Preservation and Access Systems using Phydo

AMS 2.0 Team

WGBH
Sadie Roosa
Andrew Myers
Jason Corum
Casey Davis Kaufman
Karen Cariani

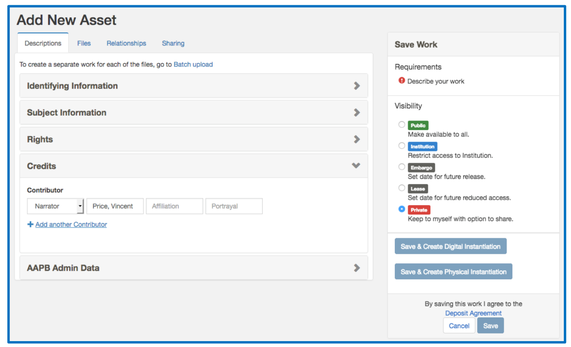
AVP
Kara Van Malssen
Adeel Ahmad

Indiana University
Jon Dunn
Jon Cameron
Chris Colvard
Maria Whitaker
Julie Hardesty

With funding from THE ANDREW W. MELLON FOUNDATION

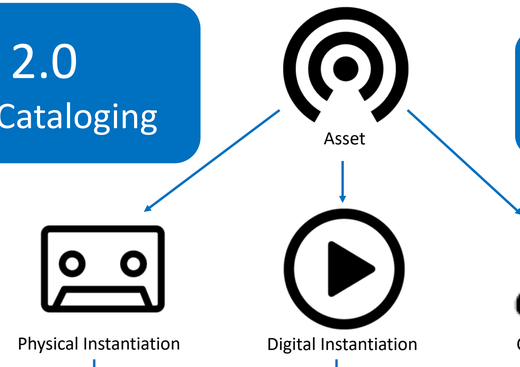
Capturing Descriptive and Technical Metadata

For the most part, entering data is done the same way as in other Hyrax-based applicated. However, to improve flow, we added buttons to allow catalogers to move easily from one work type to the next logical work type, starting with the parent Asset Type, easily adding children without having to jump through multiple screens.



For the Contribution work type, which can have a maximum for 4 fields and is often repeated many times per Asset, we thought jumping to another screen for each Contribution would be too much for cataloger. Instead, we embedded the Contribution form into the Asset form, to make it easier to add as many Contributions as are required.

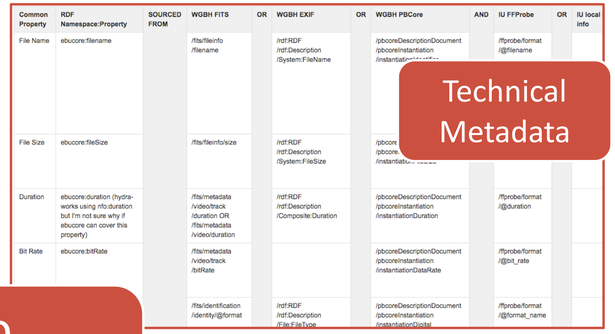
AMS 2.0
Access and Cataloging



Modeling PBCore in Hyrax and RDF

The majority of the predicates we use come from EBUCore, Dublin Core Elements, and Dublin Core Terms to represent PBCore-compliant metadata in RDF. We also used a few predicates from other ontologies, as well as a few local predicates. The entire data model for AMS 2.0 Release Candidate 1 can be found at tinyurl.com/ams2-datamodel

Using Hyrax work types, we modeled each of the main hierarchical components of PBCore records. These individual items are related to each other through parent/child relationships, which replicate the relationships implied by the nested nature of PBCore xml.



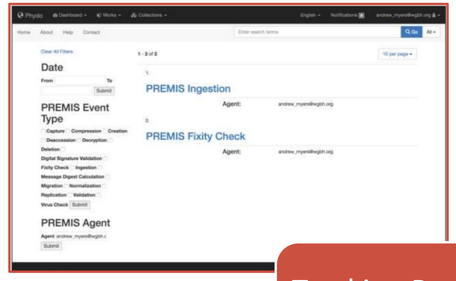
Technical Metadata

Phydo
Digital Preservation Metadata for A/V

Phydo is configurable for various technical metadata inputs. WGBH's configuration accepts PBCore instantiation xml and FITS xml, mapped to EBUCore, PREMIS and NFO RDF predicates.

Phydo was a joint development project between Indiana University and WGBH, which has been made possible in part by a major grant from the National Endowment for the Humanities: Exploring the human endeavor.

