing the location of these programmes.
- Metadata aggregation process
    - It is likely that metadata from a number of independent content creators and publishers will need to be aggregated.
    - It may result in the original metadata being changed.
- Metadata publishing process
    - An aggregated metadata set will need to be published to both the content selection and location resolution processes.
    - The content selection process will be largely concerned with the metadata describing programmes.
- Content selection process
    - It occurs through the direct involvement of the consumer or may be performed by a software agent according to consumer and his preferences inferred from the consumer's past history of content selection or by the explicit specification of preferences by the user (or a combination of the two).
    - Note that the content selection process may be, in part, affected by knowledge of the programme's location.
- Location resulotion process
    - Discovering where (or when) a programme can be found. Details of this discovery process can be found in the TV-Anytime Content Referencing Specification.

TV-Anytime metadata structure model

- Simple data modelling methodology which allows us to describe metadata structure at the high level independent of any representations.
- It clearly states the relationships between TV-Anytime entities under inheritance concepts.
- Figure 2 illustrates the entity relationship in the graph syntax

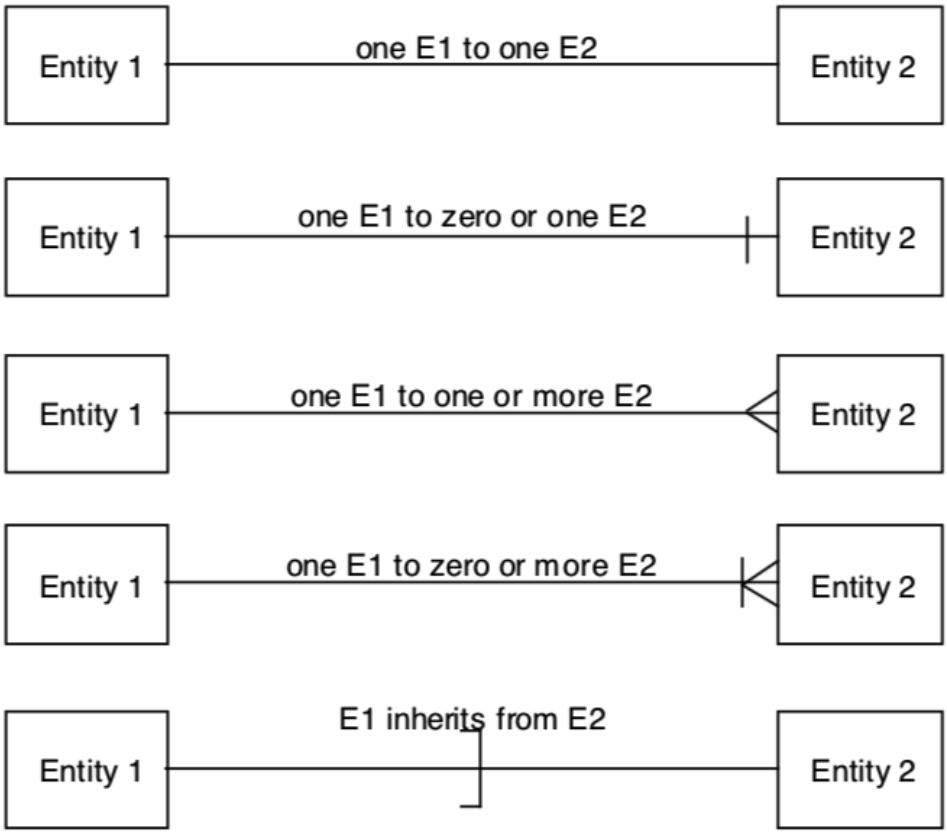


Figure 2: Basic entity relation graph syntax

TVA Identifiers

- Figure 3 depicts the relation among TVA identifiers: CRID, IMI and Locator.

