Cross-Referencing of DAB-EPG and RSS Metadata

An Approach for a Broadcast-Podcast Hybridisation

Christian Rotzoll
Bauhaus-University Weimar
Faculty of Media
Weimar, Germany
rotzoll@medien.uni-weimar.de

Günther Schatter
Bauhaus-University Weimar
Faculty of Media
Weimar, Germany
schatter@medien.uni-weimar.de

Hagen Ch. Tönnies
Bauhaus-University Weimar
Faculty of Media
Weimar, Germany
toennies@medien.uni-weimar.de

Abstract
In this paper we describe the idea of cross-referencing between the different metadata formats RSS (for Internet distributed media files, e.g. podcast) and EPG (for digital broadcast, e.g. DAB). The goal is to provide the foundation for a seamless integration of both worlds in one generalised interface.

Keywords
DAB, Digital Radio, RSS, EPG, metadata, podcasting

INTRODUCTION

The media system surrounding us keeps getting more diverse, powerful, intelligent, and personal. Especially the patterns of using the WWW as active navigation through seemingly unlimited information spaces, comfortable selection, instantaneous access and simple storing, are also influencing all the other electronic media deeply. Young people with their rapidly changing expectations seek to use alternative ways to access entertainment and news while they are abandoning traditional sources. They tend to move away from radio and TV if they cannot find new possibilities of acquiring and processing. This shift from broadcast to a multi-channel digital broadband and interactive environment does change traditional broadcast media seriously. The future of the broadcasting system is threatened by tight programme schedules and the lack of interaction, navigation and storage tools.

Since 2004 the emerging trend of podcasting – downloadable media content via RSS 2.0 [1] distribution over the Internet – has demonstrated the request for highly personalised and on-demand media on desktop PCs and mobile media players. Today not only major media companies are offering podcast content for time-shifted distribution of selected media formats. Additional thousands of user generated podcast channels are providing a broad spectrum of content. Podcasting augments the importance of niche entertainment as well thematic special content with a strong and unique style. Highly specific kinds of music and news are becoming more important than traditional generalised radio content.

Due to the availability of push- and pull technology on one device and the possibility of shifted media consumption user habits changed remarkable. Hence, a demand arose for an integration of broadband and broadcast technologies to make it available in one user interface, see Figure 1.

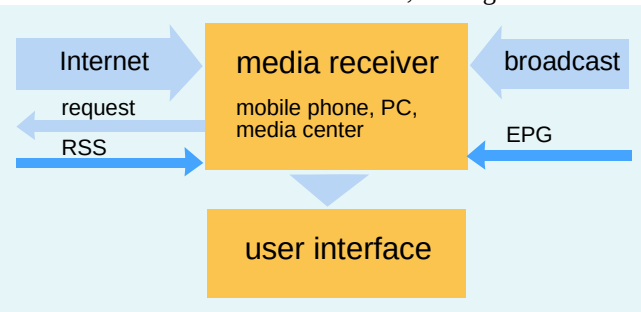

Figure 1. Integration of receiver concepts

The two main objectives of this paper are: At first we introduce the concept of cross-referencing metadata between RSS 2.0 distributed media and digital broadcasting EPG to combine the advantages of both distribution methods and to enable a seamless interface integration. Secondly we propose a hybrid transport model in order to avoid load peaks while delivering podcast content.

FOUNDATION

New Mobile Technology and Shifted Media Habits

The demand for an improved integration of RSS 2.0 distributed media and digital broadcast is based on two assumptions.

Broadband Internet and Digital Broadcast Access Combined

Considering the prior mentioned changes in the media system, the radio of the future will allow the access to each item ever broadcast via searching or recording facilities. This kind of an audible interactive medium will become an audio manager with a search engine finding every content at the very moment of a specific interest. The time-related audio access may be expanded from

- current programme (present),
- back to the last minutes (perfect) by caching,
- back by access to archives (past) via an Electronic Content Guide ECG or via WWW access and
- to expected programmes (future) by an Electronic Programme Guide EPG.

