Preserving Digital Public Television: Content Transfer Report

NDIIPP Partner’s Meeting
Content Transfer Panel
Summer 2008

Joseph G. Pawletko
New York University Libraries
PDPTV Overview:

- Preserving Digital Public Television (PDPTV)
  - preserving 20+ hours of ‘born digital’ Public Television content in prototype repository
  - many working groups breaking new ground

- Multiple Sites involved in PDPTV
  - Library of Congress
  - WNET
  - WGBH
  - PBS
  - NYU
Terminology: you probably know these already...

- **SIPs:**
  - Submission Information Packages

- **AIPs:**
  - Archival Information Packages

- **DIPs:**
  - Dissemination Information Packages
Infrastructure and Data Flow: aggregation phase

send Metadata

- WNET
- WGBH
- PBS

NYU Repository

Library of Congress

SRB (push)
Infrastructure and Data Flow: *dissemination phase*

- WNET
- WGBH
- PBS

pull DIPs

NYU Repository

Library of Congress

rsync + SSH (pull)
LC DIP Generation:

- **Generated ‘Bags’ for LC:**
  - DIPs conform to the BagIt specification
  - Wrote scripts that operate on SRB ‘collections’ to generate Bags
  - Generated both SHA1 and MD5 manifests
  - Leveraged SRB Scommands in conjunction with *NIX symlinks to create Bags that point into SRB Vault
  - Froze SRB collections during LC Bag pull
PDPTV Data Transferred to LC:

- 24 hours of content, 31 bags:
  - One Frontline episode
  - One American Experience episode
  - One Antiques Roadshow episode
  - Twelve Nature episodes
  - Sixteen New York Voices episodes

- > 900 GB of data

- Average of > 29 GB per Bag
Next Content Transfer Steps:

- **Finish upgrading SRB infrastructure:**
  - upgrading to SRB 3.5.0 at all nodes

- **Prepare and transfer additional programs and all metadata to LC**
Acknowledgements:

- **Library of Congress**
  - Carl Fleischhauer, Jane Mandelbaum, Brian Vargas

- **New York University**
  - Dean Carol Mandel, Dr. David Ackerman, Professor Howard Besser, Kara van Malssen, Rasan Rasch, Grace Yang, Unni Pillai

- **Thirteen/WNET**
  - Ken Devine, Nan Rubin, Jonathan Marmor, Ryan Weston, Winter Shanck

- **WGBH**
  - Dave MacCarn, Leah Weisse

- **PBS**
  - Bea Morse, Glenn Clatworthy, Wendy Allen, Irene Taylor, Michael Hunt
Wrap up…

Questions?

Thank you for your time!

Reference:

BagIt Specification: