

PBCore v1.1

March 28, 2007

SURA ViDe (Atlanta, GA)

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

<http://www.pbcore.org>



1

What is PBCore?

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

2

Who created PBCore?

- A metadata working group consisting of 30 representatives of local and national television and radio organizations, as well as numerous constituencies and disciplines.
- Advisors included:
 - University of Washington Information School
 - the Department of Defense's Academic CoLab
 - Rutgers University/The Association of Moving Image Archivists.
- WGBH administered the project.

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 & associated communities
 - Educate content producers, distributors and vendors about PBCore
- Project timeline: 3/2005-4/2007

What is the PBCore project at NCAM?

- The project will...
 - assess the public broadcasting system's training needs;
 - develop effective training materials and create software tools to support the use of PBCore;
 - support and encourage a PBCore Community of Practice
 - identify strategies for PBCore sustainability.

5

What, exactly, is metadata?

6

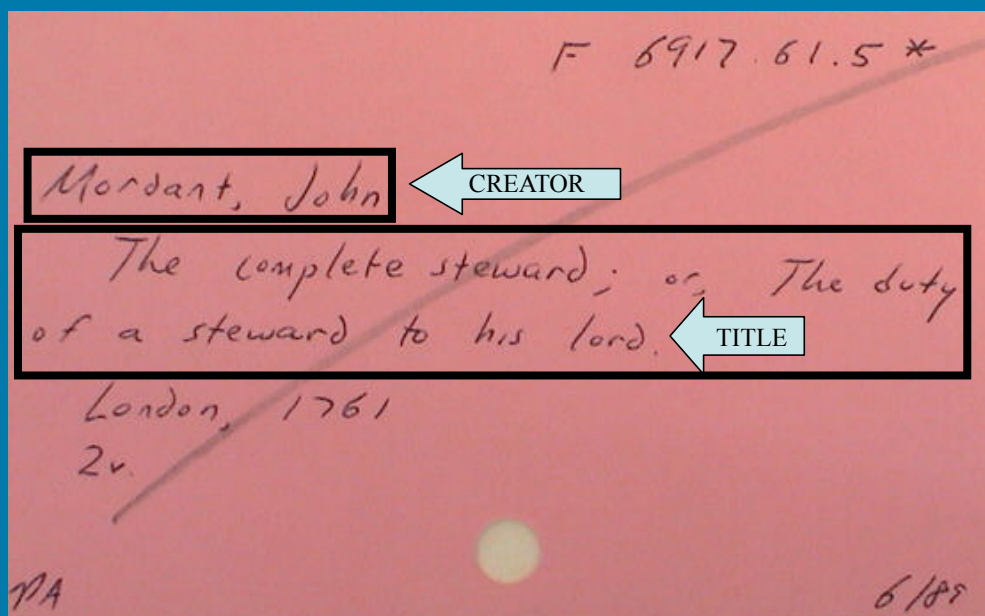
What is metadata?

- Metadata can be used to describe and identify assets or media items.
- An asset can be anything-- a book, a video or audio tape, an electronic video or audio file, a photograph, an electronic image, a Web site, interactive learning objects, etc.
- An asset is...

Essence (the thing)
+
Metadata (the description)

7

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



8

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

9

How is metadata used?

- ENHANCES FINDABILITY
 - Metadata can be used to...
 - search for an asset
 - locate an asset
 - review & evaluate an asset
 - Metadata requires Search Tools to be effective

10

This is not a search tool, nor
is findability well served...



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: Home, Browse Resources, Forums, About, and Help. Below the navigation bar is a section titled "Advanced Search".

Under "Advanced Search", there are two main sections:

- Field:** A vertical list of dropdown menus for selecting search fields: Description, Title, Subject, and Creator.
- Search Terms:** A series of text input boxes corresponding to the fields above.

Below the search fields is a section titled "Limits (optional):".

- Format:** A list of video formats: 16mm Film, Beta, BetaSP, 3/4 in, DVD-rom, and DVCPRO25.
- Title Type:** A list of title types: Virtual Pledge, Home Video, Pledge Version, Alternative, Series, and Episode.
- Program or Series Title:** A list of titles: Here and Now, In Wisconsin, Hometown Stories, Wisconsin Stories, World War II Stories, and Korean War Stories.

13

How is metadata used?

- Metadata uses elements to describe various aspects of an asset. Examples of elements are:
 - creator (Welles, George Orson)
 - publisher (Secker and Warburg)
 - title (*Citizen Kane*)
 - formatPhysical (VHS, DVD-Video)
- Individual elements, when combinend, form the complete metadata element set...
- ...such as PBCore v1.1

14

How is metadata used?

- The more metadata used to describe an asset...
 - ...the more ways that asset can be identified;
 - ...the greater the chances of a successful search
- An asset described with too little metadata may be difficult for a user to locate & evaluate.
 - e.g., including the title in an electronic record but neglecting to include the creator, publisher, physical format, or a description
 - e.g., on-screen digital TV guides with a series title, but no program title or description

15

How is metadata used?

Good metadata enhances “findability”

Good metadata is an “enabler”

16

When is metadata useful?

- Question of Sufficiency of metadata
 - when is it too much?
 - when is it too little?
 - when is it just right? (15 elements?)
- Question of Accuracy/Quality of metadata
 - well-formed data entry
 - follow metadata guidelines and standards
 - avoid Folksonomies

17

But wait...

...what, exactly, is PBCore?

18

What is the PBCore dictionary?

- PBCore is made up of 53 elements, each describing a different aspect of an asset.
- Four classes of PBCore elements
 1. Intellectual Content (16)
 2. Intellectual Property (7)
 3. Instantiations/Formats (28)
 4. Extensions

19

What is the PBCore dictionary?

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

20

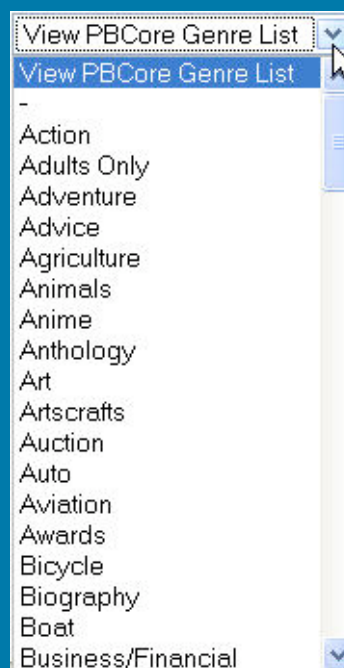
What is the PBCore dictionary?

- PBCore often uses a **controlled vocabulary** to describe an asset.
- 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 assets, it also improves the chances of a successful search.

21

What is the PBCore dictionary?

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



22

What is the PBCore dictionary?

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

23

What is the PBCore dictionary?