The major features of such a flexible time-dependent

medium are presented in Figure 2:

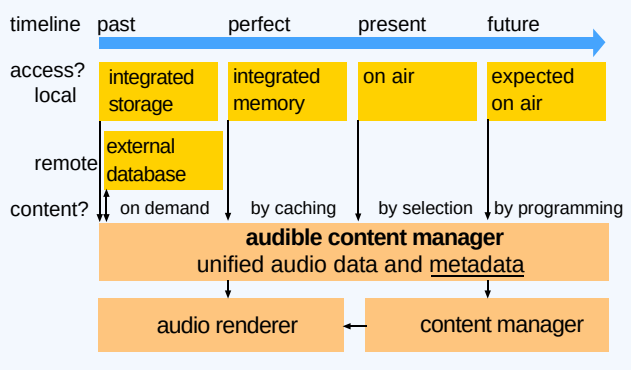

Figure 2. Types of content-access and timeline

This Audible Content Manager will receive the signals from a large number of external sources and contributors; they will convert to multi-format and multi-channel systems. Broadband Internet connection and digital broadcast technologies are expected to be more and more available on the same end user device. This trend is especially significant on the mobile phone market, where technologies like UMTS and DMB are being introduced. Furthermore, increased memory capacity is also enabling the mobile storing and archiving of media files.

Enhanced radios will be handheld devices merged with mobile phones, sophisticated car receivers or integrated in home media centres. Radio stations will have the excellent opportunity to subdivide and multiply their content into several channels for delivering – this trend is already visible today. Still, most of those channels and devices are separated. Therefore the goal arises to establish a unified method of media access – independent from time, transport channel and situation. That implies the precondition to coordinate and align the methods of content addressing via metadata.

Characterisation of Transport Systems

Digital broadcasting systems as DAB, DRM, and DVB are suitable media for delivering wide-interest content (audio, video, data services) in a cheaper way than other mobile services as UMTS or GSM. This is a prudential way, since one transmitter is able to serve an unlimited number of users and the frequency allocation is fixed. Even if the mobile Internet connection like UMTS is not fast enough for quality streaming, podcast distribution is based on media file download before the content is presented to the user.

As long as the data plan of the users mobile phone can handle huge file downloads (e.g. flat rate) the Internet connection enables RSS 2.0 distributed content. Furthermore the Internet can deliver streams and data, but without a performance guarantee in the case of a very large amount of receivers. Multicasting on the Internet is especially a great abuse of available resources [2].

Introduction into Metadata

The concept of cross-referencing is based on the metadata format RSS 2.0 for podcasting and the EPG (Electronic Programme Guide) standard for DAB (Digital Audio Broadcast).

Podcasting – RSS 2.0 Media Distribution

A podcast is a distribution method for audio and video media files via the Internet based on metadata. It is a RSS 2.0 [1] document (also called feed) extended from XML [3] that can be accessed by an URL [4]. This document (channel) contains a description of the channel itself and a list of items (podcast episodes) with further meta data (title, author, etc.).

A podcast relevant item provides also a special XML tag called ‚enclosure'. It is carrying a further URL that points to a downloadable media file. Mostly the Internet protocol HTTP [5] is used and the media file is an encoded audio or video (e.g. MP3, MPEG4, etc).

The most important qualities of RSS are as follows:

- *On-demand media:* As long as a media item is listed in the RSS document and the media file is online the content provided by this method is available to the user and can be downloaded by request. It can be used for a long time as archived media content.
- *Subscription based distribution:* A user can subscribe to a RSS document and receive new content as soon it is available.
- *Diversity of content offers a high potential of personalisation:* To provide a RSS channel with media content is almost possible for everybody on the Internet. A lot of free services are available for podcast publication. This leads to a broad diversity of content, which allows the user a very strong personalisation of his media consumption.

Receiving Podcast RSS 2.0 Distributed Media Content

To receive media content that is described and referenced by a podcast RSS document special content aggregating programmes or services called podcatchers are being used.

Figure 3. Podcast aggregation process

The user subscribes once to a podcast RSS channel. Scheduled or by user request all subscribed RSS documents are getting reviewed for changes and if new item episodes are available they are offered and/or downloaded to the users receiving device, see Figure 3. This can be a PC, a special media playback device as a portable audio player or a mobile phone. We developed two podcast solutions in April 2005 [6].

Podblogger – a Community Oriented Podcast CMS

The first solution is called "Podblogger", see Figure 4. It is a content management system programmed in PHP and specialised on community based podcast publishing. Every listener of the University podcast is able to submit her/his own audio content. The site administrator will organise the submitted content and the system takes care of providing RSS 2.0 metadata for podcast distribution for web publishing. The system is provided as an open source project [7].

