

Minnesota Public Radio
Metadata Model Overview

Public Broadcasting Metadata Model Project

Arlington, VA
2002-04-24

Alan Baker
Minnesota Public Radio
45 E 7th St.
Saint Paul MN, 55101
abaker@mpr.org

Digital Archive Initiative Background

Problem Statement

Minnesota Public Radio maintains an archive of its radio productions for both historical, mission-driven reasons as well as for purposes of research and reuse. These archived materials are recorded on a variety of media, stored in a variety of places, and indexed in a variety of ways. This diversity arises from embedded practices and from the lack of a centralized indexing and archiving procedure. Much of the success in retrieving material from the current archive depends on the researcher accessing the individual authors or knowledgeable producers on the MPR staff. As a result, there may be material in the archive that would be useful for a given feature, but which may be missed due to time constraints or inaccessibility of knowledgeable staff.

The ramification of the current MPR archive system is that the wealth of information in the Minnesota Public Radio audio archives is not efficiently exploited, to the detriment of the on-air product. Retrieving audio from the existing archive is a labor-intensive process that often takes more time and effort than staff can afford.

There are several obstacles that stand in the way of more fully utilizing this vast store of audio material:

- Most of the current Minnesota Public Radio Archive is not searchable by electronic means. To find an archived item a user must know the item's name and the approximate date it aired.
- With the upgrading of Minnesota Public Radio's production facilities to include digital production tools, it is harder to accommodate analog tape playback in all or even most of the production spaces. Therefore, loading the existing analog tape into the system for production can be difficult as well.
- In some circumstances tapes from the archive need significant repair before they are playable.

To make use of MPR's current analog/DAT/CD archive, a person must be in the MPR building and know how the archive is constructed. Only knowledgeable staff, who work at the St. Paul studios, can use the archive unassisted effectively. Public access is rare, owing to the impractical demands on staff to assist public users.

These constraints leave MPR unable to make its audio archive of more than 70,000 hours of unique and valuable audio available to much of the company, and rarely does the public get access to these materials.

Vision Statement

The media archive system will allow staff to access programs and scripts produced by Minnesota Public Radio. Users will access the media archive system from an audio workstation or a personal computer through a Web browser. From the Archive home page users will be able to search the archive index, audition audio via their browser, and read textual information such as scripts linked with each audio item. If logged on from an audio workstation computer, the user will be able to download the uncompressed audio to that workstation for further use in production or playback to air. If the search cites an item for retrieval that is stored offline (i.e. a show that was mastered to CD) the user will be given the item's location so that it can be retrieved from MPR's shelf archive.

One of the functional requirements of the MPR archive is that a special staff position not be needed to directly load items into the archive. Users will be responsible for preparing items for the archive and for sending them to the archive using the same tools that they now use to produce audio and to do other aspects of their job. After a production is completed, the producer will include the necessary descriptive information via a web browser interface and will send the item to the archive. Some situations that allow automating the "move to archive" procedure will require even less of the producer. MPR is aware that they may need to appoint staff to oversee the process and to guarantee the quality of information being added to the archive.

In the future, the MPR Archive may also serve as the repository for materials that may be requested as “audio on demand” by commercial or educational institutions or by the public at large. The business charter of Minnesota Public Radio allows for and encourages this dissemination of materials. This usage of the Archive will be considered in the design, but its implementation is outside the scope of the current project.

Phase One Overview

Phase one of the MPR Digital Archive project will create the infrastructure that will allow MPR to store and track all electronic and physical assets in a central database. We will begin storing audio as digital files with the goal of phasing out DAT tape and other unreliable physical media. The content that will be stored as digital audio files in the digital archive will initially be limited to daily news features. Once the core system is proven, we will begin transitioning newly produced content to digital storage. Likely second moves to digital storage are Midmorning, Midday. Music productions and other content currently mastered to CD will continue to be archived on CD in the near future.

Once the phase one system is up and running we will begin tracking all newly produced MPR content in the Digital Archive database. Users will access this information via an audio workstation and/or standard desktop computer using a Web browser. From the Archive home page users will be able to search the archive index, audition audio via their browser, and read textual information such as scripts linked with each audio item. If logged on from an audio workstation computer, the user will be able to download the uncompressed audio to that workstation for further use in production or playback to air. If the search cites an item for retrieval that is stored offline (i.e. a show that was mastered to CD) the user will be given the item’s location so that it can be retrieved from MPR’s shelf archive.

Undertaking the conversion of our current analog and digital content into the digital archive system will depend on our findings in our phase two project and funding.

The Metadata Scheme

Minnesota Public Radio has adopted the Dublin Core metadata standard for the description of radio programs in its media archive system. The standard has been adapted for the description of radio programs and scripts in physical and electronic formats through the introduction of element modifiers, the development of authority files, and the inclusion of data format standards.

The metadata scheme is presented below complete with element modifiers and encoding schemes, as well as information concerning the structure of element data in the media archive system database and the method of data entry.

The use of the metadata scheme to describe resources at Minnesota Public Radio is detailed in the Application Profile that follows. The authority files and information concerning their development and use are appended to the Application Profile.

