

**PBCCore, METS,**

**PREMIS, MODS, METSRights...**

**oh my!**

Kara van Malssen

Senior Research Fellow, NYU

Preserving Digital Public Television


AMIA 2008


A little bit about the **Preserving Digital**  
**Public Television** Project

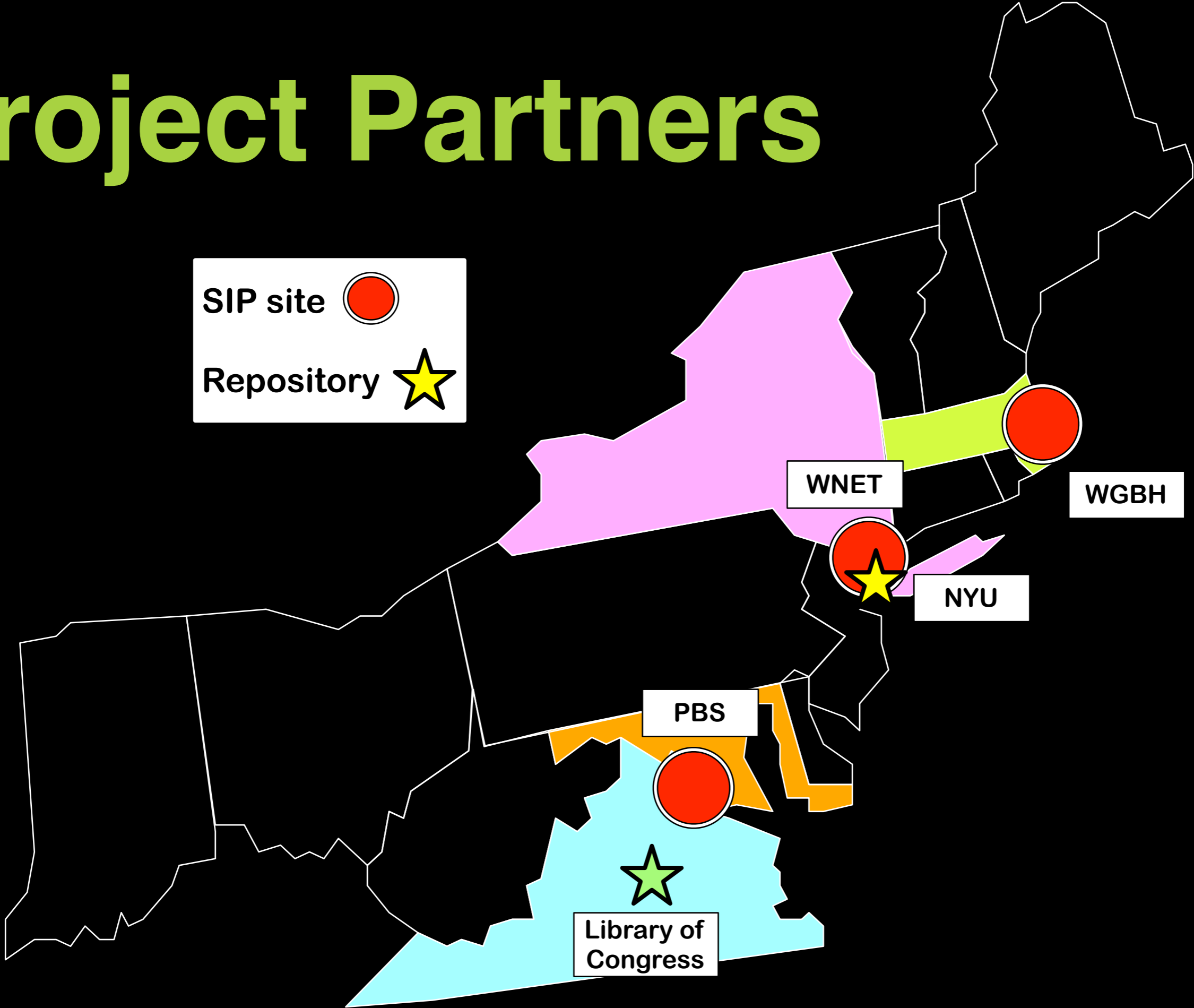
## GOALS:

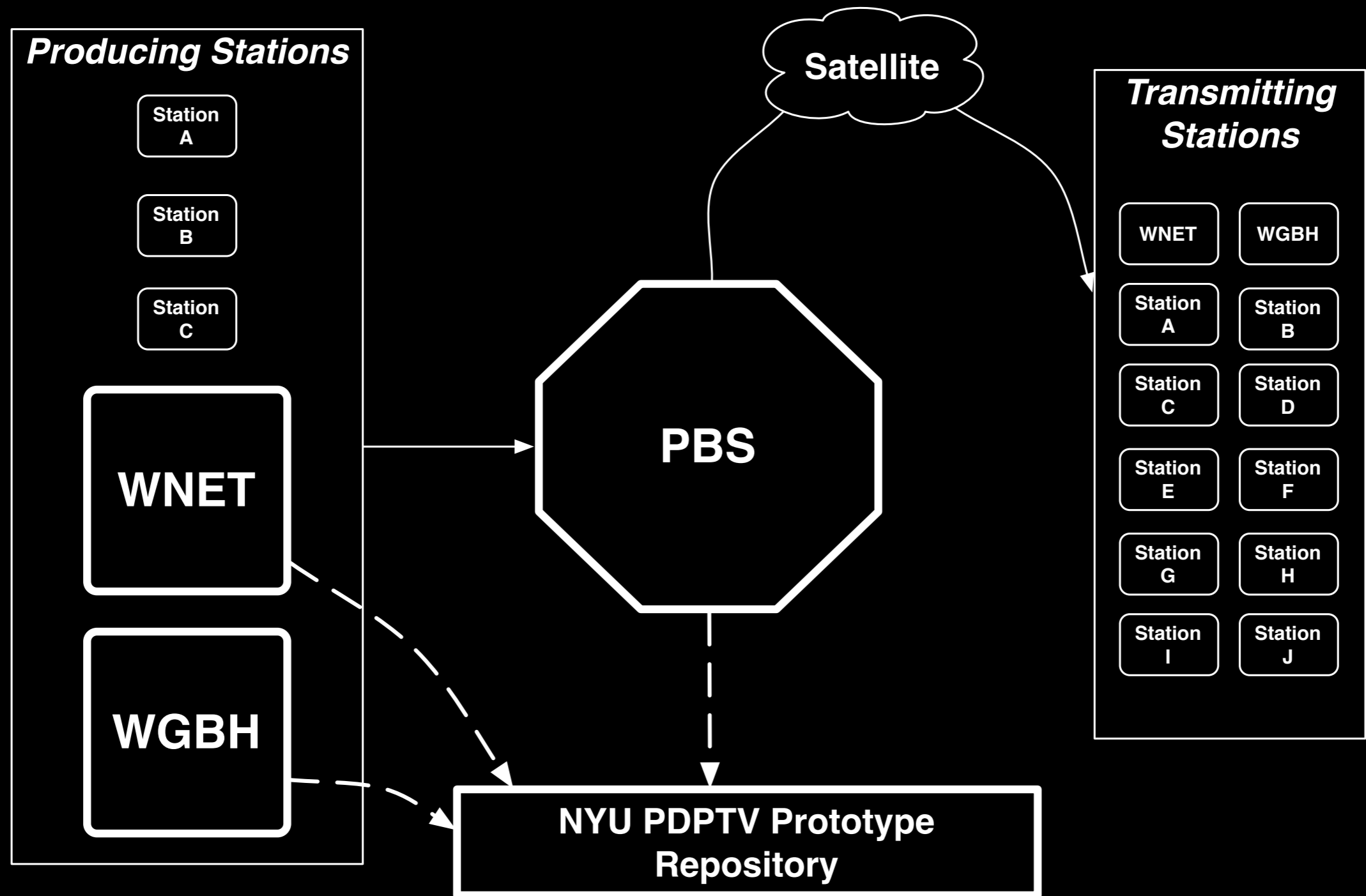
- Identify at-risk **born digital** public television content
- Build an **OAIS**-compliant prototype repository
- Explore and apply **standards**
- Create **selection** guidelines
- Research **sustainability** models, **copyright** encumbrances

# Project Partners

SIP site 

Repository 





# Submission Workflow

# NYU Goals:

- Create a **prototype repository** for long term retention
- **Aggregate** content from partner stations + PBS for sample programs
- Populate records with **metadata that already exists** (in station databases, files, scheduling systems, etc)
- **Transform** data and **package** content, while preserving relationships between items

# Important Vocabulary

- **The Repository**: NYU prototype preservation repository
- **OAIS**: Open Archival Information System
- **SIP**: Submission Information Package
- **AIP**: Archival Information Package

**OAIS  
Terms!**



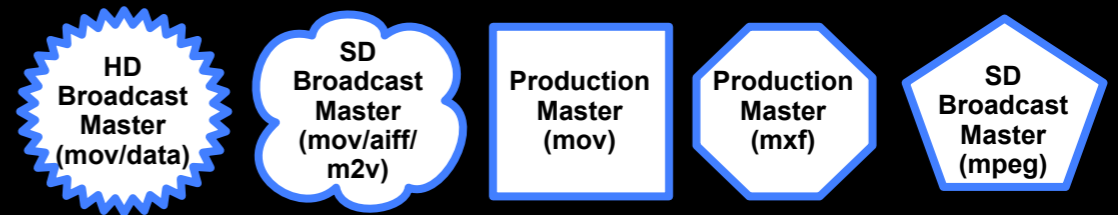
# Applying **standards**

- Normalize disparate metadata
- XML based
- One uniform scheme
- Easier to manage over the long term
- Rules, vocabularies, schemas help maintain consistency