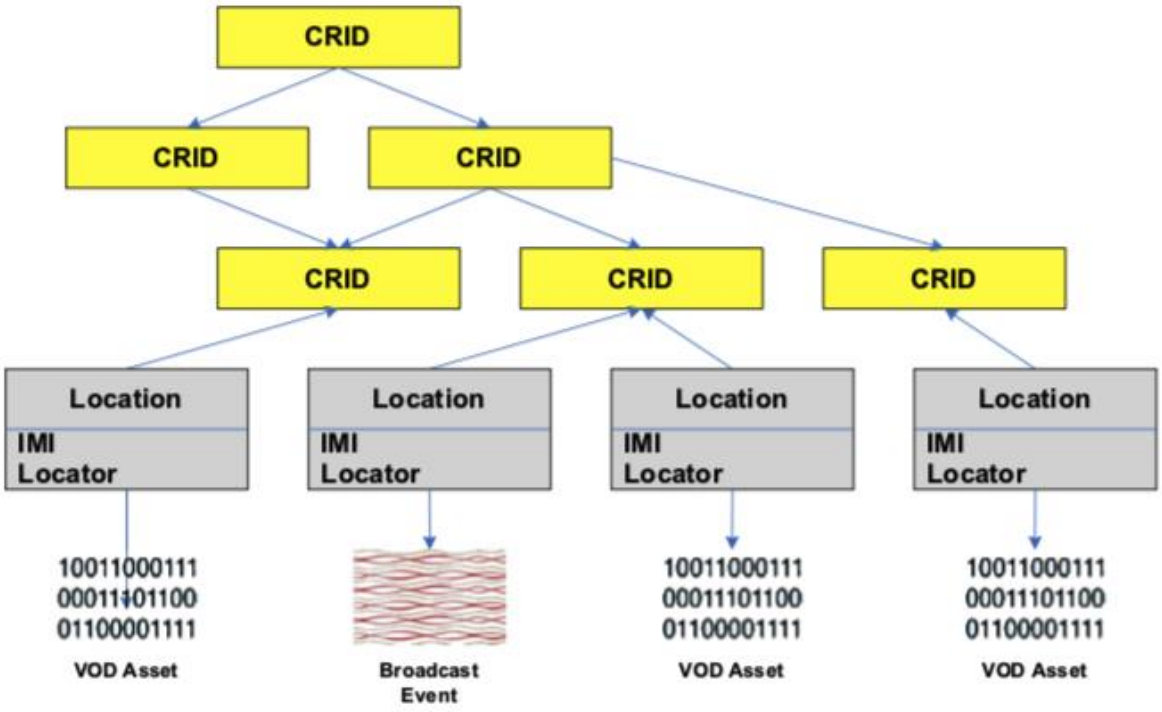


Figure 3: Relation between TVA identifiers

Content Reference Indicator (CRID)

- It's the most important topic in TV-Anytime
- CRID refers to a piece of content or one/more other CRIDs
- It links different content-related metadata descriptions
- Content-related metadata is knows as content description metadata or instance description metadata
  - content description metadata is general information about a piece of content that does not change regardless of how the content is published or broadcast.
    - It includes information such as the content's title, textual description and genre.
    - Typically, the content creator assigns content description metadata before publication.
  - Instance description metadata describes a particular instance of a piece of content, including information such as the content location, usage rules (pay-per-view, etc.) and delivery parameters (e.g. video format).
    - Instance description metadata is assigned by the content provider as a part of the publication of content. During the search and selection process, a consumer may use both general content and instance descriptions.
- Consumer related metadata includes usage history data (logging data), annotation metadata and user preferences.

❖ Syntax: CRID is an XML anyURI: CRID://<domain name>/<data>

- Whereby:
  - The entire CRID string is case insensitive.
  - <domain name>: is a registered Internet domain name or a delegated sub-domain within such a domain. (See RFC 1591 for DNS name registration).
  - The <domain name> must be a fully qualified name according to RFC 1591.
  - <data> is a free format string that is Uniform Resource Identifier (URI) compliant, and is
  - meaningful to within the given domain.

Instance Metadata Identifier (IMI)

- Each content **location** shall have an IMI, which uniquely identifies the location of an instance of the content i.e. deliverable
    - Please note that even a pull deliverable has an IMI.
- The IMI is linked to one content location.
    - So, if a recording is made of a broadcast event, both the broadcast event and the resulting VOD content location shall each have their own IMI.
    - Furthermore, when the VOD content has multiple availability windows, also each availability window shall have its own IMI.
- uniquely identifies the location of an instance of the content i.e. deliverable.
    - ❖ Syntax: IMI:[<domain name>/]<data>
        - Whereby:
            - The entire IMI string is case insensitive.
            - <domain name> shall be the same as the deployment name as defined in CRID. It shall be possible to omit the <domain name>, when it is equal to the <domain name> of the corresponding CRID
            - E.g. imi:www.eventis.nl/f47ac10b-58cc-4372-a567-0e02b111d479, shortly: imi:f47ac10b- 58cc-4372-a567-0e02b111d479, when CRID and IMI have the same domain name.

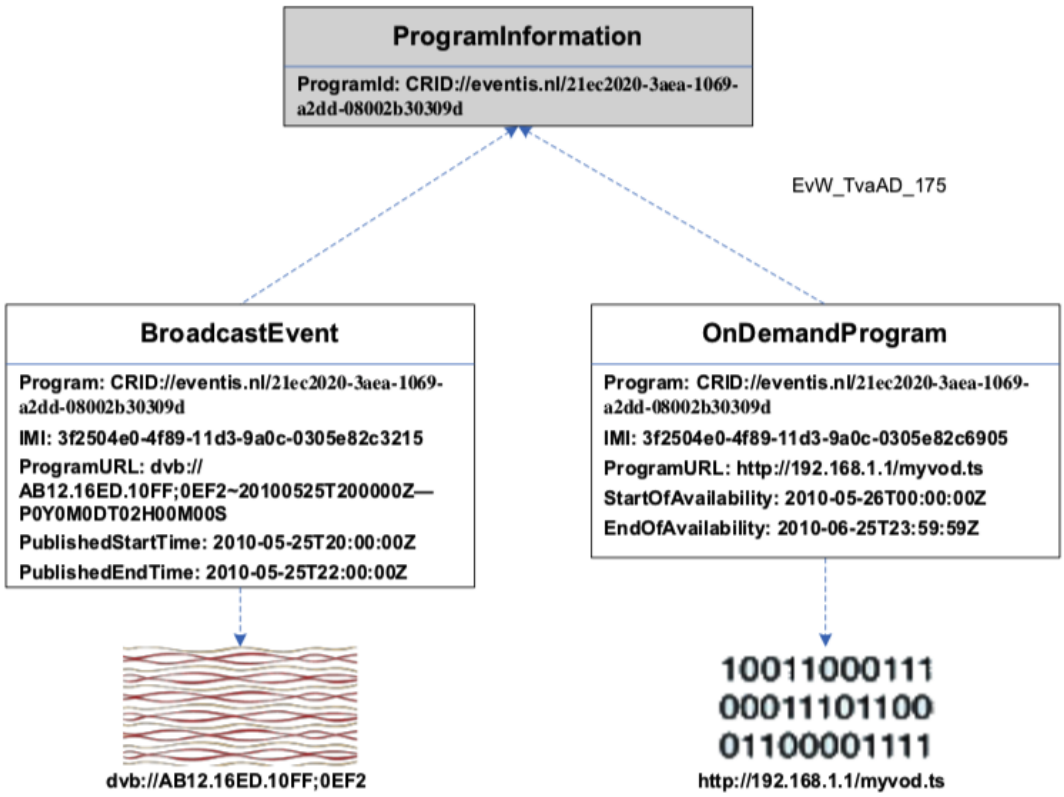


Figure 4: example of the use of CRID and IMI

TVA Locator

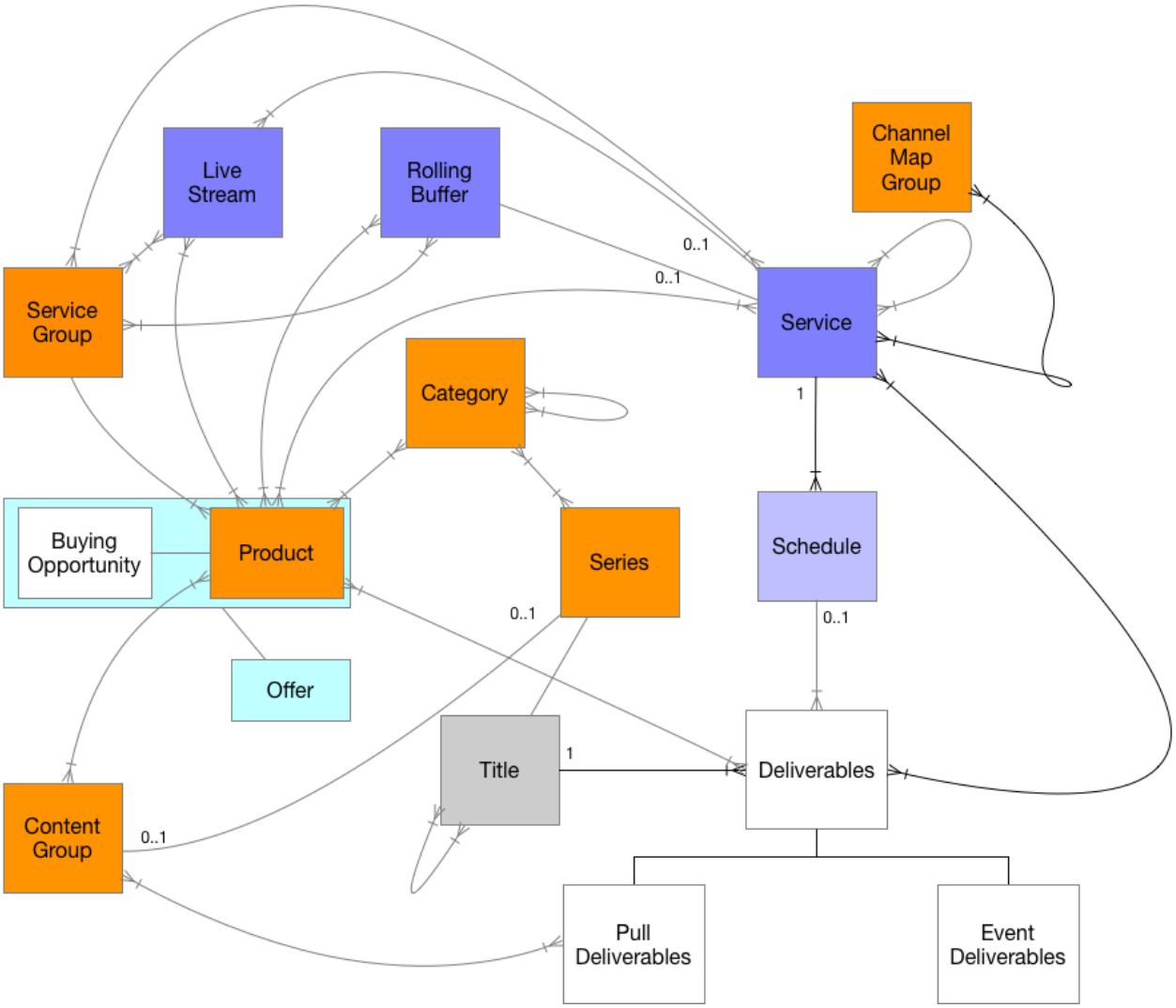
- A TV-Anytime locator identifies the location of an instance of a broadcast or VOD program (on- demand program). The locator specifies the physical location and possibly the time of availability, where and when the content in question can be acquired.
- IPTV Service Locator:
    - <transport mechanism>://<host>[:<port>][/<url-path>]
    - Transport mechanism could RTP or UDP
        - **e.g. udp://224.28.3.60:6791**
- DVB Service Locator
    - dvb://<ONID>.[<TSID>].<SID>
    - ONID (16 bits) is DVB original network identifier according to ETSI 300 468
    - TSID (16 bits) is DVB transport stream identifier according to ETSI 300 468
    - SID (16 bits) is DVB service identifier according to ETSI 300 468
    - ONID, TSID, SID are case insensitive
        - **e.g. dvb://033a.0f04.123f**
- Schedule based Content Locator
    - <service locator>[;<EID>][~<startDateAndTime>--<duration>]
    - Service loctaor typically identifies the Service URI
    - EID (16 bits) is the DVB event identifier
    - startDateAndTime format: YYYYMMDDThhmmssZ
    - duration format: PnnW or PnnYnnMnnDTnnHnnMnnS.
        - If no duration is specified, the availability is indefinitely after start date and time.
    - Example:
        - DVB event:

          dvb://033a.0f04.123f;0123~20090217T123000Z—P0Y0M0DT01H00M00S

        - IPTV event:

          udp://224.1.1.1:5555~20090328T174500Z—P0Y0M0DT01H00M00S

Entity relationship diagram



Buying opportunity	The possibility for a customer to buy a product for a given price, in a specific period of time and under specific conditions.
Deliverables	The data, actually the bits, that represents pictures, sound and text. Types of deliverables include (broadcast) events, video, audio and data of various kinds, including captions, graphics still images, text enhancements and other data as needed by an application.
Rolling buffer	A Rolling Buffer is an on demand accessible recording of a service for an encoding format allowing access to broadcast from the past without event boundaries.

Major kinds of TV-Anytimem metadata

- **Program** metadata:
    - describes information about single programmes, such as the title, genre, etc.
    - A programme is defined to be an editorially coherent piece of content which is typically acquired as a whole.
    - The programme is referenced via a programme CRID ("leaf CRID")
        - e.g. a CRID that resolves to a single programme.
    - The same programme may be found in any number of locations (e.g. schedule events), as is defined by the location resolution process.
        - This relationship is indicated via the one-to-many relationship link from "Programme" to "**Program Location**".
- Programmes can be grouped into "**Program Group**" elements such that a group may contain any number of programmes and a programme can be a member of any number of groups (many-to-many relationship).
- Furthermore, programme groups themselves can be part of other programme groups.
- Figure 5 shows the relationship between major kinds of TV-Anytime metadata.
- A programme group is uniquely identified by a group CRID[18].

❖ A programme can also be part of one or more aggregated programmes and aggregated programmes contain one or more than one programme indicating many-to-many relationship.

