
The PBCore metadata standard: A decade of evolution

Received (in revised form): 26th January, 2012



Nan Rubin

has more than thirty years' experience managing technology and facilities projects in public radio and television, including 15 years at New York public television station WNET-TV/Thirteen, working in technology planning. From 2003–2010, Ms. Rubin was Project Director of *Preserving Digital Public Television*, funded by the Library of Congress's National Digital Information and Infrastructure Preservation Program (NDIIPP), where she oversaw a team of 20 based at WNET, WGBH in Boston, the Public Broadcasting Service, and New York University designing a model preservation repository for born-digital public television programmes. She speaks frequently on digital preservation and media archives; and articles on preserving digital video have been published in *International Preservation News*, *Library Trends Journal*, and the *Technical Journal of the Association of Moving Image Archivists*. During the development of PBCore 2.0, she was responsible for outreach to users.

Community Media Services, 4700 Broadway #2J, New York, NY 10040, USA
Tel: +1 212-569-3391; E-mail: Nanrubin@gmail.com; website: www.nanrubin.com

Abstract In less than ten years, broadcasting has been completely transformed from an analogue, tape-based medium to a production and distribution environment that is digital from end to end. Early in the transition, the public broadcasting system recognised the need for a single, unified descriptive metadata schema able to deliver digital content across multiple platforms. The Public Broadcasting Metadata Dictionary Project (PBMD), comprising representatives from both public radio and public television, was convened by the Corporation for Public Broadcasting in 2001, and spent several years analysing metadata standards and collecting user requirements to create PBCore, based in large part on Dublin Core. PBCore v1.0 was released in 2005, and since then, it has been widely adopted not only by broadcasters, but also by media archives as the preferred scheme for descriptive and technical metadata. During several years of dormancy, PBCore was maintained by a loyal group of users, but with the planning of the new American Archive project, CPB is once again supporting PBCore and Version 2.0 was released early in 2011. Beyond public radio and television stations, the standard has been found to be popular for time-based media, and the community of user institutions has expanded rapidly to include audio, video and film archives. Although PBCore is proving to be very valuable in meeting its original goal as a standard for rich media that is extensible, scalable, flexible and easy to understand, it continues to need support for its ongoing maintenance and improvement and to expand its adoption.

KEYWORDS: PB Core, Dublin Core, digital media, descriptive metadata, taxonomy, media archives, broadcasting, asset management, video database; public television archives; public television metadata; public radio archives; public radio metadata.

FROM ANALOGUE TO DIGITAL: THE TRANSFORMATION OF MEDIA OBJECTS

In a public broadcasting system made up of

hundreds of independent licensees, the challenges of organising universal processes for asset appraisal, digitisation, rights clearance, preservation, etc. are myriad,

perhaps overwhelming. We did understand, however, that the foundation of any future effort in this direction would be a single, shared protocol for identifying and describing our rich media assets.

**PBCore Progress Report, presented at
Dublin Core Conference, October
20031**

When the above was written, it was not yet evident that television broadcast operations would be altered so rapidly and profoundly by digital technologies. In less than a decade, television production, distribution and preservation have undergone a radical shift, as programmes are now all shot, edited, and completed in digital formats. When all full-power television stations in the USA turned off their analogue transmitters in 2009 and began broadcasting exclusively on digital channels, the all-digital production and broadcast chain was complete.

The user expectation has changed as well, demanding that programmes be viewable on everything from the very smallest iPod screen and smartphones, to giant, wall-size, flat-panel displays, and everything in-between. Digital media files *must* be usable at every stage in the production process and across every one of these devices.

**Early recognition by public
broadcasting of a need for metadata
standards**

Without knowing how radical the changes would be, public broadcasters understood a decade ago that the emerging digital landscape would demand entirely new and different operating requirements.

To take the steps necessary to manage a digital environment, the Corporation for Public Broadcasting (CPB) convened the Asset Management Caucus in 2001, comprising broadcast engineers, technologists and traffic personnel from both radio and television stations, which quickly became the Public Broadcasting

Metadata Dictionary Project (PBMD). Their goal was to adopt a single set of metadata protocols, which would be known as the ‘Public Broadcasting Core (PBCore) Metadata Dictionary’.

The PBMD Project was a ‘cross-organisational, multi-disciplined effort to establish a metadata standard for digital public broadcasting content — both radio and television — in order that files might be more easily exchanged between software systems, institutions, and community partners’.¹

Given the constant introduction of new broadcast recording, distribution and storage equipment, there was an understanding from the beginning that a single metadata standard was necessary to ensure:

- interoperability of files *within* a station — allowing a single programme file to be used successfully by operating facilities such as network recorders; automation devices; editing systems; broadcast playout servers; and on-line interfaces (see Figure 1);
- exchanging programme files *between* entities — enabling programme files to be sent from stations to the public radio or television network; from the networks back to stations; and directly between stations.

And while not explicit, it was thought that the standard would also be important for storage, retrieval and playback of digital archival materials.

Identifying an appropriate model

Because common metadata is fundamental to the exchange of content among different users, the intent was to identify a core set of metadata descriptors applicable to most public radio and television uses that could facilitate exchange of digital programme files. To be used, the schema had to be easily understood,

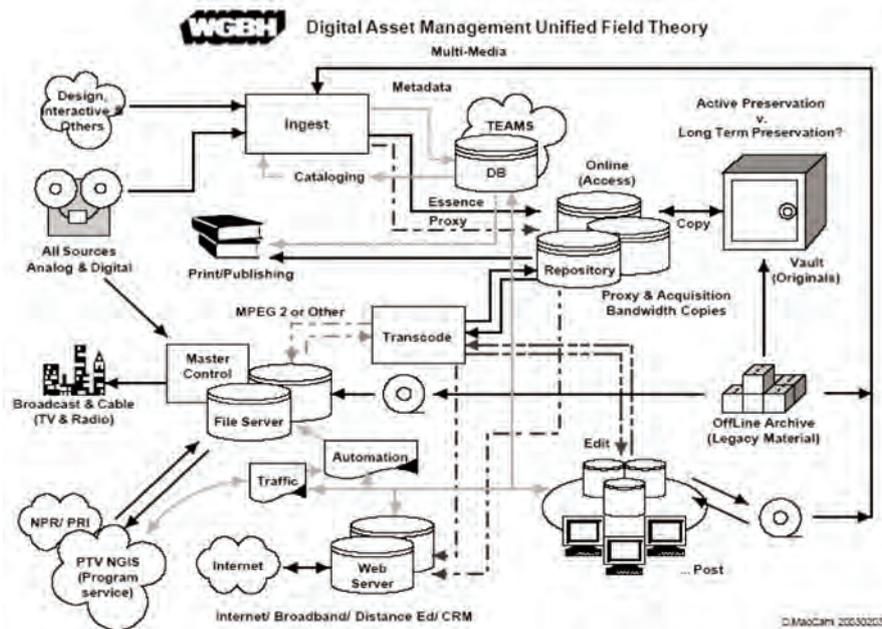


Figure 1: Internal broadcast station workflows (Dave MacCarn, WGBH, 2003)

implementable and acceptable by the public broadcasting community at large, both radio and television. The goal was to arrive at the smallest set of elements that could adequately describe and catalogue programme files, to facilitate use by internal station devices, and among such different users as stations, distributors, independent producers, and vendors who make traffic, digital asset management (DAM) and related systems.

The Project spent more than two years studying digital operating systems at stations and examining various metadata descriptors, dictionaries, and schemata, in particular metadata used to describe rich media and standards that deal with the subjects, administration and educational aspects of the assets.

They considered such standards as: OAIS, Reference Model for an Open Archival Information System;² SMEF-DM, Standard Media Exchange Framework;³ MPEG-7, 'Multimedia Content Description Interface';⁴ and MARC;⁵ used extensively by libraries.

The Dublin Core Element Set⁶ emerged as the most appropriate to meet these various concerns, and PBCore was devised as an application profile built on Dublin Core, which also retained elements from other schemata and from station-based public broadcasting digital asset management and traffic systems. The end result was a set of metadata standards with a solid foundation that is extensible, scalable, and easy to understand.

Elements of the Public Broadcasting Core

PBCore was intended to be 'simple', but not 'simplistic'. It provides a starting point that could accommodate metadata extensions for communities of specific users. PBCore is especially useful for describing digital media, including file URLs for streaming or download, and it can serve as a syndication format like RSS or ATOM but in much greater detail. Of importance for preservation, a complete PBCore record contains rich metadata on the provenance of the object, plus the