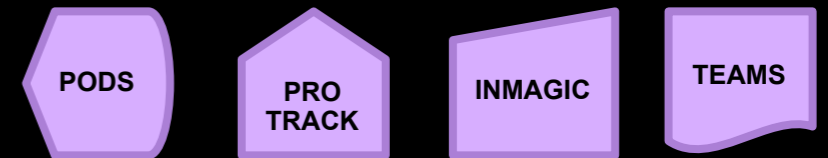
# Challenge of managing diverse SIPs:

# SIPs:

## ESSENCE FILE TYPES



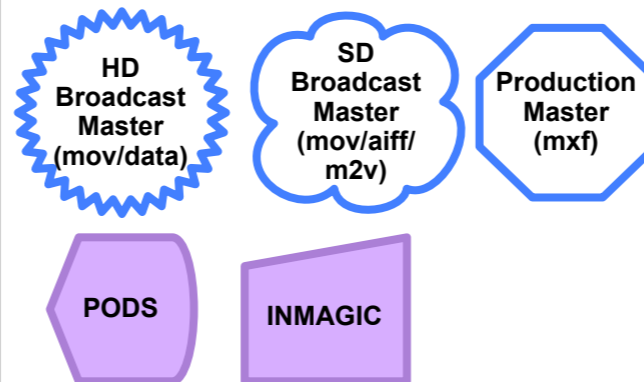
## DATABASE EXPORTS



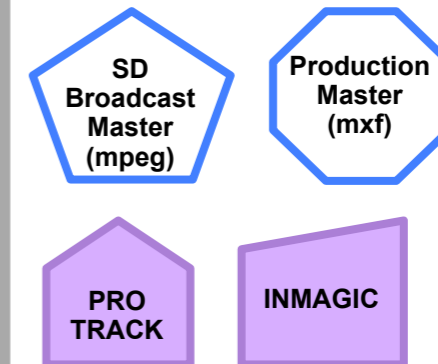
## ADDITIONAL ITEMS



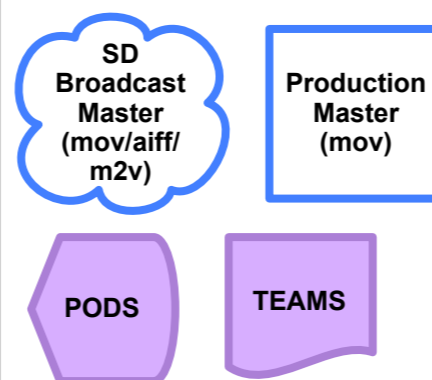
### SIP Class 1: WNET National Broadcast (Nature)



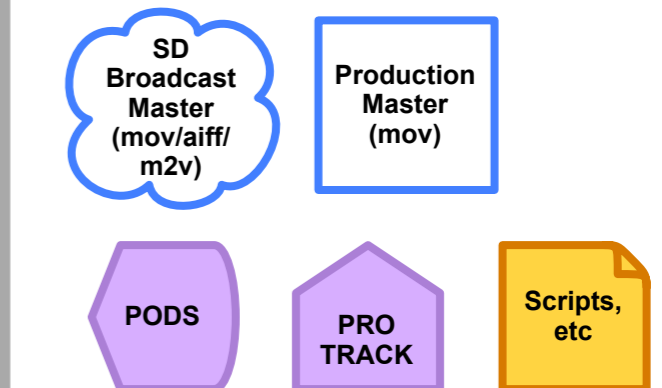
### SIP Class 3: WNET Local Broadcast (New York Voices)



