

Public Broadcasting Metadata Dictionary Project Phase Two Narrative Report (1 of 2)

Submitted by WGBH Educational Foundation, January 16, 2003

Topics in this report that relate to specific Phase 2 deliverables are indicated with an asterisk (*).

This report outlines activities conducted on behalf of the project during the period June 1 to September 30, 2002. The remainder of the project's Phase Two activities for the period of October 1 through November 22 are referenced in a second Phase 2 project narrative.*

The Working Group's activities were completed under the aegis of two committees: the Dictionary Review team and the User Requirements team. Each group achieved consensus to develop the tools, processes and timelines by which to carry out their work.

The work of each group was primarily conducted via e-mail and telephone conference calls. With much of the work transpiring over the summer months, it should be noted that the majority of Working Group members remained committed to their tasks, with several completing work on project activities and deliverables during their personal vacation time.

Recruitment of "Adjunct" Members of Public Broadcasting Metadata Working Group

CPB and WGBH worked together in June to identify additional disciplines and organizations to represent on the Working Group*, and to determine the best individual representatives to invite to participate. WGBH worked with CPB to draft the invitations, which were extended and accepted in July.

New adjunct working group members:

Nancy Baldacci/American Public Television, Sharon Blair/AMIA Local Television Task Force, Brian Callahan/WHRO, David Felland/WMVS, Tom Handy/KWSU, Steven Heard/Public Interactive, Ann Lootens-Kraus/WGBH, Chuck McConnell/NETA-OSBE, Lesley Norman/David Grubin Productions, Meg O'Hara/WNET, Michael Yoch/NPR Online, Art Zygielbaum/Nebraska Educational Television, Dave Johnston/PBS Online.

These new members participated in interviews completed by the User Requirements team to develop use case scenarios*.

Activities to Facilitate Communications for the Project

Project Communications Plan

WGBH initiated discussion with CPB about communication efforts on behalf of the project, and submitted its initial recommendations for a communications plan to CPB in August, 2002.

Presentations

PBS Annual Meeting

WGBH coordinated with PBS to arrange for the scheduling of a metadata project concurrent session.

Marcia Brooks worked with co-panelists Tim Olson, Marilyn Pierce, James Steinbach and Alison White to develop the concurrent session PowerPoint presentation titled "I Never Metadata I Didn't Like: The Promise of Digital Information Flow". This PowerPoint presentation* was developed to serve as a foundation for subsequent project presentations. The project one-sheet* was distributed as a handout at the session, and at the conference registration area.

The presentation was designed both to raise awareness of the project and the variety of system differences represented by the Working Group, and to provide an overview of Asset Management and metadata, how they are related, and the importance of each to public broadcasting stations and producers.

Roughly 20 people were in attendance, considered relatively good attendance at the 2002 Annual Meeting given the unusually high number of concurrent sessions. The audience was engaged and asked good questions, and several people stepped forward after to express their ongoing interest (Peter Kaufman from Intelligent Television, and Dr. Scott Aikens, Director, Digital Content Strategy Project for KQED).

WGBH Digital Asset Management Symposium

Alison White delivered a PowerPoint presentation* to a general session audience of approximately 180 attendees from public television and joint licensee stations, universities, the Boston Public Library, and WGBH's technology partners. The presentation focused on raising awareness of the project, including basics about what metadata is, how it's used in public broadcasting, the existing problems in its use and the project outcomes in addressing them. Major project milestones and timelines were included.

Communications with Interested Parties

WGBH held discussions with several individuals who expressed interest in the project on behalf of the organizations they represent. The common theme in these discussions centered around the lack of universal agreement about program titles at the series and episode levels. ProTrack and PubTV Online say these problems are compounded by multiple versions of a program (i.e. multiple program distributors, pledge specials, etc.)

As a matter of course, every interested party is invited to visit the project Web site.

Tracy Carter, ProTrack

WGBH initiated contact with Tracy Carter of Myers Information Systems, who expressed interest in how the metadata project will impact the restructuring of

program information, from the point it enters a station (automation interface, program guide generation, Web site updates, etc.) to beyond broadcast. Protrack's desire is to deliver information to station traffic/operations personnel even if they're not using a unique program ID.

Tracy mentioned that ProTrack is considering storing metadata for stations that aren't big production houses, but he doesn't see that potential service as an overlap or competitive one. Tracy indicated interested in sharing preliminary project recommendations.