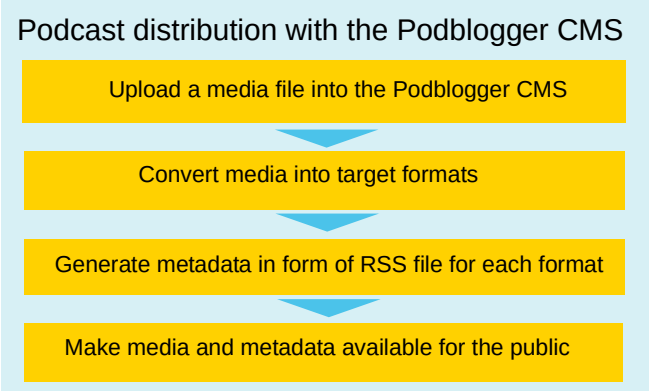

Figure 4. Content distribution with Podblogger

Podcatcher On A Stick (POAS) – a Device Centered Podcast Receiver

This podcast receiving application can be stored and is able to run directly from a storage device like for example an USB-memory stick or on most portable audio players, see Figure 5. It the application just needs an Internet enabled computer and was released in May 2005 [6].

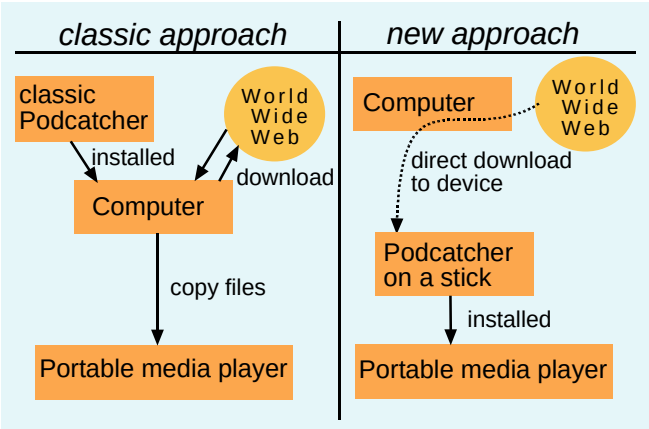

Figure 5. Podcast aggregation with POAS

EPG – Electronic Programme Guide for DAB

An EPG is an XML based document format [8]. It was defined to provide the listener additional and related content information to the broadcasted services like radio and television and is also able to refer to content and information on the WWW. It is specialised for Digital Audio Broadcasting DAB and Digital Radio Mondiale DRM [8][11][12].

The EPG document contains programme elements which are providing the description of a single broadcast transmission. To refer to information on the WWW the XML schema of EPG enables to use a special reference type which is called `urlType`. This type is used in the `programmeType` definition and there it takes place in the `linkType` [8]. This reference normally points to a HTML Website or contains an email address. To refer an EPG programme element the following identifier can be used.

The Content Reference Identifier (CRID): The CRID [13] is an identifier which is able to reference a special programme within one DAB service. The CRID is like a web URL specialised for broadcast use and defined as `cridType` in EPG XML schema. It is used within the programme element as value for the `id` attribute [8].

Comparing RSS und EPG

The RSS and EPG XML formats developed independently; however there are similarities in the underlying concepts. But there are also serious differences in the opulence of details and structure resp. capabilities of both formats, see Table 1.

Table 1. Comparison of EPG and RSS

	EPG [8]	**RSS 2.0 [1]**
Specialisation	for DAB/DRM broadcast services	for Internet applications
Standardisation	by European Telecommunications Standards Institute	user community defined, Version 2.0 is no official standard by W3C
Channel description	Does not contain direct channel description. That information is provided by service information document. Only groups can contain some descriptive content.	Contains description of channel: `title, description, link, category, webmaster, image, pubdate, lastBuildDate, ttl, copyright, generator, etc.`
Item description	`scheduleNameGroup, mediaDescription, genre, keywords, memberOf, link`	`title, link, description, author, category, comments, enclosure, guid, pubDate, source`
Complexity	complex grouping of programmes possible	simple list of items
Expandability	theoretical by XML namespaces, but limited due to backward compatibility	by XML namespaces
Availability	short time content	long time content

CROSS-REFERENCING OF EPG AND RSS METADATA

After giving a brief overview of both metadata formats we would like to introduce the concept of cross-referencing between both standards.