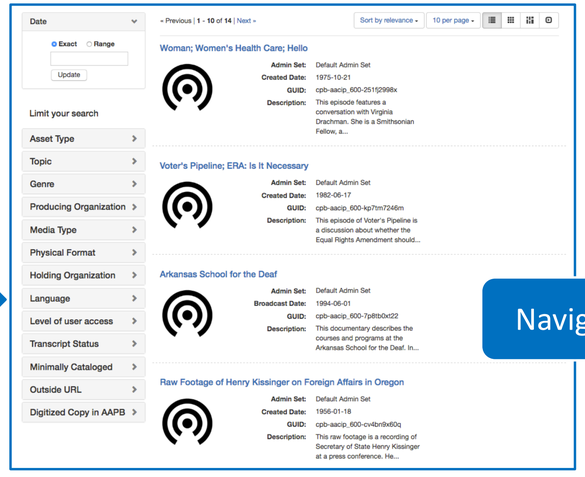
NATIONAL ENDOWMENT FOR THE Humanities



Phydo allows users to capture the results of many PREMIS Events (like transcoding and fixity checks) and run reports to plan for upcoming preservation needs.

Tracking Preservation with PREMIS Events

Phydo also uses the Hyrax Preservation gem (https://github.com/IUBLibTech/hyrax-preservation) which allows tracking, searching and sorting by PREMIS Event Type, Agent, and Date in a Hyrax application.

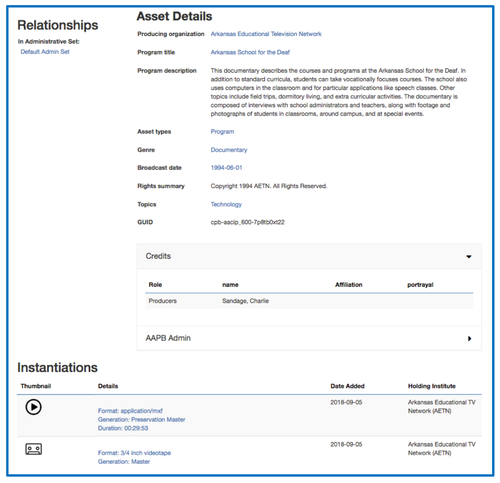


Because Instantiations and Essence Tracks can't exist without an Asset parent, we limited the search results to only Asset records. That way, anytime someone gets to a record of any work type, it's always through the context of the parent asset.

Navigating Your Records

In order to be able to search and facet on data in fields at both the Asset and child work levels, Solr indexes the children onto the parent record. This way, we can retrieve all of the results of Asset, which have a child instantiation in English or recorded on 3/4inch U-matic.

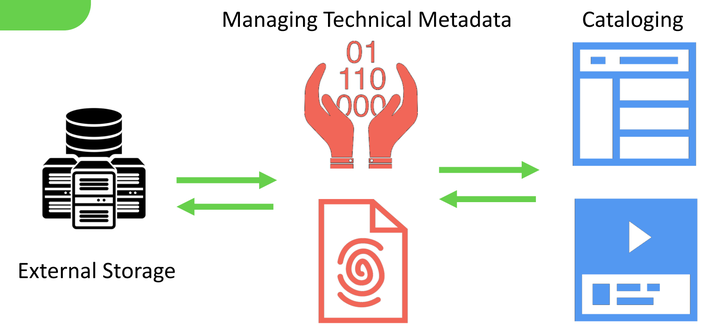
When a user decides which Asset they are interested in, they are taken to a details page for that Asset. To put all of the related works (Assets, Contributions, and Instantiations) in context, we show all of them in separate sections, with links to the fuller Instantiation records, if that's the metadata the user is the most interested in.



WGBH Integrated System
(to be built)

WGBH was recently awarded another grant from the National Endowment for the Humanities. Part of this funding will go towards implementing a version of AMS 2.0 specifically for WGBH's collection, and to integrate the AMS and Phydo applications, so that we can cohesively manage the descriptive, technical, and preservation metadata, while supporting both access and preservation workflows.

By keeping the applications separate but integrated, we will be able to tailor each application to the specific requirements of its users. Requiring one application to handle every part of the digital preservation and access lifecycle would create unnecessary challenges in development, data modeling, UX design, and migration.



Sadie Roosa
sadie_roosa@wgbh.org

WGBH Media Library & Archives