Element	Modifier	Scheme	Definition
Title			The name given to the program.
	Series	Series Title Authority File	A title applied to a sequence of related programs
	Program		A title applied to an individual episode in a program series
	Alternate		A name by which a program is otherwise known. A variant title, production title, or translated title.
Creator			Any station, production unit, or organization responsible for creating the program.
	<i>Name</i>		The name of the member station, production unit, or organization.
	<i>Role</i>	Creator Role Authority File	The role of the organization in the creation of the resource.
Description			An account of the content of the program
	Abstract		A summary of the content of the program. A promo, slug, or lead.
	Table of Contents (Text)		A formatted list of the content of the program. A list of segment titles or features.
	Table of Contents (File)		A formatted list of the content of the program. A list of segment titles or features.
	<i>Composers and Works</i>	Composers and Works Authority File	A list of works and composers.
	<i>Rundown (Text)</i>		A formatted list of the content of a program with time code and other information.
	<i>Rundown (File)</i>		A formatted list of the content of a program with time code and other information.

	<i>Cue Sheet (Text)</i>		An account of the content of a program with instructions for broadcast, such as in-cues, out-cues, and fades.
	<i>Cue Sheet (File)</i>		An account of the content of a program with instructions for broadcast, such as in-cues, out-cues, and fades.
	<i>Playlist (Text)</i>		A list of composers, works and commercial releases or sources included in the program.
	<i>Playlist (File)</i>		A list of composers, works and commercial releases or sources included in the program.
Publisher			An organization or service responsible for the publication or distribution of the program.
	<i>Name</i>	Publisher Name Authority File	The name of the organization or service responsible for the publication or distribution of the program.
	<i>Role</i>	Publisher Role Authority File	The role of the organization or service responsible for the publication or distribution of the program, ie. publisher or distributor.
Contributor			Persons or organizations that contribute to the content or presentation of the program
	<i>Name</i>	Contributor Name Authority File	The name of the contributor.
	<i>Affiliation</i>		The related organization of the contributor.
	<i>Role</i>	Contributor Role Authority File	The primary role of the contributor.
Subject			The topic of the content of the program
	<i>Heading/Code</i>	IPTC Subject Reference System	A subject term and code selected from a controlled vocabulary.
	<i>Keyword</i>		A subject term or phrase assigned by the record creator.
Date			A date associated with an event in the life-cycle of the program
	Local Broadcast	W3C-DTF (yyyy-mm-dd)	Date on which the program was broadcast locally.
	Regional Broadcast	W3C-DTF (yyyy-mm-dd)	Date on which the program was broadcast regionally.
	National Distribution	W3C-DTF (yyyy-mm-dd)	Date on which the program was distributed nationally.
	Rebroadcast	W3C-DTF (yyyy-mm-dd)	Date on which the program was rebroadcast.
	Web Issue	W3C-DTF (yyyy-mm-dd)	Date on which the program was made available on the web.

	Created	W3C-DTF (yyyy-mm-dd)	Date on which the finished program was created, if different from local broadcast date.
	Performance	W3C-DTF (yyyy-mm-dd)	Date of a performance included in the program.
Type			The nature or genre of the content of the program. Type includes terms for describing categories, genres, or aggregation levels for programs.
	Resource	DCMI Type Vocabulary	The general nature of the resource, ie. sound, image, or text.
	Program	Program Type Authority File	The structure or format of the program, ie. magazine, call-in, or documentary.
Format			The physical or digital manifestation of the program.
	Title	Format Title Authority File	The title of the format.
	Medium	Format Medium Authority File	The physical carrier or file type of the program.
	Attach File		The local path name for the digital file prior to submission to the media archive system.
	File Size		The size of the file.
	Duration	(hh:mm:ss)	The duration of the program.
	Audio Channel Configuration	Stereo or Mono	The channel configuration of the audio file or program recording.
	Audio Resolution		The bit-depth of the audio file.
	Audio Sampling Frequency		The sampling frequency of the audio file.
	Physical Description		Other physical details of the format.
	Identifier		The identifier of the format.
	Barcode		The barcode identifier of the format.
	Location		The location of the format.
	Note		Note relating to the format.
Identifier			A unique identifier assigned to the program.
	System Identifier	Asset Id	This is the identifier of the asset within the MAS.
	Archive Number		The number of the program assigned by the archive.
	Program Number		The number of the program as assigned by the producer or creator
Source			A reference to a resource from which the present resource is derived.
Language		ISO 639-2	A language of the intellectual content of the program.

Relation			A reference to a related resource.
	Type	Relationship Type Authority File	The type of relation: has version, is version of, has part, is part of, references, is referenced by
	Identifier	URI	The id of the object being referenced.
	Note		A description of the object being referenced
Coverage			The extent or scope of the content of the program
	Spatial		
	Temporal	DCMI Period Authority File	
Rights			Information about the rights held in and over the program.
Notes			
	Recording		Any information concerning the recording of the program.
	Technical		
	Award		The name, issuing body and year of any award received by a program or series.
	Record Creation		Any information concerning the creation of a record.
	Record Revision		Any information concerning the revision of a record.
Reformatting Documentation			Information concerning the source recording and conversion into the digital archive.
	Source		
	Physical Description		
	Recording Characteristics		
	Condition		
	Treatment		
	Conversion		
Record Creation			Information concerning the creation of the metadata record.
	Name	UserName	Name of the record creator.
	Date	W3C-DTF	Date of the record creation.
Record Revision			Information concerning the revision of the metadata record.
	Name	UserName	Name of the record revisionist.
	Date	W3C-DTF	Date of the record revisionist.