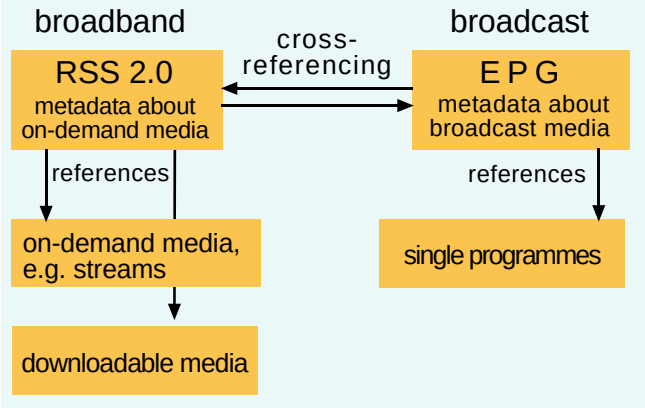

Figure 6. Cross-referencing between metadata

Cross-reference from EPG to RSS

Cross-referencing from EPG to RSS is making possible to point from a broadcast programme within an EPG to a related RSS Internet channel or channel item, see Figure 6.

Points to a RSS 2.0 channel in general:
`http://[Address of RSS document]`

Points to a single item episode within the RSS 2.0 channel:
`http://[Address of RSS document]#[GUID of RSS item]`

To add such a reference to a programme within the EPG the element `link` is used (child of element `programmeGroup`). Currently the regular use of this element is to provide the URL to a programme related website. A processing application can differ between a website and a RSS channel by the parameter `mimeValue`. For a RSS this value is always `application/rss+xml`.

```
<link
url=http://www.SomeBroadCaster.de/showname/archive.xml
mimeValue="application/rss+xml" xml:lang="de"
description="Subscription for the archive content"
expiryTime="future"/>
```

Even if a processing application is not offering special support for RSS channels the `url` can be displayed like any other website, if the RSS document contains a XML style sheet for web browsers.

Cross-reference from RSS to EPG

To refer from a RSS document on the Internet to a broadcast programme element within an EPG at first the processing application needs to identify the broadcast service that offers the requested metadata. For DAB this information is provided by the following identifier.

The Service Identifier (SID): A 16-bit, 24-bit or 32-bit code is used to identify a particular service. In the specification [9] the SID is an instance of the `contentIDType`. This type definition is used in the service information document which describes the available services that a DAB transmitter can deliver.

After the right broadcast service and the provided EPG metadata is available, the CRID can be used to identify the programme element that is the final target of the reference. To integrate such a cross-reference into a RSS 2.0 document we created a XML schema definition [10] which can be used as an additional namespace.

The main element in this schema is the `link` element. This element requires a parameter called `ref` which needs to contain an URI [14]. For a cross-reference to an EPG programme element this parameter contains the CRID. The optional parameter named `source` (requiring also an URI) is able to carry the additional SID information to provide a well-defined RSS-EPG cross-reference like in the following example.

```
<crossref:link
ref="crid://example.co.de/%E3%82%A8%E3%82%A4%E3%82%AC"
source="DAB:e1.cce15.cc221.0.1" />
```

The CRID uses its URI format as described in the EPG standard [8]. To provide more information in which kind of broadcast service the CRID is a valid reference the `source` parameter needs the prefix “DAB” (as scheme of URI) followed by the SID (as scheme-specific part of URI). This way all technical information needed are available to refer to an EPG programme element.

The `link` element can be used as an empty/single XML element like in the example above. But to add more useful information for the processing application the following elements can be child of a `link` element.

```
<crossref:available from="2007-03-04T20:15:00-01:00"
until="2007-03-04T21:00:00-01:00" />
```

The `available` element provides information about the time window in which the reference is available. This element can be used multiple times if needed. The time and date formats are based on the XML schema data type `xs:dateTime` [15].

```
<crossref:purpose type="hybridDistribution" />
```

The `purpose` element contains a parameter called `type`. This string acts as an ID describing the kind of use for this reference. This information should help the processing application to differ between certain kinds of cross-references and to present it in the right way to the user. For referencing from a RSS document to an EPG programme element as a kind of hybrid distribution (see below) it should always have the value `hybridDistribution`.

The `purpose` element can also be used as a child of the RSS element `channel` to describe the kind of use for the actual document (for example `archiveRSS` or `extendedMaterial`). This can help to determine which service is offered to the user by this RSS document when it is cross-referenced by another document like an EPG.

