

Introduction to PBCore and Metadata

The PBCore Project at the
Carl and Ruth Shapiro Family National Center
for Accessible Media at WGBH (NCAM)

<http://www.pbcore.org>

<http://ncam.wgbh.org/pbcore>



What to expect from this presentation

- In this session you will learn...
 - what metadata is and why it matters;
 - what PBCore is and why you should consider using it;
 - how PBCore is structured;
 - what PBCore looks like in actual use;
 - where to go for more information about PBCore.

What is PBCore?

- ...a metadata dictionary, used to describe many different types of media items;
- ...useful for locating, sharing and exchanging media items;
- ...available for free to anyone, not just public broadcasters.

3

Who created PBCore?

- Unprecedented collaboration between public television and radio united a working group of 30 representatives of local and national organizations, from a wide cross-section of disciplines and roles.
- Advisors included academic and metadata experts.
- Before its launch, PBCore had wide industry review and key test implementations.
- WGBH administered the project.

4

What is the PBCore project at NCAM?

- Project objectives:
 - Define, promote and advocate for the use of PBCore
 - Ensure PBCore becomes well known to the public broadcasting system
 - Educate content producers, distributors and vendors about PBCore
- Project timeline: 3/2005-4/2007

5

What is the PBCore project at NCAM?

- The project is currently...
 - assessing the public broadcasting system's training needs;
 - developing effective training materials and creating software tools to support the use of PBCore;
 - identifying strategies for PBCore sustainability.

6

Part I

Background:

What, exactly, is metadata?

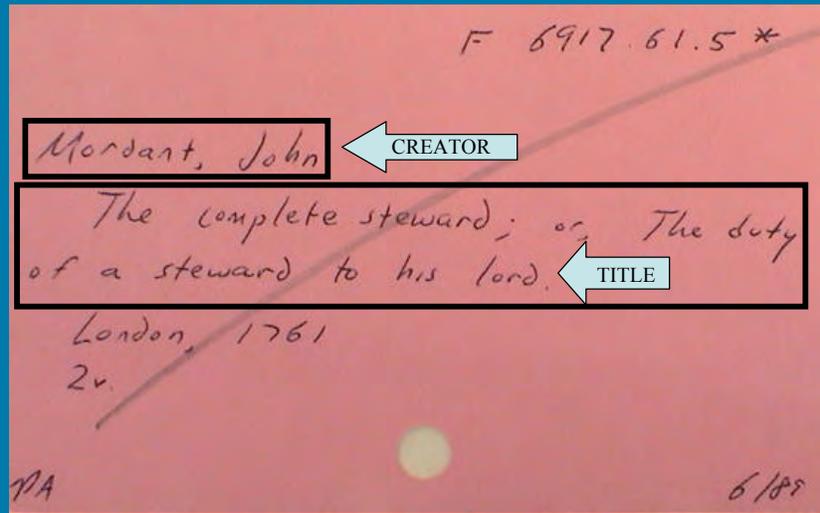
7

What is metadata?

- Metadata can be used to describe and identify media items.
- A media item can be anything-- a book, a video or audio tape, an electronic video or audio file, a photograph, an electronic image, a Web site, etc.

8

This catalog card (circa 1989) contains metadata about a book



This electronic record (circa 2006) contains metadata about a videotape

Media Library Online	
Home	Browse Resources
Forums	About
Help	
Tape Number:	FF0084
Title:	Budget Energy #3
Title Type:	Field Footage
Program or Series Title:	In Wisconsin
Creator:	Freyberg, Frederica//Producer
Date Created:	2005-03-08
Publisher:	Board of Regents UW//Copyright holder Wisconsin Public Television//Presenter
Subject:	energy and resources alternative energy renewable energy

10

How is metadata used?

- Metadata can be used to search for, or **locate**, a media item.

11

This card catalog (circa 1930)
is a metadata search tool



12

This electronic search engine (circa 2006) is a metadata search tool

The screenshot shows a web-based search interface with a navigation bar at the top containing links for Home, Browse Resources, Forums, About, and Help. Below the navigation bar is a section titled "Advanced Search". This section is divided into two main parts. The first part, "Field:", contains four rows of dropdown menus for selecting search fields: Description, Title, Subject, and Creator. To the right of these dropdowns is a "Search Terms:" input field. The second part, "Limits (optional):", contains three more dropdown menus: "Format:" with options like 16mm Film, Beta, BetaSP, 3/4 in, DVD-rom, and DVCPRO25; "Title Type:" with options like Virtual Pledge, Home Video, Pledge Versior, Alternative, Series, and Episode; and "Program or Series Title:" with options like Here and Now, In Wisconsin, Hometown Stories, Wisconsin Stories, World War II Stories, and Korean War Stories.

How is metadata used?

- Metadata uses **elements** to describe various aspects of a media item. Examples of elements are:
 - creator (Welles, George Orson)
 - publisher (Secker and Warburg)
 - title (*Citizen Kane*)
 - formatPhysical (VHS, DVD-Video)
- Individual elements form the complete **metadata element set**.

How is metadata used?

- The more metadata used to describe a media item...
 - the more ways that item can be identified;
 - the greater the chances of a successful search.
- A media item described with too little metadata may be difficult for a user to locate.
 - For example, including the title in an electronic record but neglecting to include the creator, publisher or physical format.

15

When is metadata useful?

- Sufficiency of metadata
 - too much, too little, just right?
- Accuracy and quality of metadata
 - correct data entry; following guidelines and standards

16

PBCore Metadata Primer

<http://www.pbcore.org>; click on "User Guide" button

A Primer for Understanding the PBCore Metadata

- ✦ **What is Metadata, Anyway?**
- ✦ **What is a Metadata Dictionary?**
- ✦ **What is a Metadata Element?**
- ✦ **What is a Content Class and an Element Container?**
- ✦ **What is an Authority, Controlled Vocabulary & Structured Syntax?**
- ✦ **What are Element Attributes?**
- ✦ **What is an Application Profile?**
- ✦ **What is the Dublin Core (DCMI)?**
- ✦ **What is the Public Broadcasting Core (PBMD Project)?**
- ✦ **How do I Review the Metadata Elements of the PBCore?**
- ✦ **What is a XML Schema Definition (XSD)?**
- ✦ **What is a Namespace?**
- ✦ **Do the PBCore Elements have a Hierarchical Tree Structure?**
- ✦ **Do PBCore Elements Map to other Metadata Schemas?**
- ✦ **More Resources for the Curious and Compulsive**

17

Part II What, exactly, is PBCore?

18

What is PBCore?

- PBCore is a metadata **dictionary** that can describe many kinds of media items: video or audio tapes and electronic files, text, PDF documents, electronic images, etc.
- PBCore can be used by anyone, not just public broadcasters or television and radio stations.

19

Why use PBCore?

- Public broadcasters that use PBCore to describe their media items can...
 - track them from creation to distribution to archiving;
 - more easily share them within an organization, with other public broadcasters, and other content and service partners.
- All public-broadcasting entities are encouraged to implement PBCore.

20

Why use PBCore?

- The PBCore metadata dictionary is available for free.
- Licensing information can be found at the PBCore Web site (<http://www.pbcore.org>; click on “Licensing PBCore” button).

21

What is the PBCore vocabulary?

- PBCore v1.1 (<http://www.pbcore.org>; click on “User Guide” button) was published January 24, 2007
- Version 1.1 of PBCore maintains 53 elements arranged in 15 containers and 3 sub-containers, all organized under 4 content classes.

22

What is the PBCore dictionary?

- PBCore is made up of 53 elements, each describing a different aspect of a media item.
- Three sample PBCore elements are:
 1. **title**: the media item's name;
 2. **creator**: the person or organization responsible for creating the media item;
 3. **publisher**: the person or organization responsible for distributing the media item.

23

What is the PBCore dictionary?

- You can see the complete list of the 53 PBCore elements at the [PBCore Dictionary](http://www.pbcore.org) (<http://www.pbcore.org>; click on “User Guide” button).
- You don't need to use all 53 elements when describing a media item with PBCore.

24

What is the PBCore vocabulary?

- PBCore often uses a **controlled vocabulary** to describe a media item.
- A controlled vocabulary is a prescribed set of words or phrases that may be used for description.
- Controlling the terms used, and how they are presented, not only ensures consistent description of media items, it also improves the chances of a successful search.

25

What is the PBCore vocabulary?

- A **picklist**, or drop-down menu, is an example of a controlled vocabulary. Here is a picklist for the PBCore element 05.01, genre.



26

What is the PBCore vocabulary?

- PBCore also uses a **structured syntax**.
- A structured syntax follows specific rules of punctuation, grammar or data entry in order to provide unambiguous descriptions.
- For example...
 - creator (element 15.01): last name, first name middle name (Welles, George Orson)
 - title (element 02.01): Citizen Kane
 - dateCreated (element 25.01): YYYY-MM-DD (1941-05-01)

27

Can I see a PBCore cataloging application?

- CWIS (Collection Workflow Integration System) at Wisconsin PTV

The screenshot shows a web form titled "Add New Resource" with the following fields and options:

- Buttons: "Add Record" and "Cancel"
- Text: "A '*' next to a label indicates a required field."
- Form fields:
 - Tape Number: [text input]
 - Title*: [text input]
 - Title Type: [dropdown menu with options: Alternative, Collection, Episode]
 - Program or Series Title: [dropdown menu with options: Here and Now, Hometown Stories in Wisconsin, Korean War Stories]
 - Project Number: [text input]
 - Search/Add Creator(s): [dropdown menu with "None Selected"]
 - Search/Add Contributor(s): [dropdown menu with "None Selected"]
 - Date Created: [text input]
 - Broadcast Date: [text input]
 - Search/Add Publisher(s): [dropdown menu with "None Selected"]
 - Subject(s): [text input with "Add More Subjects" button]

28

Can I see a PBCore cataloging application?

- CWIS (Collection Workflow Integration System) at Wisconsin PTV

Rights Summary:

Audience Level: College, Educator, Female, General

Audience Rating: TV-14 (D), TV-14 (L), TV-14 (S)

Format Standard: Audio/MPEG, Audio/SDS, MPEG video, NTSC video

Language: eng (English), fre (French), german, spa (Spanish)

Date Of Record Release*: 2006-10-11 08:19:25

Date Record Checked*: 2006-10-11

Date Available Ends:

29

Can I see a PBCore cataloging application?

- Media Exchange (NETA/UEN) http://www.uen.org/media_exchange

media exchange PROVIDER

Back To Results

Title: Glen Canyon: A Dam, Water

This title is a(n): Program

Comes from a collection name:

Media type: Video

Duration of media item: 00:56:04

Funding credit:

SUBMISSION ID INFO:

Submission ID:

Submission ID source:

DESCRIPTION:

Description narrative: Straddling the Utah-Arizona border, the Glen Canyon Dam rises over 700 feet from solid bedrock to block the flow of the Colorado River. Behind its wall of

SUBJECTS:

Save

30

Can I see a PBCore cataloging application?

- Media Exchange (NETA/UEN) http://www.uen.org/media_exchange

The screenshot shows the 'media exchange SEEKER' interface. On the left, there is a search bar with 'Quick Search:' and a 'Find' button, and a 'Catalogs' section listing: Antelope Island (4 items), KUED-TV (30 items), PBS (18 items), Utah Collections (31 items), and Workplace Essential Skills (27 items). The main content area displays 'KUED-TV - 30 documents' with a 'View: Summary' dropdown and a 'Page 1 of 2' indicator. Two document thumbnails labeled 'New Document' are shown. The first document details are: Title: Artists of the West Program; Genres: Art Biography Documentary Educational Animals Horse History Nature Western; Digital or physical types: Digital Video; Duration of media item: 00:55:46; File size: 336.1 MB; Format unique ID/file name: KUED_AOTW.mov. The second document details are: Title: Atomic Stampede Program; Genres: Documentary Educational History Law Nature Politics; Digital or physical types: Digital Video; Duration of media item: 00:55:38; File size: 333.8 MB; Format unique ID/file name: KUED_AS.mov.

31

Can I see a PBCore cataloging application?

- Media Exchange (NETA/UEN) http://www.uen.org/media_exchange

The screenshot shows the 'media exchange SEEKER' interface with detailed information for a document. On the left, the search and catalog sections are visible. The main content area displays 'Back To Results' and 'Info' for the document 'Glen Canyon: A Dam, Water and the West'. The details include: Title: Glen Canyon: A Dam, Water and the West; Format unique ID/file name: KUED_GC_ADWWW.mov; This title is a(n): Program; Comes from a collection named: ; Digital or physical types: Digital Video; Genres: Documentary Educational History Law Nature Politics Community Science Environment Western; Description narrative: Straddling the Utah-Arizona border, the Glen Canyon Dam rises over 700 feet from a red bedrock to block the flow of the Colorado River. Behind its wall of five-million yards of concrete are the waters of Lake Powell. Storing enough water to meet the needs of almost thirty million families for a year, it is one of the largest man-made reservoirs in the world. The program offers a comprehensive history of the dam from early compromises between the government and environmental activists to archeological efforts to document the history of Glen Canyon. The Glen Canyon Dam created one of the world's largest man-made bodies of water, while at the same time preserving a thousand years of human history... and a million years of natural history. Teams from the Museum of Northern Arizona and the University of Utah race against time to salvage artifacts for future generations to study. But ancient Indian dwellings and rock art were inevitably lost to the rising waters of Lake Powell. Illuminating both sides of the debate over use and development of natural resources, Glen Canyon features interviews with two key players: David Brower, director of the Sierra Club during the 1950s and Ray Dornay, the Commissioner of the Bureau of Reclamation during the dam's construction. Some three million people a year visit the dam and Lake Powell. There's always been an important

32

What is PBCore compliant?

- When setting up a metadata structure, follow PBCore dictionary attributes and rules.
 - Underlying elements have matching meanings, even if their names vary.
 - Elements have matching data structures
 - data types
 - range of values
 - known authorities
 - matching vocabulary terms (picklists)
 - matching syntax
 - If variances must be implemented, document them.

33

Part III Summary and Resources

34

PBCore is...

- a metadata dictionary used to describe many different types of media items;
- useful for locating, sharing and exchanging media items;
- available for free to anyone, not just public broadcasters.
- Licensing information can be found at the PBCore Web site (<http://www.pbcore.org>; click on “Licensing PBCore” button).

35

PBCore listserv



- Be part of a community of practice where PBCore adopters share information and resources
- To **join** :
 - Go to the PBCore Web site (<http://www.pbcore.org>) and click on the “Join Listserv” button.

36

Upcoming PBCore training sessions

- PBCore 101: An Overview

- PBCore 102: The Elements

- The PBCore XML Schema

- Cataloging Tools for PBCore

No.	Topic	Description	Live WebEx Time	On-Demand Availability
I	<i>PBCore 101: An Overview</i>	An introduction to the nature of metadata and the PBCore.	Wednesday 2007-01-31/2:00pm EST Repeated Thursday 2007-02-01/2:00pm EST	QuickTime Windows Media PowerPoint
II	<i>PBCore 102: The Elements</i>	A more detailed look at the structure of PBCore, what each of the elements are, and how to use them to describe media objects of all kinds.	TBA	QuickTime Windows Media PowerPoint
III	<i>The PBCore XML Schema</i>	A general discussion of what an XML Schema is, and a look at how the PBCore XML Schema can be used to implement the PBCore metadata dictionary within information and cataloging systems. Also explains how important the XML Schema is for sharing metadata and interoperability.	TBA	QuickTime Windows Media PowerPoint
IV	<i>Cataloging Tools for PBCore</i>	A detailed discussion of tools that use PBCore. The tools facilitate the markup of descriptions for media objects and have data export capabilities for sharing and interoperability.	TBA	QuickTime Windows Media PowerPoint

Resources

General

- [PBCore.org \(http://www.pbcore.org/\)](http://www.pbcore.org/)
- The [PBCore listserv \(http://www.pbcore.org\)](http://www.pbcore.org)
- The [PBCore project at NCAM \(http://ncam.wgbh.org/pbcore\)](http://ncam.wgbh.org/pbcore)
- The [PBCore XML Schema \(draft; http://www.pbcore.org; click on “Resources and Links” button\)](http://www.pbcore.org)

Resources

Metadata

- PBCore metadata resources (<http://www.pbcore.org>; click on “Resources and Links” button)
- "The Case for Shared Metadata Standards" (<http://www.current.org/tech/tech0209metadata.html>)
- A Primer for Understanding PBCore
- PBCore update (<http://www.pbcore.org>; click on “User Guide” button)

39

Resources

Metadata search tools

- Wisconsin Public TV's Media Library Online (<http://wptmedialibrary.wpt.org/>), a PBCore implementation through CWIS (<http://scout.wisc.edu/Projects/CWIS/>)
- NETA/Utah Education Network Media Exchange http://www.uen.org/media_exchange
- RELOAD (<http://www.reload.ac.uk/>)

40

Contact information

Marcia Brooks
Project Director, PBCore Project
marcia_brooks@wgbh.org