### SIP Class 2: WGBH National Broadcasts



### SIP Class 4: Religion and Ethics





# PDPTV metadata model



**METS: Metadata Encoding and Transmission Standard**

Structural and administrative

+



**PBCore: Public Broadcasting Metadata Dictionary**

Descriptive and technical

+

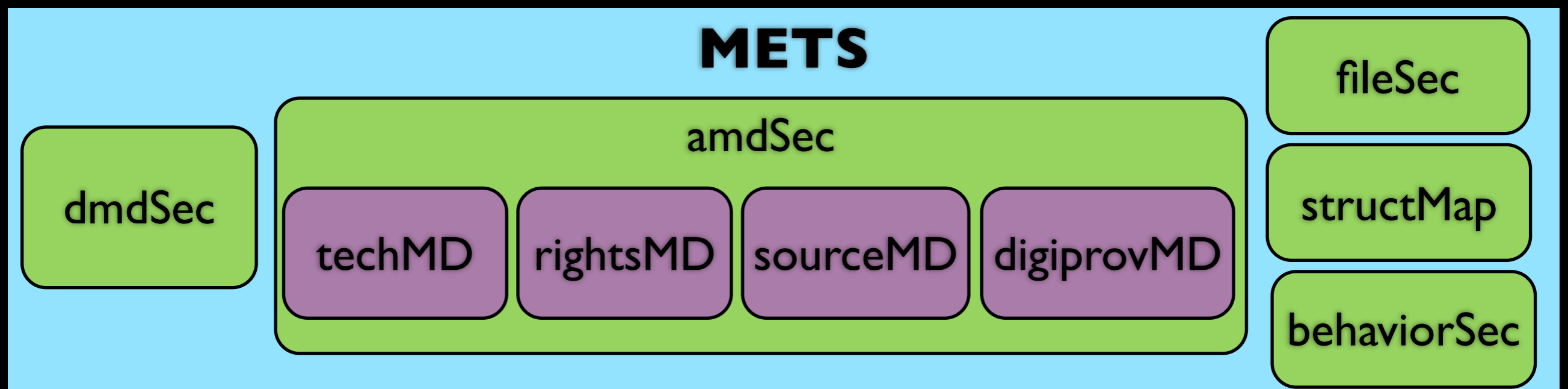


**PREMIS: Preservation Metadata Implementation Strategy**

Technical preservation metadata

# METS: Metadata Encoding and Transmission Standard

- Provides a structure to bundle all content (essence + metadata) in one AIP
- Identifies types of metadata, but not the terms to define them (with a few exceptions)



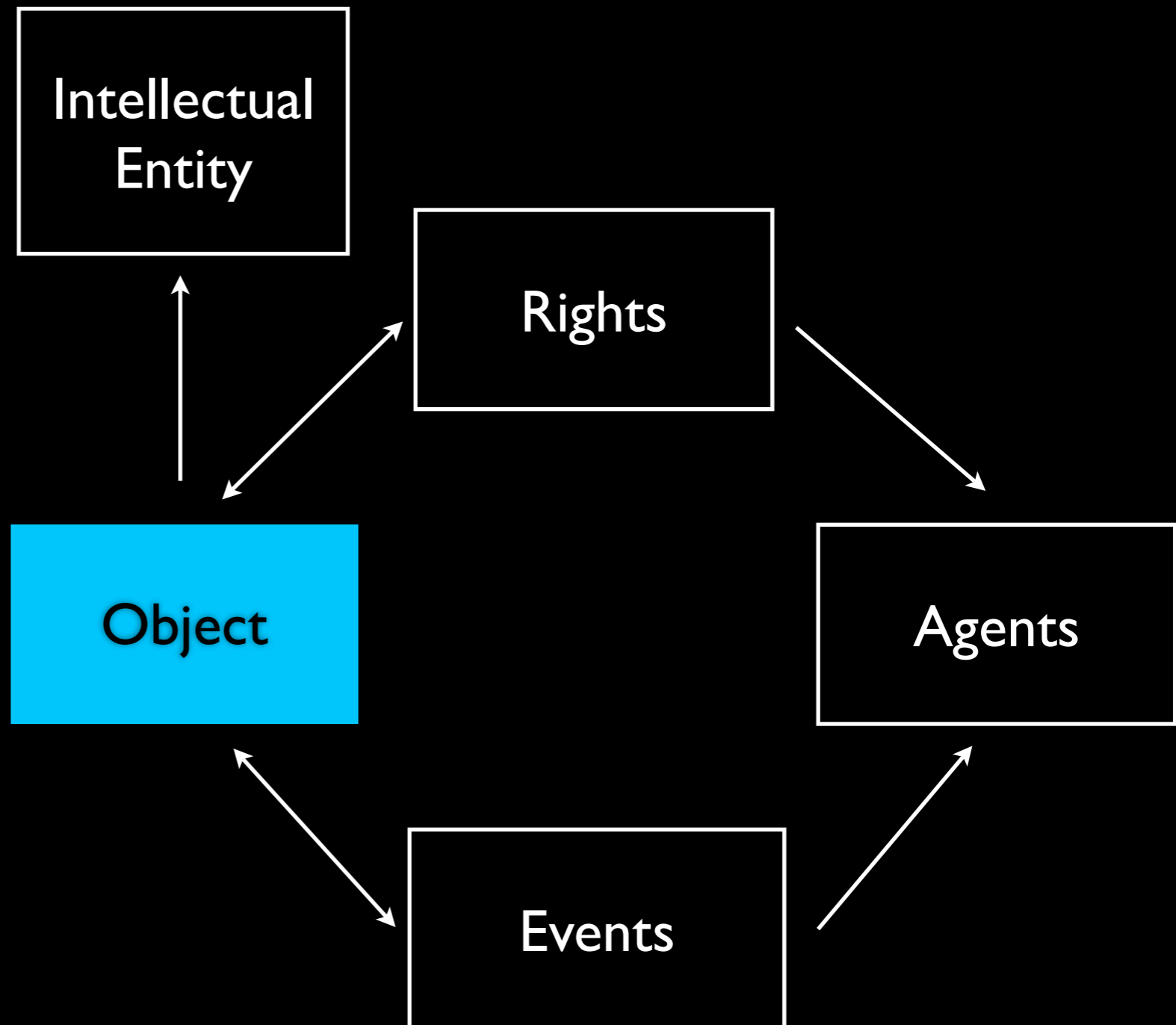
# PBCore: What is it good for?