BROADCAST-PODCAST HYBRIDISATION

Referencing from EPG to RSS is adding on-demand and diversity to a broadcast channel through alternative or additional podcast delivery. To determine why referencing from RSS to EPG is reasonable a closer look at the derivation and performance of RSS 2.0 media distribution is needed.

Optimising Performance of RSS Distribution

Distributing downloadable media files via a RSS 2.0 document was designed for small community environments. It is a classic narrow casting technique to reach a limited audience. If a channel is used in a broadcast-like manner the following effect occurs.

Every user of an RSS channel is requesting a unique download of the media content. So the web server traffic volume will increase rapidly with a growing number of users. Due to the RSS subscription mechanism most of those requests are happening in a short timeframe close to the release date of a new channel episode, see Figure 7. The traffic peak can become critical if the number of users is very high.

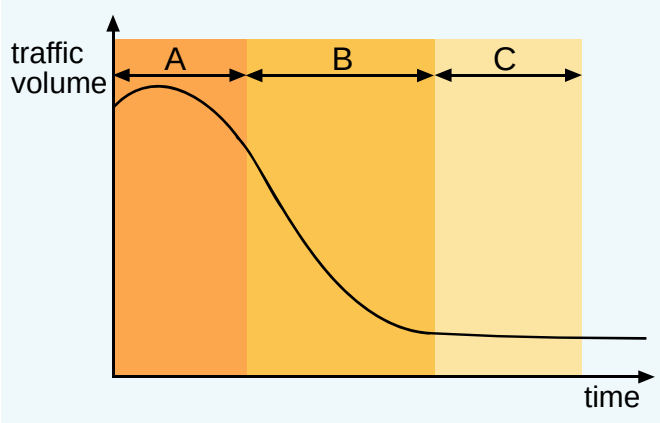

Figure 7. Traffic peak on release of RSS media content

Broadcast media seems to be the perfect fit for this peak time. Instead of having a linear user-traffic dependency, the data volume to distribute one episode stays constant with a growing number of listeners.

So if the same content is provided via broadcast and RSS distribution one partial solution to mitigate this peak is to provide the content first by broadcast media and having a cross-reference (from RSS item to EPG programme) in the RSS document before the actual broadcasting date. This way all podcatching applications can prioritise the broadcast recording of content instead of downloading it. After the broadcast date the normal RSS enclosure file reference is added to make the content available for users who were not able to receive it.

Hybrid Podcast-Broadcast User Interface

Future user interfaces need to integrate classic and enhanced broadcast services like live consumption, programming and automated recording as well as managing and downloading podcast content under one surface, see Figure 8.

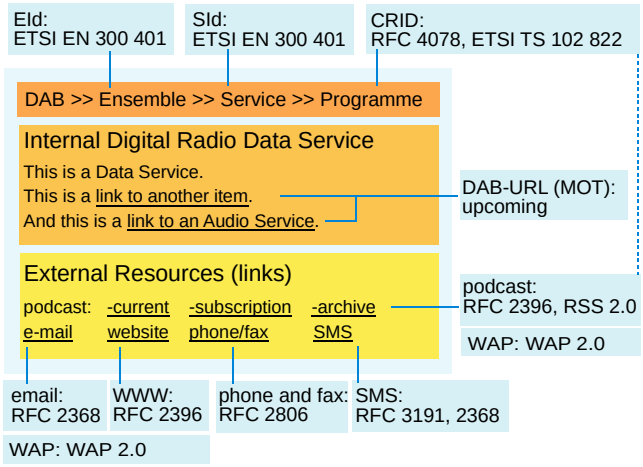

Figure 8. Addressing of resources for an Audible Content Manager

Instead of presenting both distribution methods only side by side, through cross-referencing a user interface can detect connections between both kinds of media and offer enhanced options and interactions that are giving the user more comfort and a seamless feeling. Cross-referencing is adding the following advantages to user interface:

Access to on-demand programme archive: If a user is watching a live or recorded broadcast programme and for example a reference to an „archive RSS“ is available (see element `purpose` in XML schema [10]), then the user interface can offer the option to browse the archive and select single episodes for on-demand access. An „archive RSS“ can also provide metadata including programmes that are not longer provided in the EPG when the user is searching on all broadcast channels for special content.