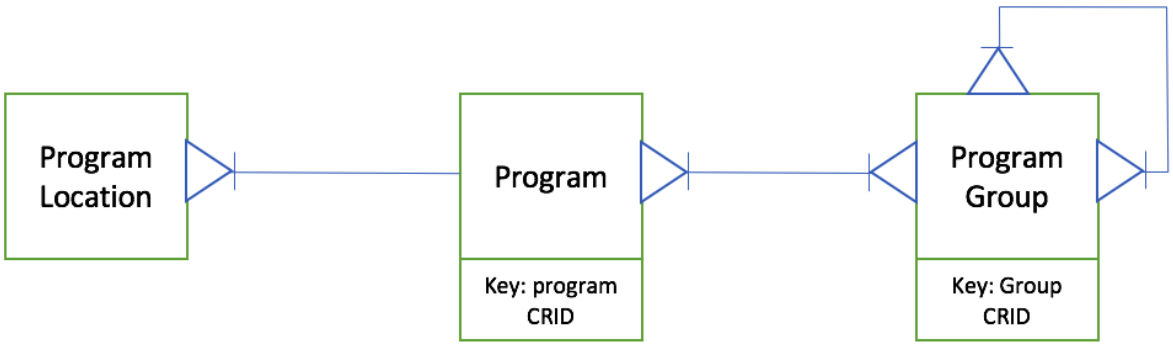


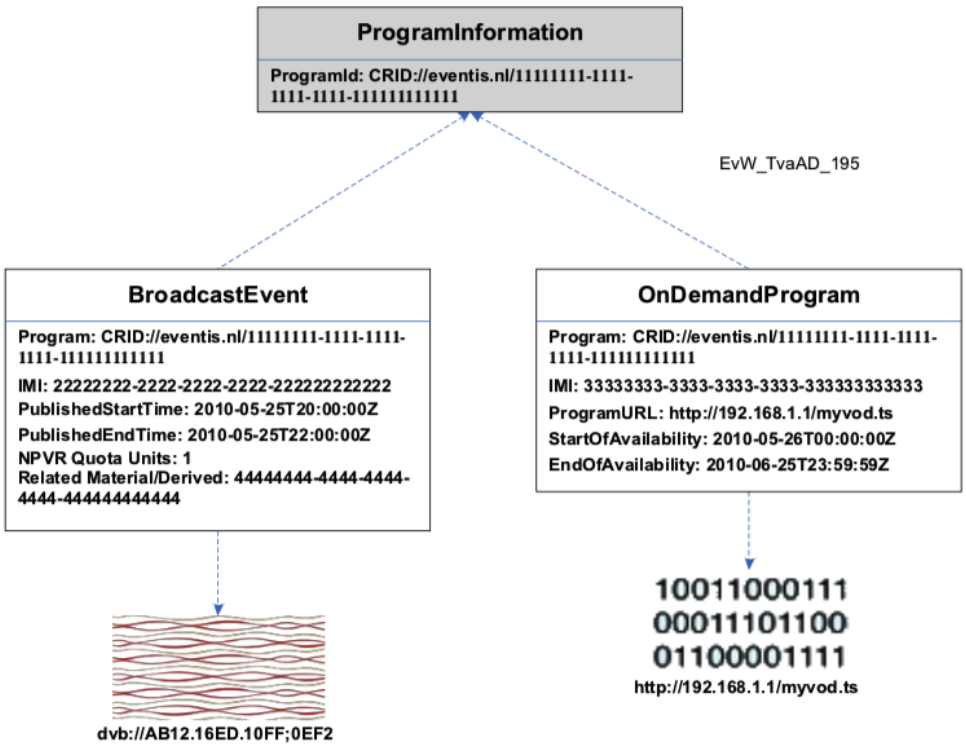
Figure 5: TV-Anytime content description model

[18] ETSI TS 102 822-4: "Broadcast and On-line Services: Search, select, and rightful use of content on personal storage systems ("TV-Anytime"); Part 4: Phase 1 - Content Referencing".

Content description model

- Requirement: the content description model must be able to represent the following concepts:
    - A simple programme.
    - A programme with a number of different versions
        - e.g. edits for sex, violence, language, director's cut, etc.
    - A programme that has been divided into a number of parts for publication
        - e.g. a 3hr film shown in 2 parts on different days
    - A programme that is a concatenation of a sequence of other programmes identified as an **aggregated programme**.
    - A **series of programmes** that can be ordered
        - e.g. episodes in a numerical order or unordered and bounded or unbounded.
    - A **collection of series** and individual programmes that have the same programme concept
        - i.e. a **show** (e.g. all series of "Only Fools and Horses" together with the Christmas specials).
    - A publication of a programme that may have publication dependent attributes
        - e.g. a film showing as tribute to a recently deceased actor which would have a different description.

Example:



TV-Anytime metadata schema specification

- Example of simple and complex utility types syntax used in TV-Anytime schema:

```
// A simple type used to indicate uniqueness within a
metadata
<simpleType name="TVAIDType">
   <restriction base="string">
   </restriction>
</simpleType>

//A complex type that allows a reference to be made to a
CRID.
<complexType name="CRIDRefType">
   <attribute name="crid" type="tva:CRIDType"
use="required"/>
</complexType>

<complexType name="FlagType">
   <attribute name="value" type="boolean" use="required"/>
</complexType>

<complexType name="TVATimeType">
   <sequence>
// to designate a point in time
     <element name="TimePoint"
type="mpeg7:timePointType"/>
// to designate a period in time
     <element name="Duration" type="mpeg7:durationType"
minOccurs="0"/>
   </sequence>
</complexType>

//A type defining the price proposed for a particular
content item. The same price can be expressed e.g. several
times in different currencies.
<complexType name="PriceType">
   <simpleContent>
      <extension base="float">
        <attribute name="currency"
type="tva:currencyCodeType" use="required"/>
      </extension>
   </simpleContent>
</complextype>
```