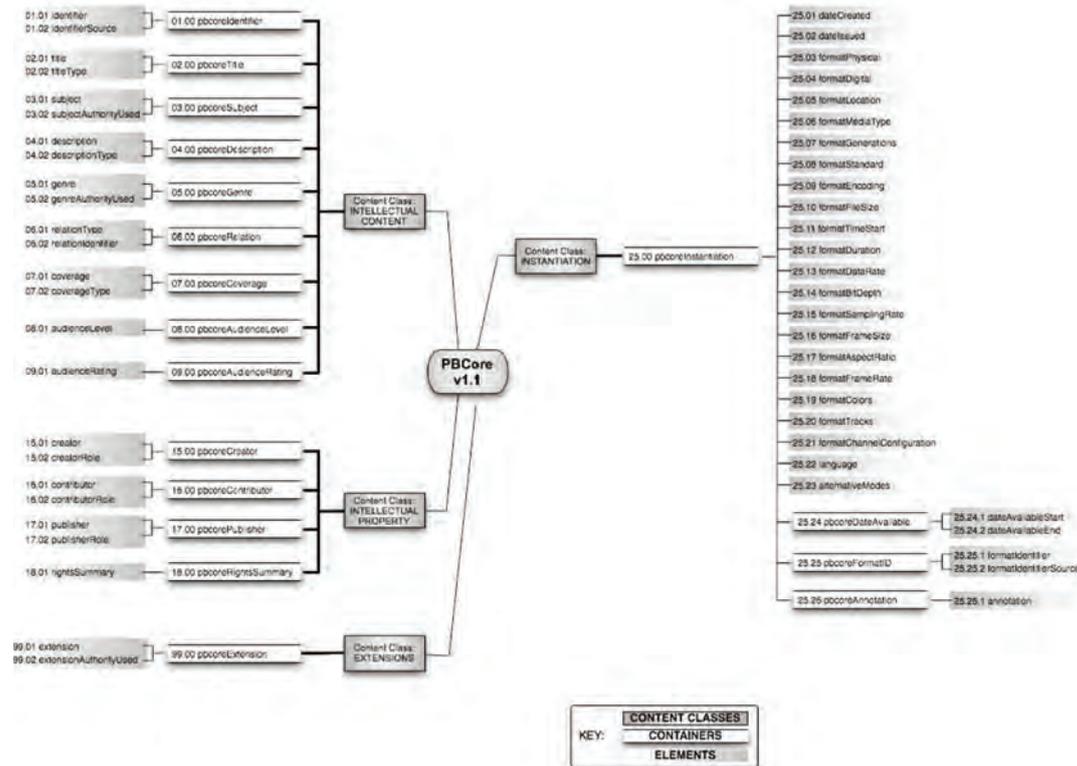


Figure 2: Schema: PBCore v1.1

location of media ‘instantiations’ whether digital or analogue — for example, a PBCore asset record can specify a media asset published on the internet, along with the location of the file and ownership of its source.

PBCore contains 58 elements placed in three categories:

- *Content*: 20 elements describe the actual intellectual content of a media asset or resource.
- *Intellectual Property*: nine elements relate to the creation, creators and usage of a media asset or resource.
- *Instantiation*: 29 elements identify the nature of the media asset as it exists in some form or format, in the physical world or digitally.

These, in turn, are organised by Content Class, each of which consists of Containers

holding individual Elements (Figure 2).

A fragment of PBCore XML shows the pattern of its schema: nested pairs of Elements inside Containers:

```

<pbcoreIdentifier>
  <identifier>Division of Broadcasting
  Sound Recordings, Record Series
  13/6/5, Tape 040</identifier>
  <identifierSource>University of Illinois
  Archives</identifierSource>
</pbcoreIdentifier>
<pbcoreTitle>
  <title>My Experience in the Labor
  and Radical Struggles of the Thirties,
  by A.J. Muste, February 9,
  1966</title>
  <titleType>Program</titleType>
</pbcoreTitle>
<pbcoreSubject>
  <subject>ACLU</subject>
  <subjectAuthorityUsed>Illinois Public
  
```

```

Media</subjectAuthorityUsed>
</pbcoreSubject>
<pbcoreDescription>
<description>On the campus of the
University of Illinois at
Urbana-Champaign, A.J. Muste, a
social justice activist, describes the
development of labor unions during
the Great Depression. Muste became
an avowed Marxist-Leninist, and he
was instrumental in forming the
American Workers Party. However,
Muste says he began to question the
ability of Marxism in carrying out
efforts to help workers. ‘I was
convinced that the concentration on
political maneuvering,’ he explains,
‘would virtually eliminate our
effectiveness in the mass struggle
taking place in basic
industries.’</description>
<descriptionType>Abstract</descriptionType>
</pbcoreDescription>
<pbcoreGenre>
<genre>Speech</genre>
<genreAuthorityUsed>Illinois Public
Media</genreAuthorityUsed>
</pbcoreGenre>
<pbcoreCoverage>
<coverage>1930s</coverage>
<coverageType>Temporal</coverageType>
</pbcoreCoverage>
<pbcoreCoverage>
<coverage>United States of
America</coverage>
<coverageType>Spatial</coverageType>
</pbcoreCoverage>
<pbcoreAudienceLevel>
<audienceLevel>General</audienceLevel>
</pbcoreAudienceLevel>

```

Container pairs for subject, description and other elements are repeatable, allowing for extensive annotation and the inclusion of multiple domain-specific taxonomies and descriptive formats within a single PBCore record. ‘Instantiations’ in PBCore represent the physical or digital versions

that exist for a given title or asset. A complete PBCore record, then, contains a description of the intellectual content of the asset, followed by intellectual property metadata (who created it, who published it, who owns it, etc), followed by one or more Instantiation sections describing its physical or digital characteristics and locations. PBCore Instantiations take this form (although some are much more complex):

```

<pbcoreInstantiation>
<pbcoreFormatID>
<formatIdentifier>brasilccmovie_512kb.
mp4</formatIdentifier>
<formatIdentifierSource>pbcore XML
database online asset
</formatIdentifierSource>
</pbcoreFormatID>
<pbcoreFormatID>
<formatIdentifier>8ad6a40f-4fc0-4aa5-ba
a7-05c02e10da36</formatIdentifier>
<formatIdentifierSource>pbcore XML
database UUID</formatIdentifierSource>
</pbcoreFormatID>
<formatDigital>video/mp4
</formatDigital>
<formatLocation>e37d15ad-db8e-4af0-8d
dc-5338f9c5081f/brasilccmovie_512kb.
mp4</formatLocation>
<formatTracks>2</formatTracks>
<pbcoreEssenceTrack>
<essenceTrackType>video</essenceTrackType>
<essenceTrackIdentifier>1
</essenceTrackIdentifier>
<essenceTrackIdentifierSource>ID
(Mediainfo)</essenceTrackIdentifierSource>
<essenceTrackEncoding>AVC
</essenceTrackEncoding>
<essenceTrackBitDepth>8 bits
</essenceTrackBitDepth>
<essenceTrackFrameSize>384 pixelsX240
pixels</essenceTrackFrameSize>
<essenceTrackAnnotation>Chroma-subsam
pling: 4:2:0</essenceTrackAnnotation>
</pbcoreEssenceTrack>
<pbcoreEssenceTrack>

```

```

<essenceTrackType>audio
</essenceTrackType>
<essenceTrackIdentifier>2
</essenceTrackIdentifier>
<essenceTrackIdentifierSource>ID
(Mediainfo)</essenceTrackIdentifierSource>
<essenceTrackEncoding>AAC
</essenceTrackEncoding>
<essenceTrackAnnotation>Channels: 2
channels</essenceTrackAnnotation>
</pbcoreEssenceTrack>
<pbcoreAnnotation>
<annotation>Container: MPEG-4
</annotation>
</pbcoreAnnotation>
<pbcoreAnnotation>
<annotation>File_size: 48.1 MiB
</annotation>
</pbcoreAnnotation>
<pbcoreAnnotation>
<annotation>Comment: license:
http://creativecommons.org/licenses
/by-nc-sa/2.0/</annotation>
</pbcoreAnnotation>
</pbcoreInstantiation>

```

With a focus on organising descriptive metadata, the objective was to facilitate sharing media items by giving users complete, well-thought-out, descriptions. The combination of descriptive, intellectual property, and technical metadata in PBCore is a ‘core’ of descriptors used to categorise media items so that other parties can successfully identify, search for, and find them.

RELEASE OF PBCORE 1.0

The PBCore 1.0 data dictionary was published in spring, 2005 (see: <http://pbcore.org>). It was presented as a single, streamlined standard to which other database structures, including those of PBS, NPR, major producing stations and other asset/content management systems, could be mapped. It was also considered a useful guide for launching an archival or

asset management system at a station or production group.

In this form, PBCore v1.0 worked as a basic ‘starter kit.’ However, it was recognised that v1.0 would require improvements once stations and others began to use it. The addition of extensions to the existing set of metadata elements was planned as a means to accommodate such practical needs. Over the subsequent two years, there was uneven but slow acceptance across public radio and television, especially as stations increased the volume of content being offered on-line — audio and video clips in particular.

At the same time, PBCore was discovered by moving image and media archivists as an especially useful cataloguing format for film and video collections. From 2005–2008, PBCore sessions were presented widely at conferences such as the Association of Moving Image Archivists; the National Educational Television Association; the annual PBS Technology Conference; and the Integrated Media Association, which spurred interest in its use far outside public broadcasting. Such high visibility led regional and academic film archives, along with the Library of Congress, to begin treating PBCore as an ‘access and archival media’ metadata standard.

Kept active by users

Development was put on hold from 2007–2009, when funding from CPB was dormant. During this period, the standard was kept alive by a community of users who were applying PBCore to web distribution, educational access, and archival preservation.

The user group was large enough to support itself through information sharing and ‘unofficial’ websites. One site, PBCore Resources (see: <http://pbcoreresources.org>) was created as a common area ‘to collaborate and share resources

related to PBCore, a metadata standard developed to exchange information about audio/video media objects'. An informal site for news and discussion, it has become a primary sounding board for PBCore users to share information on their experiences testing and implementing PBCore in a variety of environments.

Another site developed by Dave Rice and Mike Castleman was the online 'PBCore database'. Powered by Ruby on Rails, Apache Solr, MySQL, nginx, and other third-party modules, the 'PBCore Repository Tool' (as it has become known in the user community) is an open-source application available for free download at <http://pbcore.vermicel.li/>. This public repository holds examples of 1,200+ film and video entries, and the cataloguing tool can generate records in XML or in a standard viewable format. This tool has since been adopted at several high-profile public media archives, notably at WNYC Radio in New York.

PBCORE 2.0

In 2009, CPB began planning the creation of The American Archive, a new initiative with a goal to organise and preserve America's social history as reflected in the extensive programme legacy of 50 years of public radio and television programmes. It was the first time CPB had ever allocated funds for preservation activities, and a number of American Archive initiatives are currently under way to create a basic framework for the new operation.

The very first project of the *American Archive* was to re-establish funding for PBCore and continue its development. With this restored support, Version 2.0 was released in January 2011. While still needing improvement, PBCore 2.0 is considered an important upgrade over the initial schema (see Figure 3).

Since the autumn of 2010, the *American Archive Content Inventory Project*,

coordinated at WGBH, has been spearheading a national effort to build an inventory of public radio and television collections across the country, and to date, more than 3m assets have been counted. See <http://americanarchiveinventory.org/project/>

A critical requirement for programmes to be accepted into the *American Archive* is that programme records be PBCore-compliant. Consequently, the basic information for each entry is a PBCore record. The inventory is being rapidly expanded and new entries are being made daily, which has also resulted in an instant increase in PBCore users among public broadcast stations and programme producers (see Figure 4).

PBCORE IN A PRODUCTION MAM

Much of the implementation of PBCore has been with archival and pre-recorded media. Adapting it for use in a production environment requires a different approach.

WNET in New York is the producer of *NATURE*, *American Masters*, *Great Performances*, *Religion & Ethics Newsweekly*, and a significant number of other major programmes that, over the years, have helped make up the national PBS prime-time schedule. The station has decades of production experience, and it recognised that shifting from tape-based to digital file-based production workflows was necessary to provide streamlined and cost-effective creation, distribution, and archiving methods for all the media produced within the organisation — not solely broadcast, but education, online, promotion and other units as well. To achieve this goal, WNET introduced a video media asset management (MAM) system for production and archiving.

WNET wanted a MAM that would address the dual needs of managing video files both for programme production and for preservation. Therefore, the workflow

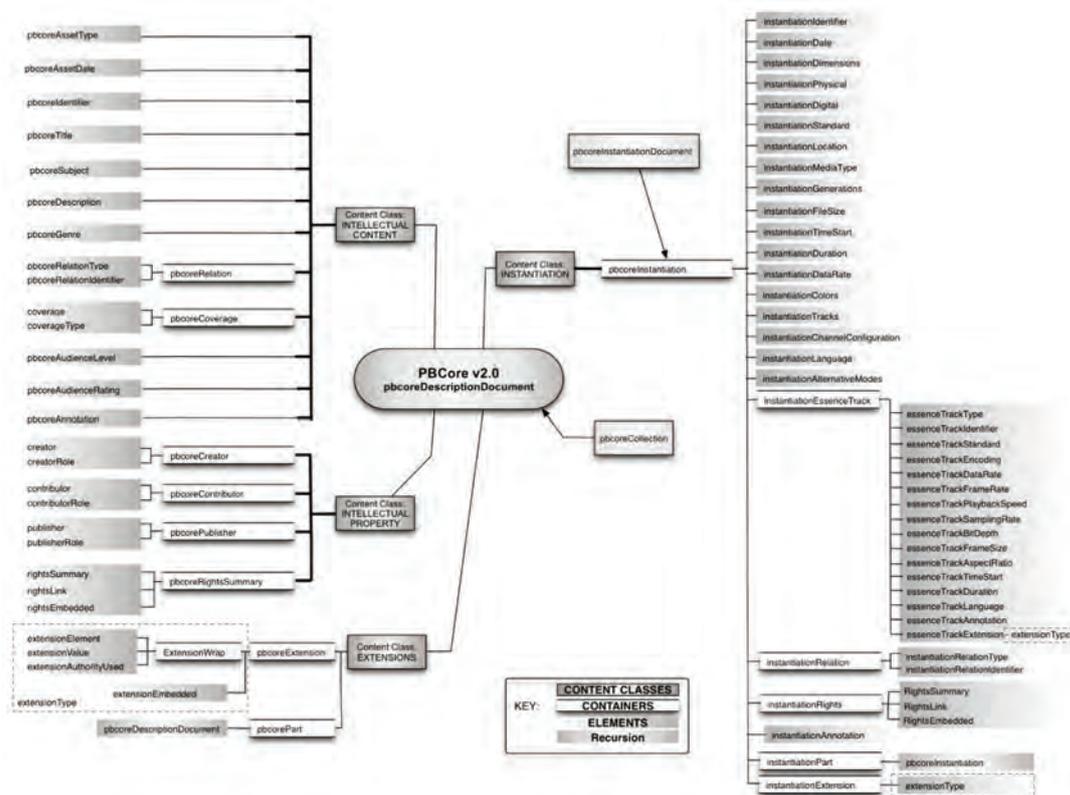


Figure 3: Schema: PBCore v2.0

had to include: the systematic collection of metadata throughout production; managing a limited number of standard file formats; technical tools that would facilitate the flow of media and metadata from inception to archiving; safe and efficient internal content storage; and automated exports of rich media packages for distribution and to a preservation repository.⁷

Among other requirements, WNET wanted its production MAM configured to be ‘preservation-compliant’, meaning the system had to be optimised to create standardised packages of rich media and metadata that could be delivered to a preservation repository as a Submission Information Package (SIP) ready for ingest. These packages would have to include all the information the repository required to provide long-term preservation and access to the content in the package.

From its experience of building a model preservation repository through the Library of Congress-funded project *Preserving Digital Public Television*,⁸ WNET had already learned the importance of collecting uniform metadata about programme productions. A key conclusion of the project was that ‘diverse file formats and inconsistent and haphazard metadata records that did not conform to any standard was simply unsustainable’.⁷ Adopting standards for metadata collection was thus a key requirement that had to be implemented.

It was decided to use PBCore as the basic schema for descriptive and technical metadata. The general process for designing the metadata structure and fields within the MAM focused on specific in-house production use cases. Based on the analysis of the use case, a concise set of fields appropriate for that use case was

American Archive Content Inventory Project - Wireframe for Web Entry

What's on this Asset? - Intellectual Content

*type media type

*title

id source of id

description

subject source of subject

genre source of genre

language coverage - location

date created date broadcast

*RED entry fields are REQUIRED, remaining are recommended

Note: do not include national programming which your organization did not produce and which is still within distribution rights.

Upload a photo (jpg) associated with this Asset:

Browse:

Who made this Asset? - Intellectual Property

publisher role

creator role

contributor role

rights notes

About this Asset - Instantiation

*unique id *unique id source condition

*format physical and/or digital

*generation

*duration Approximate? Yes No

*location

born yes unsure no

digital? no

encoding

file size

Extra Questions Are there any ancillary materials associated with this asset? Yes No

If yes, what kind?

<input type="checkbox"/> Awards	<input type="checkbox"/> Computer Disks	<input type="checkbox"/> Maps	<input type="checkbox"/> Legal materials	<input type="checkbox"/> Research materials
<input type="checkbox"/> Books	<input type="checkbox"/> Hard Drives	<input type="checkbox"/> Visual Art	<input type="checkbox"/> Production materials	<input type="checkbox"/> Scripts
<input type="checkbox"/> Costumes	<input type="checkbox"/> Computer files	<input type="checkbox"/> Educational materials	<input type="checkbox"/> Promotional material	<input type="checkbox"/> Transcripts
<input type="checkbox"/> Merchandise	<input type="checkbox"/> Images	<input type="checkbox"/> Financial materials	<input type="checkbox"/> Proposals	<input type="checkbox"/> Logs

To the best of your knowledge, is this item unique? Yes No

If yes, how so?

Figure 4: Sample record American Archive Content Inventory Project

identified and corresponding PBCore elements would then be used as a template for final field design.

DALET, selected in 2010 as the MAM system, was configured to support import of PBCore XML as well as export of valid PBCore for delivery to internal systems. With this system in place, PBCore became

the preferred format for metadata delivered from external systems. However, because of database limitations, the full one-to-many asset-to-instantiation model enabled by PBCore could not be utilised in the import and export design.

Instead, a single PBCore record is output for each instantiation. When

importing PBCore records that contain multiple instantiations for a single asset (such as an episode of a TV programme with tape, file, and DVD instantiations), each instantiation must be turned into its own PBCore document containing all asset level (descriptive and rights) information. (At this point, a growing number of MAM systems promote being 'PBCore ready'.)

Slow acceptance among production units

Implementing the MAM has not been an easy task. Public media content is produced in a highly decentralised manner, with large producers like WNET and small independent producers each having their own individualised production workflows and systems for collecting required and optional metadata. Even within a single entity like WNET, various production teams can function autonomously or even on an ad-hoc basis.

Consequently, in the absence of any common procedures or best practices adopted station-wide, each production unit has been free to collect programme-related metadata in its own idiosyncratic fashion. Necessary metadata being carried at each stage of the production workflow was outside the traditional production process, so sharing it was both impractical and unreliable.

As a result, even though the MAM system was configured to optimise metadata standardisation, encouraging producers to alter their workflow habits and training them to conform to a more uniform set of practices has been challenging. Rather than sending e-mails, saving text documents or keeping notebooks, which have been their long-accepted working methods for record-keeping, producers are now required to enter their metadata directly into the MAM using PBCore vocabularies

and other taxonomies. At the time of this writing, the station continues to experience resistance to this change.

Making sure every user is trained specifically on input, data sharing and collaboration using the MAM is key to its success. Because long-term acceptance and trust by producers is still a hurdle, implementing the MAM across numerous production units at WNET remains a work in progress.

OTHER PBCORE USERS

Today, PBCore is being used by a variety of organisations, broadcasters and archives. With its public broadcasting orientation, most of the groups adopting it are not-for-profit organisations.

These groups include the following:

- **Illinois Public Media** at the University of Illinois began using PBCore in 2006 for its online media collections and assets. Built into its content workflow is a simple cataloguing tool that creates web pages with unique URLs for each asset, along with syndication methods like RSS and Atom, and archival-grade metadata expressed as PBCore XML records.
- **WHYY Radio** — Since 1975, WHYY Radio has been producing *Fresh Air with Terry Gross*, creating thousands of ¼-inch analogue reels, DAT tapes, CDs, and digital files, as well as countless Microsoft Word and Excel documents reflecting rights, inventory, descriptive, and technical information. In 2010, WHYY began creating a digital preservation and access archive of these materials through a new DAM system, configured specifically with PBCore as the primary metadata standard for descriptive and technical metadata. A number of specialised tools have been developed to transform FA's extensive written records including transcripts,

guest lists, and similar programme materials, into searchable PBCore database records.

- **WNYC Radio** — On the air since 1924, WNYC is a public radio station formerly owned by the City of New York. The station is a central repository for thousands of recordings, photographs, memorabilia, and institutional records, and the archives hold a collection of more than 50,000 recordings in a variety of formats, from early lacquer and acetate discs to reel-to-reel tape, DBX 700, F1 Beta, DAT, MD and CD formats. The WNYC Archives, the largest at a public radio station, is also using PBCore as its central metadata schema for materials as they become digitised.
- **Witness** — A human rights organisation that uses video to transform personal stories into powerful images for justice and policy change, Witness uses PBCore as a primary tool in their catalogue of over 4,000 hours of video footage. By aligning their catalogue fields with the hierarchies, elements, and meanings specified in PBCore, they can create PBCore-compliant, human- and machine-readable exports that can be exchanged and used by others.
- **Public access media centres** — Public access production centres bring public voices and community events to hundreds of local cable systems across the country. For several years, a working group of the Alliance for Community Media has been adopting PBCore as a module for Drupal, an open-source Content Management System (CMS) used at many facilities for managing databases and web content. Their work, developed collaboratively for exchanging files in a local production environment, has resulted in a free PBCore module which extends the functionality of any Drupal-based media

website. Members of the working group have posted training videos⁹ to illustrate how to adapt PBCore to a user's own Drupal configuration, and they encourage others to contribute. It has been so well received that the group has taken up extending the use of PBCore to exchange program files between facilities as well.

- **Northeast Historic Film** is a Maine-based archives and study centre saving New England's moving image heritage, with a vault holding over 10m feet of film. Using a web-based interface, they are using PBCore to catalogue their collections and display the entries by linking them directly to their online finding aids.
- **Dance Heritage Coalition** is an organisation of ten dance-related collections, sharing resources both to preserve their collections and to make them more accessible. They used PBCore to standardise descriptive vocabularies, restructure data and map fields from a wide range of disparate metadata schemata including MARC21, MARCXML, custom SQL, Microsoft Access databases, and a number of custom databases. PBCore provides consistent results for searching and finding objects across collections.

However, PBCore is not being used solely by non-profits. It is also being discovered by other video producers, including:

- **Broadway Video Digital Media** — When producing a promotional video for the New York City Department of Transportation, BVDM used PBCore as the data model for DOT's Library Access Platform, which stored the various production elements. They found it was perfect to allow both technical and non-technical users to search through and view proxies of the raw footage used in the production.

STILL A NEED FOR CONTINUED SUPPORT

Despite the acceptance of PBCore by archives and others, among public broadcasters it is not universally understood that digital media files require standardised metadata. The question ‘What is PBCore for?’ continues to be raised, especially if a station already has a functional media database.

Public broadcasting stations and producers are not required to follow any specified system-wide metadata policies except those relating to programme distribution, so most stations have created their own conventions for cataloguing and managing their assets. Yet only common standards make it possible to share digital assets regardless of applications, distribution platforms or operating devices, which is why PBCore was developed — specifically to address this very need.

This raises two basic concerns — first, that many groups may need assistance if they do not actually have an appropriate programme-cataloguing database. The second is a larger issue. Many groups may have adequate database applications, but may not realise that it is not the database *per se* that facilitates file interoperability — it is having *standard schema for the entries in the database* that makes it work.

Through the concrete demonstration provided by the *American Archive Inventory Project*, for the first time stations are required to use PBCore and are seeing the value of adopting a shared metadata standard. Because of this, many stations and producers are now starting to look for guidance on organising metadata, if for no other reason than to manage their own burgeoning sets of digital recordings and production files.

The evolution of PBCore over the past five years has relied on the input and active participation of an expanding user community, and as PBCore improves, each

year new institutions are discovering it. These groups are not solely public broadcasters, but reflect a broader array of media producers and collections, and in response to a growing market, new open-source and vendor-provided PBCore tools are increasingly available

Continuing to support the use and improvement of PBCore is an important component in this larger picture. Ongoing support from CPB, and promoting it to public radio and television as an important metadata tool, remain especially critical now that its value has been demonstrated.

Acknowledgments

Special thanks are due to Marcia Brooks, Project Director of the *PBCore 2.0 Project*, Dave MacCarn, Chief Technologist at WGBH, and Jack Brighton, Director of New Media & Innovation, Illinois Public Media, University of Illinois Urbana-Champaign, for their assistance and advice in preparing this article. Additional material was provided by Jonathan Marmor, MAM Technical Project Manager, WNET, and Kara Van Malsen, Senior Consultant, AudioVisual Preservation Solutions.

References

1. PB Core — the Public Broadcasting Metadata Initiative: Progress Report 2003, available at: http://www.pbcore.utah.edu/resources/items/PBCore_DC03PaperFinal.pdf (accessed 9th March, 2012).
2. Consultative Committee for Space data Systems (2002) ‘Reference Model for an Open Archival Information System (OAIS), CCSDS, Washington, DC, available at: <http://public.ccsds.org/publications/archive/650x0b1.PDF> (accessed 30th July, 2011).
3. SMEF-DM: Standard Media Exchange Framework Data Model (2007), 15th August, available at: <http://www.bbc.co.uk/guidelines/smef/> (accessed 30th July, 2011).
4. International Organisation for Standardisation (2004) MPEG-7: ‘Multimedia Content Description Interface’ ISO/IEC JTC1/SC29/WG11, October, available at: <http://mpeg.chiariglione.org/standards/mpeg-7/mpeg-7.htm> (accessed 30th July, 2011).
5. Library of Congress – Network Development and MARC Standards Office (no date) MARC: Machine Readable Cataloging, Library of Congress, Washington, DC, available at: <http://www.loc.gov/marc/> (accessed 30th July, 2011).
6. Dublin Core Metadata Initiative (2010) ‘Dublin Core Metadata Element Set, Version 1.1:

- Reference Description’, DCMI, Dublin, OH, available at: <http://www.dublincore.org/documents/dces/> (accessed 30th July, 2011).
7. National Digital Information Infrastructure and Preservation Program: Preserving Digital Public Television Project (2011) ‘A Preservation-Compliant Media Asset Management System for Television Production’, WNET, 27th April, available at: <http://cn2.wnet.org/thirteen> /ptvdigitalarchive/files/2011/04/Preservation CompliantMAM2.pdf (accessed 30th July, 2011).
 8. Preserving Digital Public Television Project website, available at: <http://www.thirteen.org/ptvdigitalarchive/> (accessed 9th March, 2012).
 9. Groups.Drupal (2010) ‘How-to set up PBCore & Genres for OMP’ 30th July, available at: <http://groups.drupal.org/node/83804> (accessed 30th July, 2011).

APPENDIX

PBCore v2.0 schema chart, pbcoreCollection, pbcoreDescriptionDocument

Content Class	Containers	Elements
Intellectual Content	pbcoreAssetType	pbcoreAssetDate pbcoreIdentifier pbcoreTitle pbcoreSubject pbcoreDescription pbcoreGenre
	pbcoreRelation	pbcoreRelationType pbcoreRelationIdentifier
	pbcoreCoverage	coverage coverageType pbcoreAudienceLevel pbcoreAudienceRating pbcoreAnnotation
Intellectual Property	pbcoreCreator	creator creatorRole
	pbcoreContributor	contributor contributorRole
	pbcorePublisher	publisher publisherRole
	pbcoreRightsSummary rightsLink	rightsSummary rightsEmbedded
Instantiation	pbcoreInstantiation	(pbcoreInstantiationDocument) InstantiationIdentifier InstantiationDate InstantiationDimensions InstantiationPhysical InstantiationDigital InstantiationStandard InstantiationLocation InstantiationMediaType InstantiationGenerations InstantiationFileSize InstantiationTimeStart InstantiationDuration InstantiationDataRate InstantiationColors InstantiationTracks InstantiationChannelConfiguration InstantiationLanguage
	InstantiationEssenceTrack	essenceTrackType essenceTrackIdentifier essenceTrackStandard essenceTrackEncoding essenceTrackDataRate essenceTrackFrameRate essenceTrackPlaybackSpeed essenceTrackSamplingRate essenceTrackBitDepth

APPENDIX (continued)

PBCore v2.0 schema chart, pbcoreCollection, pbcoreDescriptionDocument

<i>Content Class</i>	<i>Containers</i>	<i>Elements</i>
Extensions	instantiationRelation instantiationRights instantiationPart instantiationExtension pbcoreExtension ExtensionWrap extensionValue pbcorePart	essenceTrackFrameSize essenceTrackAspectRatio essenceTrackTimeStart essenceTrackDuration essenceTrackLanguage essenceTrackAnnotation essenceTrackExtension - instantiationRelationType instantiationRelationIdentifier RightsSummary RightsLink RightsEmbedded instantiationAnnotation pbcoreInstantiation (Recursion) extensionType extensionEmbedded extensionElement extensionAuthorityUsed pbcoreDescriptionDocument (Recursion)