Jon Marston – WBUR

Jon called WGBH regarding his work on a program schedule system for WBUR, built as a relational database and tied into WBUR's existing content management system. While Jon's comfortable with what he's built, he wants to do a reality check for the formats in which people are describing content, i.e. how to distinguish between name of a series and episode title. Jon's interested in knowing what's on the horizon for pulling metadata so it doesn't need to be entered by hand, and wants to be sure he interfaces with the appropriate systems and protocols.

Jon expressed interest in staying apprised of the project, and said he would check back.

Jak Schibley- PubTV Online

WGBH initiated contact with Jak by e-mail, to share information about the project, including the project one-sheet. In exchanges of e-mail, Jak indicated interest in querying a live database not only via an interface for live users, but programmatically as well. Many of Jak's requests centered around the premise that a live database would exist on an ongoing basis, which WGBH advised is beyond the scope of the project. Jak also indicated interest in addressing local programming as well as the smaller program distributors, and echoed ProTrack's concerns about duplicate records for a single show (multiple distributors, etc.) Jak hopes to see genre information at the episodic level because some series' content differs at the episodic level. He seeks consistent rules, i.e. to how to classify programs that fall between a special and a series, and programs with multiple versions.

Jak was advised that the project will ultimately help provide standards for public broadcasting entities that provide and exchange content, including title issues, once the authority files are completed. Jak was invited to visit the project Web site, was given the project timeline and major milestones, and was advised that the final project phase is anticipated to produce recommendations for the maintenance of the metadata dictionary beyond the scope of the project.

Project Web Site and Listserv

WGBH worked with KUED and its client CPB to support the August launch of the project Web site <<http://www.utah.edu/cpbmetadata>> and listserv. WGBH's efforts for the Web site included: providing feedback on content, information architecture, and permissions for access; and aggregating content for the site. Efforts for the listserv included driving decisions about approved participants. Efforts for both

included providing assistance in identifying conditions of use, and in drafting related communications to the Working Group.

Working Group Committee Members and Committee Tasks/Assignments

Dictionary Review Team

Alan Baker/Minnesota Public Radio, Marty Bloss/National Public Radio, Paul E. Burrows/KUED Media Solutions, Dave MacCarn/WGBH Educational Foundation, Marilyn Pierce, Cate Twohill/ Public Broadcasting Service. Team advisor: Efthimis Efthimiadis/University of Washington. Additional support provided by: Scott Bridgewater/NPR, Carrie Lowe/Public Broadcasting Service, and Thom Shepard/WGBH Educational Foundation. Team facilitator: Marcia Brooks.

The Dictionary Review Team began its work in early June 2002, and demonstrated a high level of motivation, productivity, and commitment to consensus throughout its active work in Phase 2.

The team began its work by reaching agreement on the format suggested by Alan Baker/MPR for each member to document and exchange their organization's internal metadata schemes/models. Upon completion, this resulted in a combined total of 467 metadata elements, which WGBH compiled into an Excel document, grouped by Dictionary Team members' contributing organizations.

The group's review revealed that approximately 90 percent of the aggregate fields overlapped. To better facilitate continued data manipulation, sharing and reporting, Paul Burrows/KUED Media Solutions imported the data into a FileMaker database,* which became both the de facto tool to facilitate and track the Dictionary's ongoing work, and one of several anticipated formats in which the Dictionary will be delivered. Paul developed a Grouping Tool to facilitate the identification of those metadata fields common to the representative organizations of the Dictionary Team.

Each Team member conformed, to the extent possible, their organization's metadata dictionaries to match the Dublin Core Metadata Element Sets, and, using the Grouping Tool, worked with stakeholders in their organizations to group similar fields, reducing the number of separate metadata elements from 467 to 249.

The Dictionary Review Team produced a report* to summarize its activities for presentation at the 2nd Working Group Meeting held September 12-13 in Arlington, VA.

User Requirements Team

Bea Morse/PBS, Robin Mudge (replaced in January 2003 by Michael Connet)/onCourse, Tim Olson/KCTS, Marilyn Pierce/PBS, Richard Ruotolo (replaced in September 2002 by Brent Trinacty)/Public Radio International, James Steinbach/WHA, Steven Vedro/Consultant, Tracy Vosburgh/WPSX. Team advisors:

Grace Agnew/AMIA and Judy Brown/Academic CoLab (SCORM). Team facilitator: Marcia Brooks.