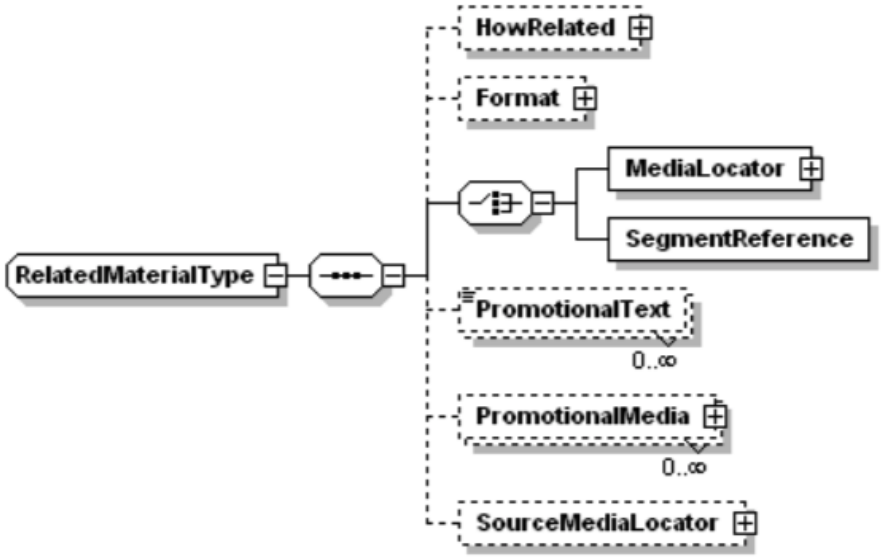
Example: GenreType syntax

```
<complexType name="GenreType">
   <complexContent>
     <extension base="tva:ControlledTermType">
        <attribute name="type" use="optional"
default="main">
          <simpleType>
             <restriction base="string">
                <enumeration value="main"/>
                <enumeration value="secondary"/>
                <enumeration value="other"/>
             </restriction>
          </simpleType>
        </attribute>
        <attribute name="metadataOriginIDRef"
type="tva:TVAIDRefType" use="optional"/>
     </extension>
   </complexContent>
</complexType>

<simpleType name="SynopsisLengthType">
   <restriction base="string">
     <enumeration value="short"/>
     <enumeration value="medium"/>
     <enumeration value="long"/>
   </restriction>
</simpleType>
```

Example: RelatedMaterial Type syntax

- Related materials can be used to refer a piece of content such as a movie to its trailor(s)

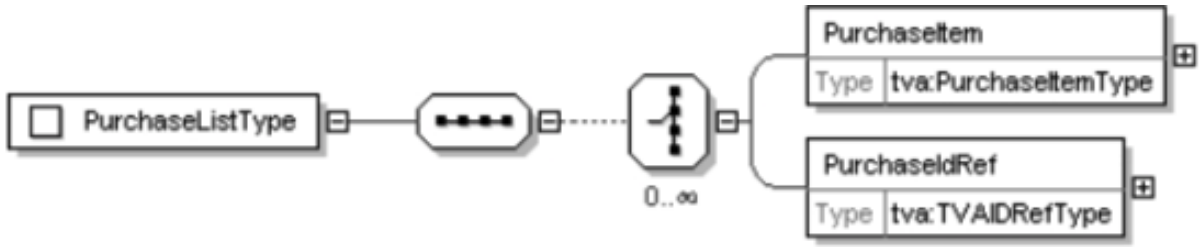



```
<complexType name="RelatedMaterialType">
   <sequence>
// Specifies the nature of the relationship between the
described AV content and the related media assets
     <element name="HowRelated" type="tva:ControlledTermType"
minOccurs="0"/>
     <element name="Format" type="tva:ControlledTermType"
minOccurs="0"/>
     <choice>
        <element name="MediaLocator"
type="mpeg7:MediaLocatorType"/>
        <element name="SegmentReference"
type="tva:SegmentReferenceType"/>
     </choice>
//Provides promotional information about the link, which can be
used as an additional attractor (e.g. record "Pride and
Prejudice" series).
     <element name="PromotionalText" type="mpeg7:TextualType"
     minOccurs="0" maxOccurs="unbounded"/>
<element name="PromotionalMedia" type="mpeg7:TitleMediaType"
minOccurs="0" maxOccurs="unbounded"/>
<element name="SourceMediaLocator" type="mpeg7:MediaLocatorType"
     minOccurs="0"/>
   </sequence>
</complexType>
```