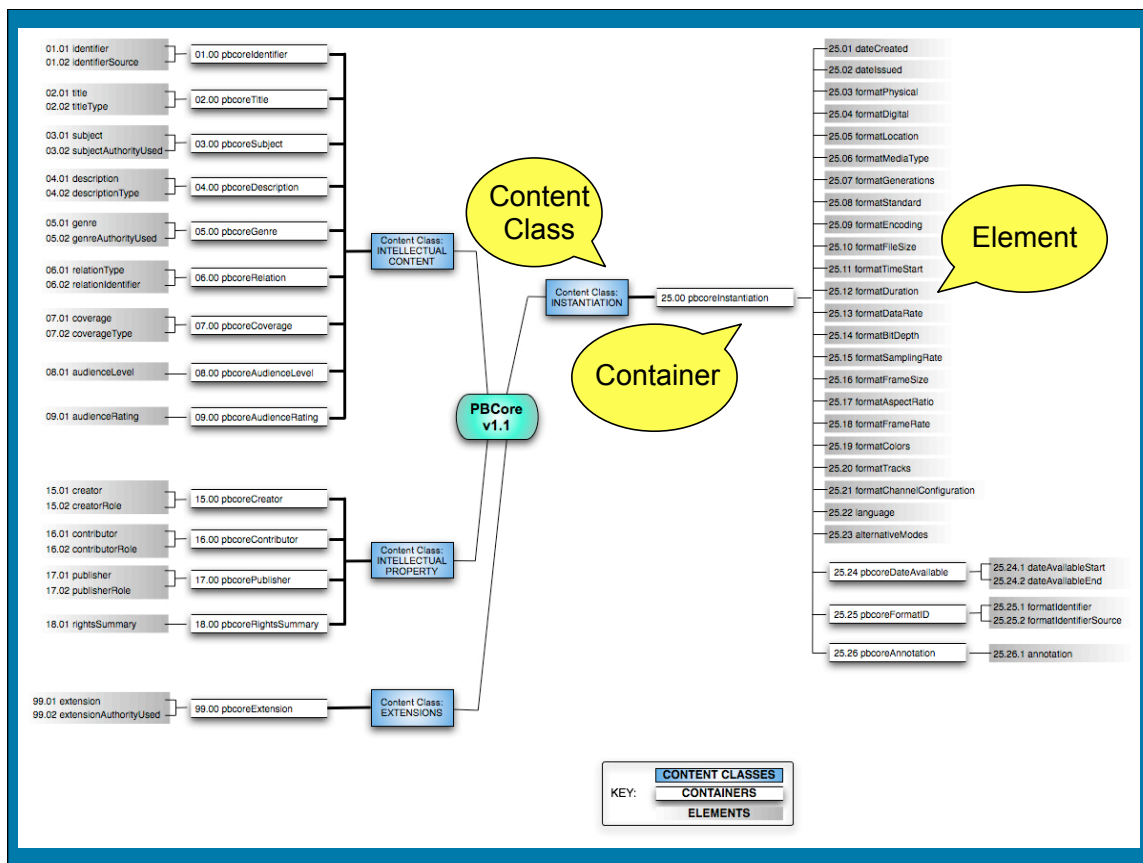
- PBCore also uses established **authorities**.
- An authority is a pre-existing set of terms and usage guidelines that is considered standardized and in widespread use.
- For example...
 - Library of Congress “Subject Authorities”
<http://authorities.loc.gov/help/contents.htm>
 - Library of Congress “Name Authorities”
<http://authorities.loc.gov/help/name-auth.htm>
 - PBS PODS “Subject Terms”
<https://secure.connect.pbs.org/pbsdocuments/pbs/PBCore/>
 - ID3 (PRI PDS) “Music Tags”
<http://www.mpx.cz/mp3manager/tags.htm>

24

What's Up So Far...

- PBCore v1.1 is Published January 24, 2007
- v1.0 was released April 2005
 - Version 1.1 of PBCore maintains:
 - 53 elements
 - arranged in hierarchies of 15 containers and 3 sub-containers
 - all organized under 4 content classes.

25



26



What's Up So Far...

- PBCore User Guide v1.1
 - Metadata Primer
 - Element Views
 - Element Documentation
 - Training Ops
 - Cataloging Tools
 - Case Examples

✓ QUICK JUMPS in USER GUIDE...

USER GUIDE

- Metadata Primer
- Element Views
 - (1) quick index view
 - (2) graphical view
 - (3) content class view
 - (4) alphabetical view
 - (5) mandatory elements list
 - (6) changes for PBCore v1.1
- Element Documentation
 - (1) full documentation for elements
 - (2) cheat sheet for elements
 - (3) attributes of elements explained
 - (4) metadata examples
 - (5) XML Schema Definition
 - (6) XML primers
 - (7) hierarchies & interdependencies
 - (8) mappings & crosswalks
 - (9) namespace
 - (10) glossary
- Training for PBCore
- Cataloging Tools
- Case Examples

BACKGROUND OF THE PROJECT

RESOURCES & LINKS

CPB ASSET MANAGEMENT WEBSITE

PBCORE HOME PAGE

27



Metadata Primer...

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

28



Element Views...

Views of the PBCore Elements

PBCore Element Views

Quick Index View
Graphical View
Content Class View
Alphabetical View
Mandatory Elements List
Full Documentation for each Element
Cheat Sheet for Using the Elements
Hierarchies & Element Interdependencies
Metadata Examples
Changes for PBCore v1.1

29



Element Documentation...

Documentation for PBCore Elements

For full documentation of PBCore elements, on an element-by-element based, select an entry from the index of metadata descriptors on the left-hand side of this page.

Listed below are other sections of the PBCore User Guide that document the elements.

➤ **Element Documentation**

Of interest to information specialists is the Full Documentation for the PBCore elements. If you are looking for a one-page quick reference for the definitions of each element and some data examples, use the Cheat Sheet. Other documentation is provided to round out the understanding of PBCore's metadata elements.

- **Full Documentation**
 - **Cheat Sheet for Elements**
 - **Attributes of Elements Explained**
 - **Content Classes Described**
 - **Element Containers Described**
 - **Metadata Examples**
 - **XML Schema Definition (XSD)**
 - **XML Primers**
 - **Hierarchies & Element Interdependencies**
 - **Mappings & Crosswalks**
 - **Namespace for PBCore**
 - **Glossary of Metadata Terms**
-

30

Cataloging Tool(s)...

TOOLS FOR DESCRIBING MEDIA ITEMS USING PBCore :



FileMaker Pro Database (posting soon)



Microsoft XCEL Spreadsheet (posting soon)



Microsoft Access Database (posting later)



PBCore XML Schema (and XML Primer)



Mappings of PBCore to other Metadata Standards

31

Can I see a PBCore cataloging application?

- CWIS (Collection Workflow Integration System) at Wisconsin PTV
<http://wptmedialibrary.wpt.org/>

Add New Resource

A "*" next to a label indicates a required field.

Tape Number:

Title: *

Title Type: Program or Series Title:

Alternative
Collection
Episode

Here and Now
Hometown Stories
In Wisconsin
Korean War Stories

Project Number:

[Search/Add](#) Creator(s):

[Search/Add](#) Contributor(s):

Date Created:

Broadcast Date:

[Search/Add](#) Publisher(s):

Subject(s): *

32

Can I see a PBCore cataloging application?

- CWIS (Collection Workflow Integration System) at Wisconsin PTV <http://wptmedialibrary.wpt.org/>

Rights Summary:

Audience Level: College, Educator, Female, General

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

Format Standard: Audio/MPEG, Audio/SDDS, 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 End:

33

Can I see a PBCore cataloging application?

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

media exchange PROVIDER

Find All My Settings Logout New Document

Quick Search: Find Search In Current Results

Search

Advanced Form TQC Content

Use Tree: Provider

KET - Kentucky Network

KUED-TV through University of Utah

PBS

University of Utah Media Solutions

University of Utah University of Utah Press

Utah State Historical Society

Search In Current Results

Back To Results 10 of 30

Info History

Apply Template: Select Go

Title: Glen Canyon: A Dam, Water

This title is a(n): Program

Comes from a collection named:

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

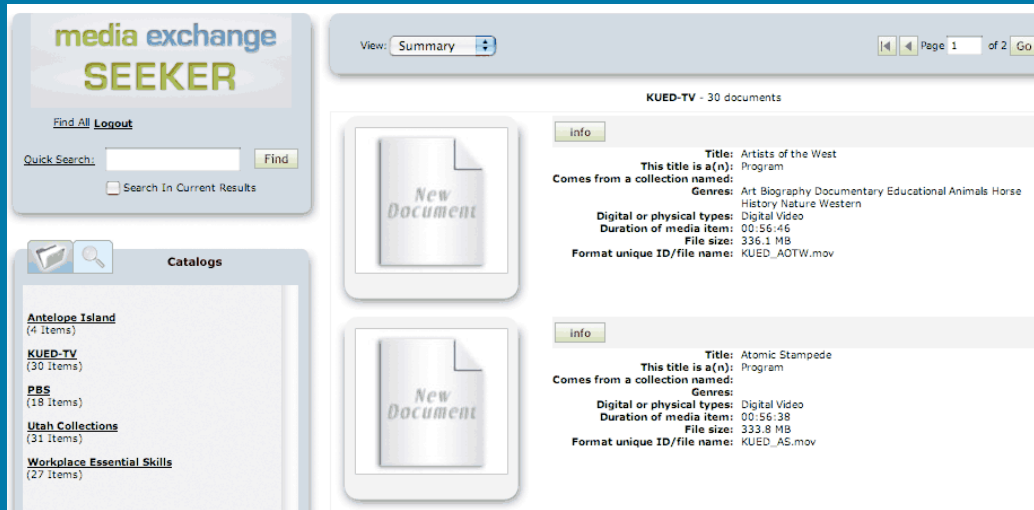
SUBJECTS...

Save

34

Can I see a PBCore cataloging application?

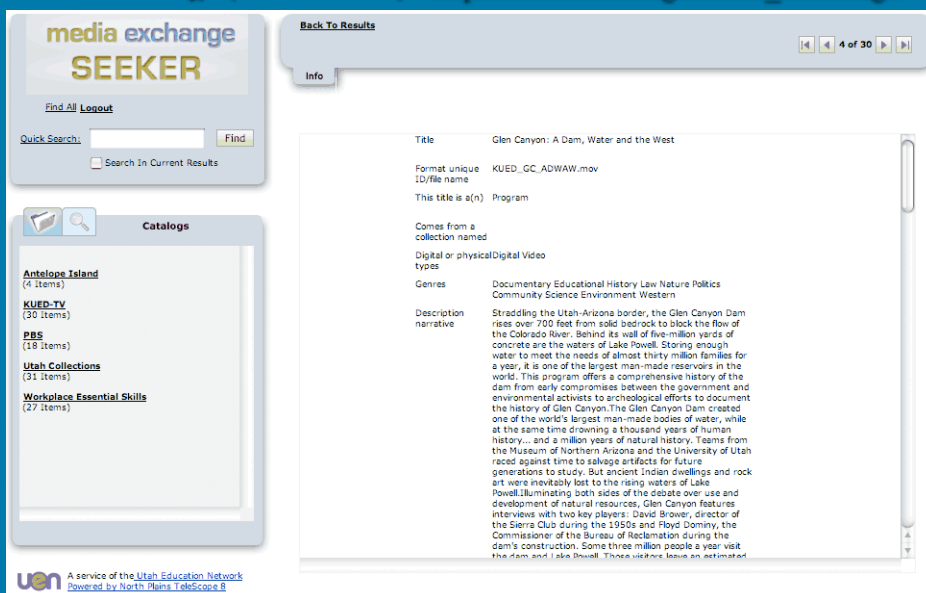
- Media Exchange (NETA/UEN) http://www.uen.org/media_exchange



35

Can I see a PBCore cataloging application?

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



36



Mappings & Crosswalks...

MAPPINGS FROM PBCORE TO OTHER METADATA STANDARDS:

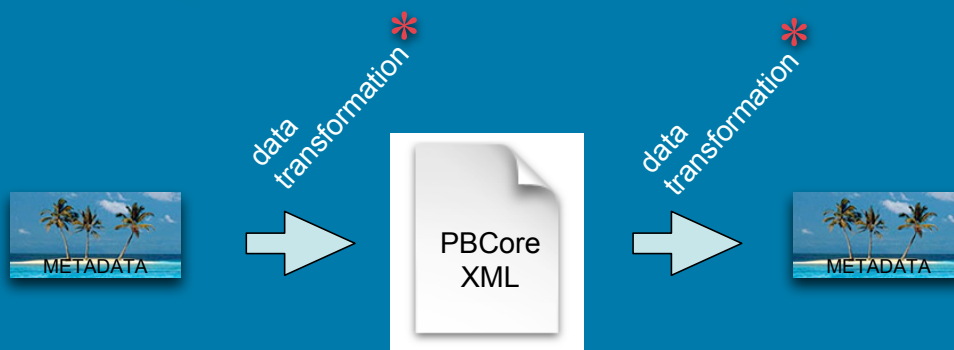
PBCore v1.1	↔	Dublin Core & Qualified Dublin Core (DC Terms) (posting soon)
PBCore v1.1	↔	PBS PODS: Public Broadcasting Service's Program Offer Data Service (posting soon)
PBCore v1.1	↔	Content Depot PSD: Public Radio Satellite Service and the Program Service Data (posting soon)
PBCore v1.1	↔	SMPTE DMS-1 (MXF) (posting later)
PBCore v1.1	↔	Moving Image Collections (MIC) (posting later)
PBCore v1.1	↔	SCORM (posting later)
PBCore v1.1	↔	MPEG-7 (posting later)
PBCore v1.1	↔	MODS (posting later)
PBCore v1.1	↔	METS (posting later)
PBCore v1.1	↔	ARIADNE (posting later)
PBCore v1.1	↔	USMARC (posting later)
PBCore v1.1	↔	IEEE-LOM (posting later)
PBCore v1.1	↔	IMS (posting later)
PBCore v1.1	↔	ViDe (posting later)
PBCore v1.1	↔	MPEG-21 (posting later)

37



XSD Data Exchanges...

Documents the dictionary definitions
in a specific framework for data sharing...



PBCore XML document
(XSD-XML Schema Definition)

* Not trivial

38



What is PBCore Compliant?

- Be able to interpret the PBCore XSD
 - When sharing data...
...parse, transform, then export your metadata into the PBCore XSD
 - When ingesting shared data...
...parse, transform, then import the PBCore XSD compliant metadata into your structures.



What is PBCore Compliant?

- When setting up a metadata structure, follow PBCore Dictionary attributes
 - Elements have matching meanings, even if their names vary (author vs creator)
 - 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.



Case Examples...

CASE EXAMPLES OF PBCORE AS USED IN REAL WORLD APPLICATIONS :



Case Example No. 1 (posting by early summer 2007)



Case Example No. 2 (posting by early summer 2007)



Case Example No. 3 (posting by early summer 2007)



Case Example No. 4 (posting by early summer 2007)

41



Training...

- **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	PowerPoint-zip PDF QuickTime Video-zip QuickTime/iPod Video-zip MP3 audio-zip
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.	Wednesday 2007-02-28/2:00pm EST Repeated Thursday 2007-03-01/2:00pm EST	PowerPoint-zip PDF QuickTime Video-zip QuickTime/iPod Video-zip MP3 audio-zip
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	PowerPoint-zip PDF QuickTime Video-zip QuickTime/iPod Video-zip MP3 audio-zip
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	PowerPoint-zip PDF QuickTime Video-zip QuickTime/iPod Video-zip MP3 audio-zip

42



Training...

- To *register* for a session:•
 - go to <http://pbcore.webex.com>•
 - click on the "Upcoming" tab•
 - enter your registration information•
 - you will be prompted to install the **WebEx Java applet**; install any time prior to your session (not platform specific)•
 - contact Geoff Freed (geoff_freed@wgbh.org) if you have questions

43



Listserv...



- For questions, answers, opinions & announcements
- To *join* :
 - <http://www.pbcore.org>
 - *Click on button for CONTACTS*

44



Listserv...



- small and large public TV and radio stations
- regional program producers & distributors
- national program producers & distributors
- educational institutions/libraries/librarians
- cable operators
- non-profit organizations
- vendors

45



Community of Practice (CoP)

- refers to the process of social learning that occurs when people...
 - who have a common interest in some subject or problem...
 - collaborate over an extended period...
 - to share ideas...
 - to find solutions...
 - to build innovations

Wikipedia article
http://en.wikipedia.org/wiki/Community_of_Practice

46



Community of Practice (CoP)