- Descriptive metadata elements that are specific to public broadcasting
- Controlled vocabularies with broadcast terms
- Easy to map to from legacy station databases
- Granular technical metadata (PBCore 1.2)
  - ➔ Accurately represents the file specific metadata
  - ➔ Can be auto populated using technical metadata extraction tools & stylesheets

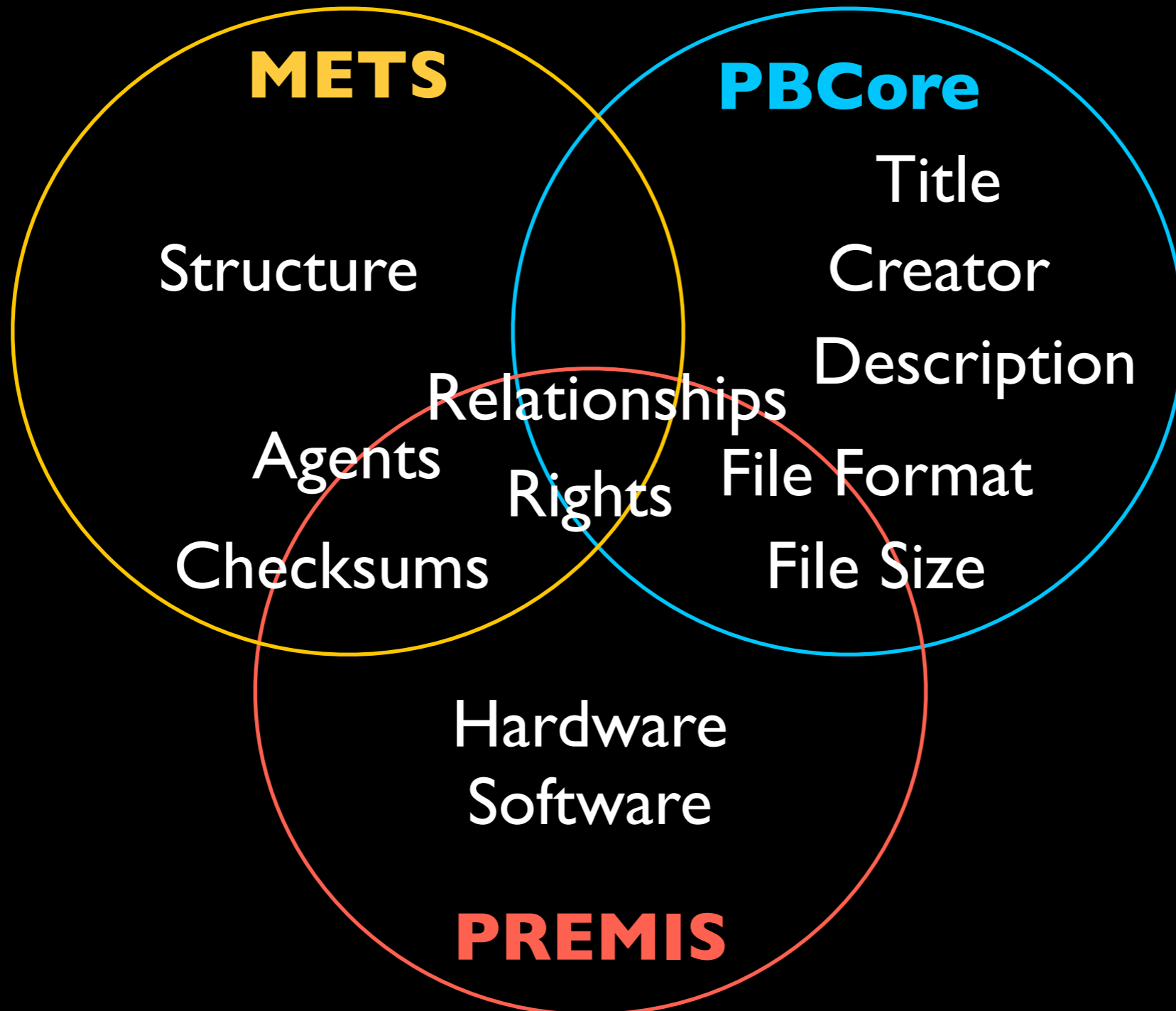
# PREMIS: Preservation Metadata Implementation Strategy