Example: PurchaseItem (PI) Type syntax


```
<complexType name="PurchaseItemType">
   <sequence>
      <element name="Price" type="tva:PriceType"
maxOccurs="unbounded"/>
      <element name="Purchase" minOccurs="0"
type="tva:PurchaseType"
      maxOccurs="unbounded"/>
      <element name="Description" type="mpeg7:TextualType"
minOccurs="0"
      maxOccurs="unbounded"/>
      <element name="PricingServerURL" type="anyURI"
minOccurs="0"
      maxOccurs="unbounded"/>
      <element name="DRMDeclaration"
type="tva:DRMDeclarationType" minOccurs="0"/>
   </sequence>
   <attribute name="start" type="dateTime" use="optional"/>
   <attribute name="end" type="dateTime" use="optional"/>
</complexType>
```

Example: PurchaseList (PL) Type




```
<complexType name="PurchaseListType">
   <sequence>
      <choice minOccurs="0" maxOccurs="unbounded">
        <element name="PurchaseItem"
type="tva:PurchaseItemType"/>
        <element name="PurchaseIdRef" type="tva:TVAIDRefType"/>
      </choice>
   </sequence>
</complexType>
```

Example: TVA XML snippets

1. an example of the PI elements for buying an opportunity.

```
<PurchaseItem start="2010-06-30T00:00:00Z" end="2010-07-
10T23:59:59Z"> <Price currency="EUR">1.99</Price>
<Purchase>

     <PurchaseType
     href="urn:eventis:metadata:cs:PurchaseTypeCS:2010:transac
     tional"> <Name>Transactional offer</Name>

     </PurchaseType>
</Purchase>
<Purchase>

     <PurchaseType
     href="urn:eventis:metadata:cs:PurchaseTypeCS:2010:playFor
     Period"> <Name>Play for period</Name>

     </PurchaseType>
     <QuantityUnit
     href="urn:tva:metadata:cs:UnitTypeCS:2007:minutes">

          <Name>Minutes</Name>
        </QuantityUnit>
        <QuantityRange max="1440" />

</Purchase>

<Description>This is an example of a buying
opportunity.</Description>

</PurchaseItem>
```

2. an example of how the encoding profiles are represented within TVA XML Product link example for Schedule Event, ServiceInformation and Group – Note: the encoding profile (list) is part of the purchase list

```
<PurchaseItem>
<Price currency="EUR">0</Price> <Purchase>

<PurchaseType
href="urn:eventis:metadata:cs:PropertyCS:2010:profile">
<Definition>Cable_SD</Definition>

      </PurchaseType>
  </Purchase>

<Purchase>
<PurchaseType
href="urn:eventis:metadata:cs:PropertyCS:2010:profile">

<Definition>Cable_HD</Definition> </PurchaseType>

    </Purchase>
</PurchaseItem>
```

3. The OnDemandProgram, Rolling Buffer, LiveStream example when the encoding profile is part of the Genre element:

```
<Genre href="urn:eventis:metadata:cs:PropertyCS:2010:profile"
type="other”>

<Definition>IpadMediaFormat</Definition> </Genre>
```