Backup access for Personal Media Recording: Also the on-demand RSS access can improve the supply of recorded broadcast content. Especially mobile devices cannot guarantee „always on“ and „always a clean signal“ for broadcast channels. So if the user has subscribed on broadcast programme episodes via EPG (the idea of Personal Media Recording) and the broadcast signal is not available or distorted, a referenced RSS can provide the missing content. Such a backup solution can operate fully automatically without bothering the user.

Supply of extended media content: If a user is watching a live or recorded broadcast programme episode and a reference to an „extended material RSS“ is available (see element `purpose` in XML schema [10]), while playback or at the end of the episode the user is getting offered the download or the subscription of those referenced content. For example the user is consuming an informational magazine dedicated to a special topic.

Figure 9. Possible user interface design

After the show the user can subscribe to a podcast on that topic and will be provided with additional media as it gets available, see Figure 9.

CONCLUSIONS AND FUTURE PROSPECTS

Backward Compatibility for RSS and Broadcast

Even if RSS applications or broadcast receivers are not able to handle the cross-reference information, all produced RSS or EPG data are still valid and can be processed the usual way. For example if a broadcaster is producing RSS documents to enrich his broadcast content such documents can be used by almost all RSS applications that are already in use. The produced content can be used in multiple ways.

Open Standard – Future Developments

The presented cross-reference XML schema is just a first approach to show the possibilities of cross-referencing between Internet and broadcast distributed media. The schema is open for improvement and extension.

The given examples are specifically designed to show the cross-reference between RSS and DAB. The idea can be extended to all other digital broadcast formats like DVB or DMB.

A look at the `link` element shows that the optional `source` parameter is expecting an URI. For DAB this is always 'DAB:[ServiceID]'. In the same way it should be possible to identify also other broadcast services.

Also the `purpose` element offers a free defined ID string. It is possible to define further ideas of application for a cross-reference (for example refer between different RSS documents).

At the moment the cross-referencing as shown in this document is a very simple and practical approach to refer between metadata. But it should help to test and discuss the development of a unified method of media access – independent from time, transport channel and situation.

Opening up Broadcast for External Media

The cross-referencing is opening up broadcast content for external media on the Internet in a seamless way. Thousands of already produced Internet channels are available. This offers a lot of new possibilities for broadcast publishers to enrich their content and even integrate and moderate user generated content into existing infrastructures. Users can submit own generated content (e.g. audio fan reports, video commentary, etc.) and by using cross-referencing it can be merged into the existing broadcast services.

Cross-referencing is one practical approach to open a window to Internet media resources like podcasts and combine them with the advantages of digital broadcast. The ability to refer to content independent of its location and time will improve the capability of the media system with a listener-focused noticeable progress.

REFERENCES

[1] Winer, D., RSS 2.0 Specification, `blogs.law.harvard.edu/tech/rss`

[2] Barletta, A., Java–applications in Digital Audio Broadcasting, EBU Technical Review September 2001.

[3] Yergeau, F. et al., XML 1.1, W3C Recommendation, `www.w3.org/TR/2004/REC-xml11-20040204/`

[4] RFC 1738 Uniform Resource Locators (URL)

[5] RFC 2616 Hypertext Transfer Protocol (HTTP/1.1)

[6] Rotzoll, C., Tönnies, H., Podcast[@]University. Building an Infrastructure for Instant Audio Content. Bauhaus-Universität Weimar 2005.

[7] Podblogger Project Website, `www.podblogger.de`

[8] ETSI TS 102 818 Digital Audio Broadcasting (DAB); XML Specification for DAB Electronic Programme Guide (EPG).

[9] ETSI TS 102 371 Digital Audio Broadcasting (DAB); Digital Radio Mondiale (DRM); Transportation and Binary Encoding Specification for Electronic Programme Guide (EPG)

[10] Cross-Referencing Project Webpage providing XML schema definition, `www.crossreferencing.org/schema/crossReference-0.1.xsd`

[11] ETSI ES 300 401 Radio Broadcasting Systems; Digital Audio Broadcasting (DAB).

[12] ETSI ES 201 980 Digital Radio Mondiale (DRM), System specification.

[13] RFC 4078 The TV-Anytime Content Reference Identifier (CRID)

[14] RFC 2396 Uniform Resource Identifiers (URI), Generic Syntax

[15] XML Schema Part 2: Datatypes Second Edition. `www.w3.org/TR/xmlschema-2/`