## Object Entity:

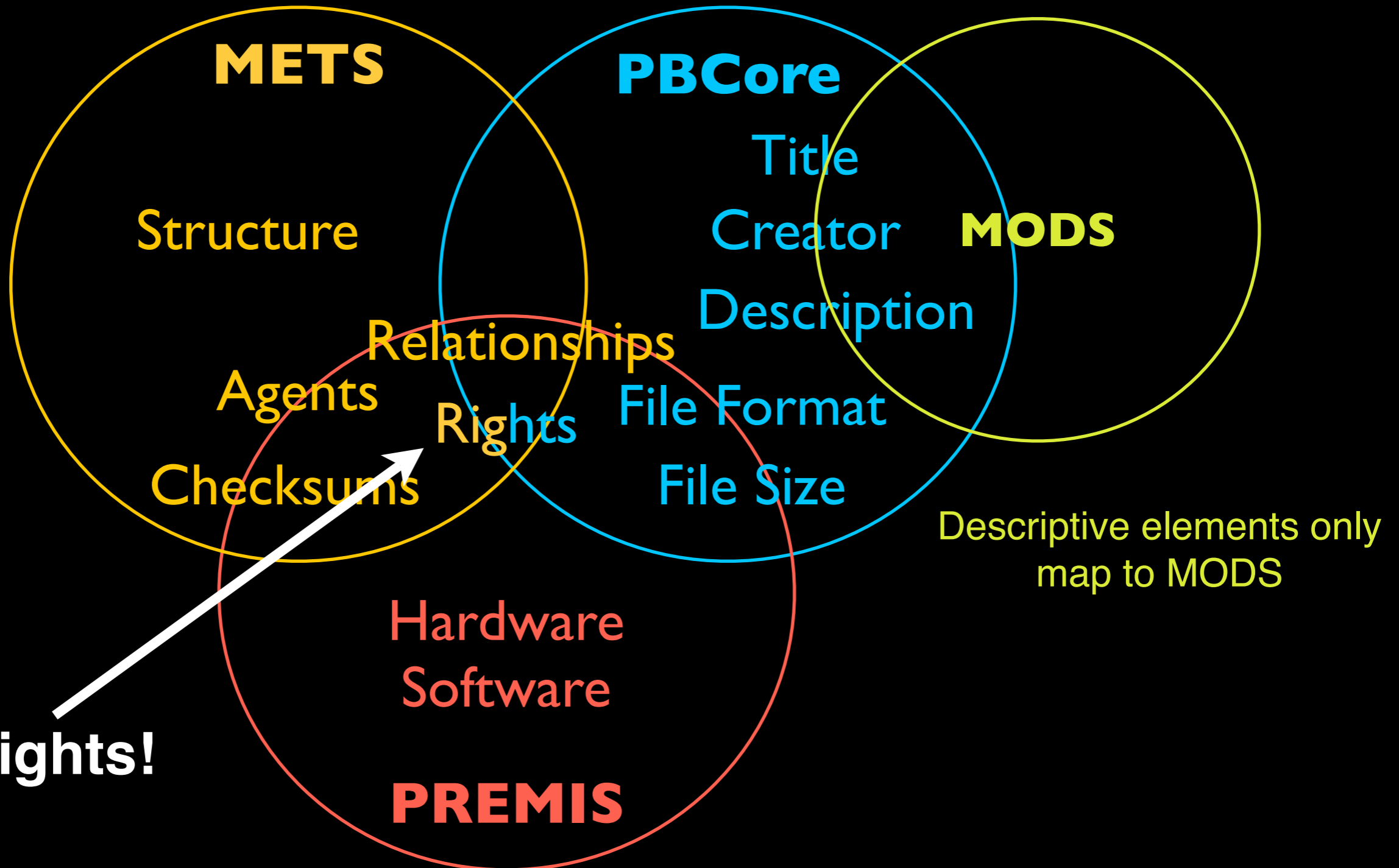
- Creating application info
- Playback environment (hardware and software)



# Issue of **Redundancy** between standards



# Putting it all **together**



# METS

dmdSec

amdSec

techMD

rightsMD

sourceMD

digiprovMD

fileSec

structMap

behaviorSec

# METS

**dmdSec**



amdSec

**techMD**



**rightsMD**



fileSec

structMap

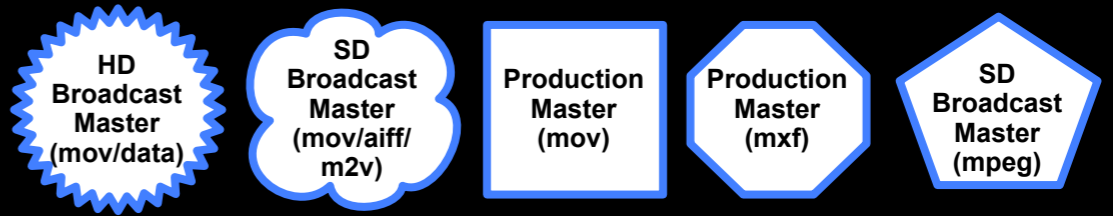


# AIP creation simplified

1. Content submitted, verified
2. METS automatically generated (checksums into METS attributes)
3. Source database exports automatically converted to PBCore
4. Technical metadata extracted from files using MediaInfo, converted to PBCore
5. MODS created from completed PBCore
6. Rights metadata (METSRights), preservation metadata (PREMIS) created
7. AIP complete