The User Requirements Team began its work in mid-July 2002, and was tasked with identifying key user requirements for the metadata dictionary. To ascertain these needs, the team was assigned to conduct use case interviews* among the new members of the working group, to summarize use case scenarios, and to identify issues and recommendations to inform the Dictionary Team's work. The team then identified the specific metadata fields mentioned in the use case interviews.

In order to conduct this work, the team was asked to review a "strawman" list* of data sources, and internal and external constituents. The team was also asked to comment on sample use case formats that would facilitate the comparison and summary of information gathered.

The User Requirements Team recommended that the Dictionary implement a small, mandatory set of core metadata, but allow for modularity and special case extensions useful to other communities. The team also recommended that any metadata must be able to pass through existing transport and application standards already in place for public broadcasting dovetail with MPEG-7, Dublin Core, SMPTE and EBU P/Meta schemes, and that XML be used to "crosswalk" metadata from one application profile to another.

The User Requirements Team produced a report* to summarize its activities for presentation at the 2nd Working Group Meeting held September 12-13 in Arlington, VA. Following that meeting, work continued to develop the User Requirements matrix*, a document that maps the Dictionary fields to the discipline-specific needs identified in the use case interviews in order to identify any potential gaps in the Dictionary.

Committee and Adjunct Meeting Dates, Participants, Agenda and Outcomes

Working Group Meeting #2

The Working Group convened September 12-13 in Arlington, VA. WGBH began making the necessary arrangements in June, and continued preparations (for hotel accommodations and meeting space, meals, agenda* development in coordination with the project's meeting facilitator, etc.) until the meeting date.

Project participants in attendance at the meeting*:

Alan Baker/MPR, Marty Bloss/NPR, Marcia Brooks/WGBH, Judy Brown/Academic CoLab (SCORM), Paul Burrows/KUED Media Solutions, Efthimis Efthimiadis/University of Washington, Dennis Haarsager/Washington State University-CPB, Dave MacCarn/WGBH, Bea Morse/PBS, Tim Olson/KCTS, Marilyn Pierce/PBS, Lou Pugliese, Michael Connet/onCourse, Amy Rantanen/WGBH, Brent Trinacty/PRI, Cate Twohill/PBS, Steven Vedro/Consultant, Tracy Vosburgh/Penn State Public Broadcasting, and Alison White/CPB. In order to maximize the meeting's outcomes, it was decided to keep the meeting attendee size manageable by limiting Working Group members' attendance to the Core Working Group.

The meeting was facilitated and subsequently reported on* by Bill Keens/Wolf, Keens & Co. who also facilitated the first meeting of the Working Group in April 2002. With Bill's skill in facilitating the group, and with the Group's commitment to consensus building, the meeting participants were motivated, productive, and collegial.

Dave MacCarn presented an overview of standards activities and digital archives initiatives*, to help the Working Group frame its work. In order to apprise the Core Working Group as a whole, Tim Olson presented the User Requirements Team report* and Paul Burrows presented the Dictionary Team Report*.

During a working lunch, participants suggested priorities for breakout discussion, which were then prioritized into three topics tasked to three breakout groups:

Group One: Focus on a process for continued collapsing of the 249 metadata elements, illustrate the suggested approach, and identify issues for the full Working Group to address during the remainder of the meeting.

Group Two: Focus on the Dictionary's constituents, how to prioritize and accommodate key users, and to illustrate the suggested approach.

Group Three: How to ensure compliance, provide incentives and anticipate future needs of the to-be-published Dictionary -- beyond the life of the grant project.

Each group presented their deliberations* to the full Group. It was determined that the User Requirements Group would continue work to produce a more complete User Requirements matrix*; that the Dictionary Review Team would meet in October to continue collapsing the Dictionary fields (later referred to as the "Beantown Smackdown*"); and that an Education and Communications Committee would be formed from members of the entire working group, including the Adjunct members. A potential Compliance and Policy Group was discussed, with no formal decision made as to its formation.

The meeting concluded with discussion about ongoing refinement, requests for comments and testing of the Dictionary for the remainder of the grant project, with the decision for the entire Working Group to meet early in 2003.