- Encourage adoption of PBCore
- Share experiences about implementations
- Clarify current definitions and applications of PBCore v1.1 elements
- Connect with broader communities outside of public broadcasting and the USA
- Connect with individuals of differing professional backgrounds and expertise
- Broaden our metadata horizons
- Broaden our digital asset management horizons
- Become engaged and involved

47



Community of Practice (CoP)

- How does your participating in the CoP impact PBCore “beyond its funding phase”
- Active funding from CPB ends this year...
 - what are the funding needs into the future
 - what funding sources within the public broadcasting system and outside are possible

48



Community of Practice (CoP)

- How does your participating in the CoP impact PBCore “beyond its funding phase”
- The current leadership team will no longer be under contract
 - who will continue advocacy
 - who will continue coordination
 - who will continue communications
 - who will continue training
 - who will play match-maker between PBCore and possible implementations

49



Community of Practice (CoP)

- How does your participating in the CoP impact PBCore “beyond its funding phase”
- www.pbcore.org is the PBCore Port-of-Call
 - how to maintain its presence
 - how to continue with announcements, RSS feeds, and conference presentation postings
 - how to manage and communicate change orders to the PBCore elements

50



Community of Practice (CoP)

- How does your participating in the CoP impact PBCore “beyond its funding phase”
 - How will the PBCore ship be steered?
 - should PBCore be folded into the efforts of existing metadata and asset management initiatives and projects
 - should PBCore retain its status as a stand-alone entity

51



Community of Practice (CoP)

- How does participating in the CoP impact PBCore “beyond its funding phase”
 - As adopters and enablers of PBCore...
 - think beyond your own needs and specific implementations
 - consider how to advocate and encourage others to embrace PBCore
 - contemplate the long-term sustainability of PBCore

52



Community of Practice (CoP)

- How does participating in the CoP impact PBCore “beyond its funding phase”
 - As adopters and enablers of PBCore...
 - what recommendations can we make to CPB to keep the PBCore ship navigating the waters of metadata and information management
 - what guidelines can we make to ensure its continued viability

53



Community of Practice (CoP)

- How does participating in the CoP impact PBCore “beyond its funding phase”
 - Our Call to Action...
 - is to remain engaged with PBCore
 - is to engage others
 - is to broaden the use of PBCore
 - is to sustain PBCore

54



Listserv...



- Our Call to Action...
 - is to join the PBCore Listserv
 - is to submit questions, answers, opinions, recommendations, and guidelines
 - www.pbcore.org ➔ Click on *CONTACTS*

55

Resources

56

Resources

General:

- [PBCore.org](http://www.pbcore.org/)
- (<http://www.pbcore.org/>)
- The [PBCore listserv](http://www.pbcore.org/contacts/index.html)
- (<http://www.pbcore.org/contacts/index.html>)
- The [PBCore project](http://ncam.wgbh.org/pbcore) at NCAM
(<http://ncam.wgbh.org/pbcore>)
- The [PBCore XML Schema](http://www.pbcore.org/PBCore/PBCoreXMLSchema.html)
(draft: <http://www.pbcore.org/PBCore/PBCoreXMLSchema.html>)

57

Resources

Metadata

- [PBCore metadata resources](http://www.pbcore.org/PBCore/PBCoreXMLSchema.html#XMLSchemaPrimers)
(<http://www.pbcore.org/PBCore/PBCoreXMLSchema.html#XMLSchemaPrimers>)
- ["The Case for Shared Metadata Standards"](http://www.current.org/tech/tech0209metadata.html)
(<http://www.current.org/tech/tech0209metadata.html>)
- [A Primer for Understanding PBCore](http://www.pbcore.org/PBCore/PBCorePrimer.html)
(<http://www.pbcore.org/PBCore/PBCorePrimer.html>)

58

Resources

Metadata search tools:

- Wisconsin Public TV's [Media Library Online](http://wptmedialibrary.wpt.org/) (<http://wptmedialibrary.wpt.org/>), a PBCore implementation through [CWIS](http://scout.wisc.edu/Projects/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/>)

Contact information

Marcia Brooks

Project Director, PBCore Project

marcia_brooks@wgbh.org