# AIPs:

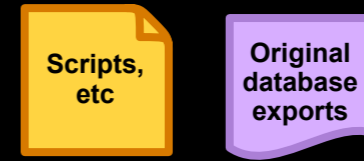
## ESSENCE FILE TYPES



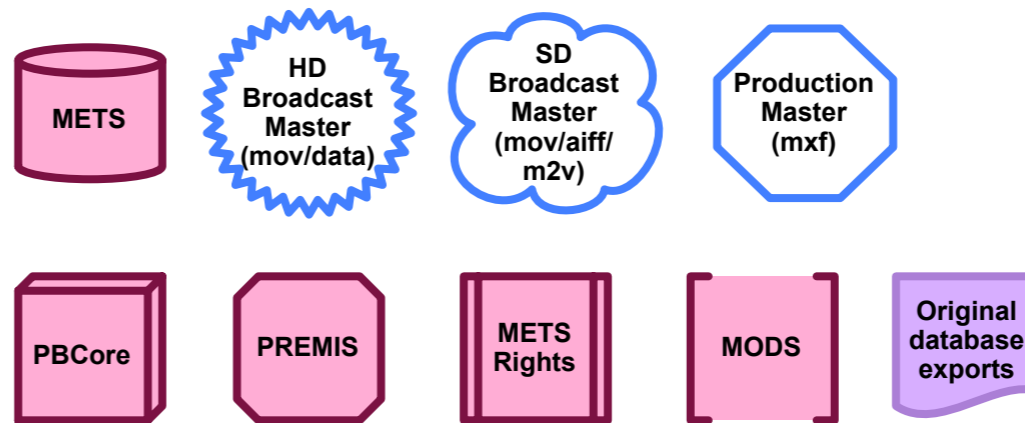
## METADATA



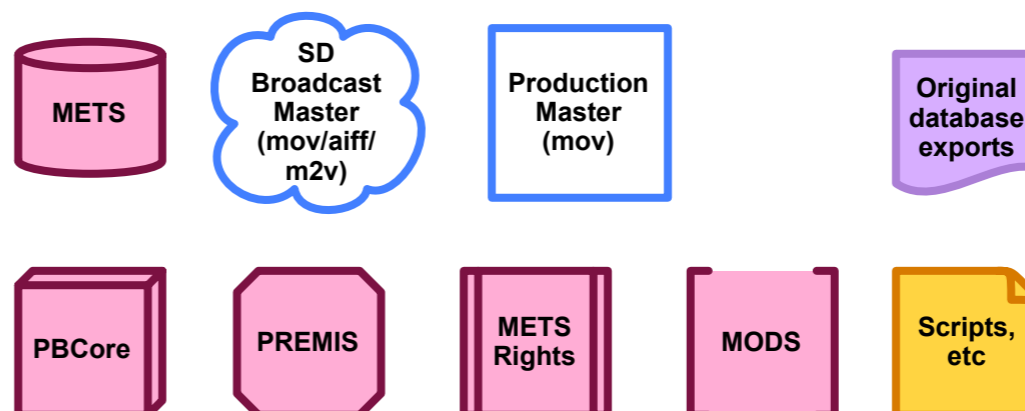
## ADDITIONAL ITEMS



### AIP Class 1: Nationally distributed content (Nature)



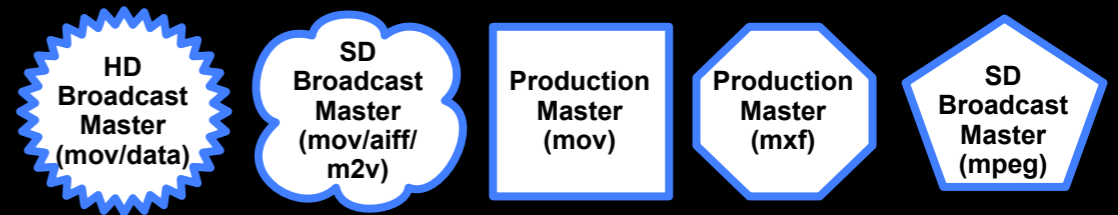
### AIP Class 4: Religion and Ethics



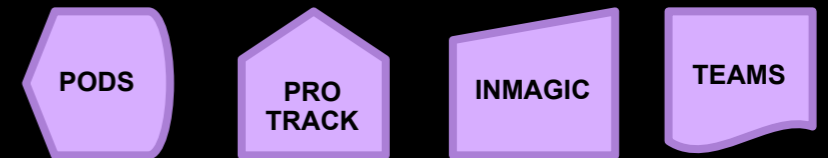
# Challenge of managing diverse SIPs:

# SIPs:

## ESSENCE FILE TYPES



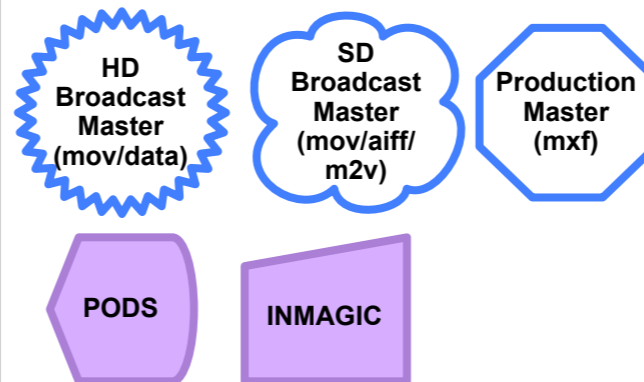
## DATABASE EXPORTS



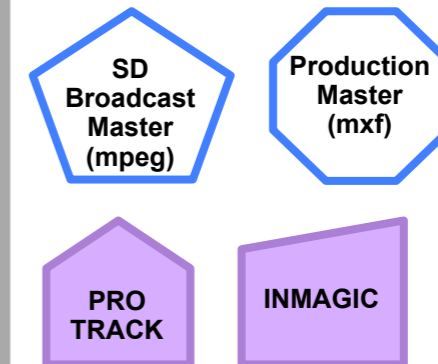
## ADDITIONAL ITEMS



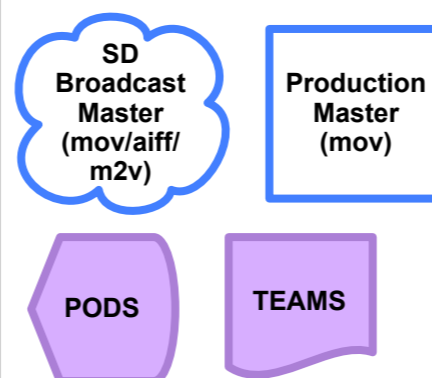
### SIP Class 1: WNET National Broadcast (Nature)



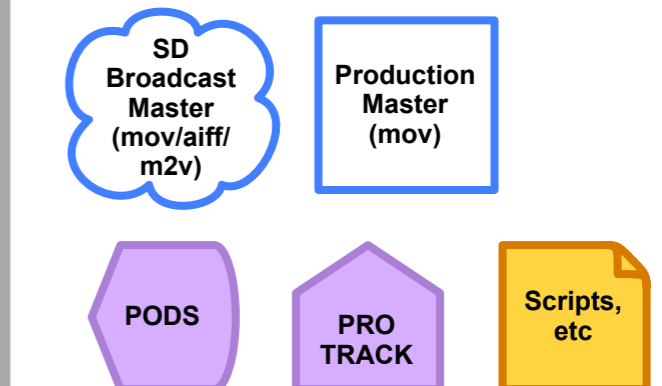
### SIP Class 3: WNET Local Broadcast (New York Voices)



### SIP Class 2: WGBH National Broadcasts



### SIP Class 4: Religion and Ethics



# Thank You!

[kara.vanmalssen@nyu.edu](mailto:kara.vanmalssen@nyu.edu)

[www.ptvdigitalarchive.org](http://www.ptvdigitalarchive.org)