




Open Source Digital Preservation Workflows

Gather up the bits, standardize the
AIPS, and preserve them indefinitely?



Nathan Tallman
iPRES 2015, UNC Chapel Hill
November 6, 2015
<https://goo.gl/sPmbR6>

Basic Digital Preservation Workflow @ UC

Gather the bits

- DSpace
- LUNA
- HTML/CSS GUI
- Hydra/Fedora
- Filesystems
- BitCurator

Make the AIPS

- Archivematica
- Bash scripts
- XSLT
- ImageMagick
- PDFtk
- FFMPEG
- and more

Preservation

- Fedora
- Geo-spatially redundant storage
- Fixity checks
- Forward migrations

Ignore the man behind the
curtain, he has problems



Gather the Bits

- Exports all look different
- Sometimes little or no metadata
- Base64 encoded binary bitstreams
- Rights information
- Restricted bitstreams



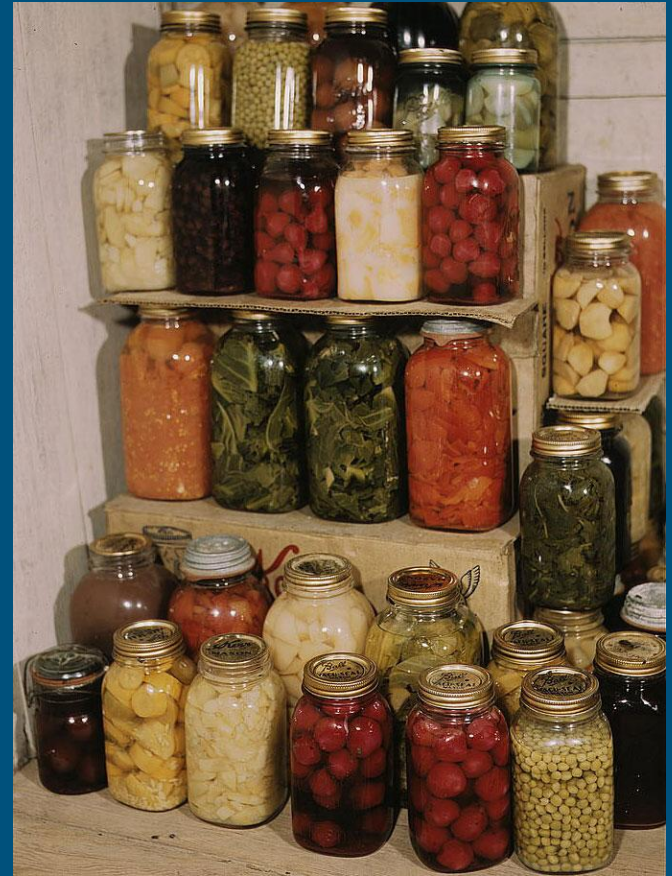
Make the AIPS



- Boutique scripts don't scale
- Derivatives (include?)
 - Normalization
- Versioning of content or metadata

Preservation

- Local copies
- Remote copies
- Tie-breakers
- Synchronization



Some of the major concerns

How many copies



Metadata changes



Logistics



It's Complicated



Possible Partial Solutions

- Send DIP to ATOM with notes on where it has been sent
 - Binder?
- Git-like versioning for metadata
 - possible stored separately
- Rely on backups for disaster recovery, only include master files in AIP
 - Risk of holey or corrupt backups

What about

- Rights metadata?
- Base64?
- Normalization?
- How many copies?
- Bi-directional synchronizations?

It's Expensive





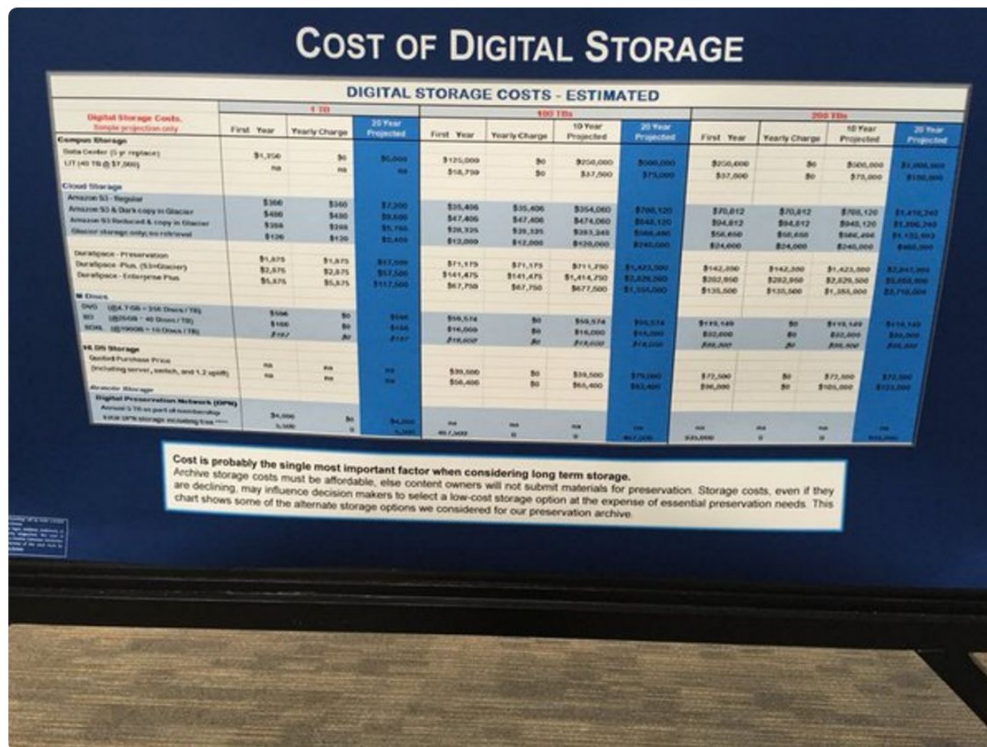
Nicholas Graham

@nmgraham



+ Follow

Handy chart from BYU looking at long-term cost of preservation: 200TB for 20 yrs, \$1.4 million in Amazon #ipres2015



What are you doing? Surely we're not the first

Are these all local implementation decisions?

Communities of best practice?

Can we have confidence in our due diligence?

Do you compute cost/benefit analysis for some of these choices?

Thank You

Nathan Tallman
Digital Content Strategist
University of Cincinnati Libraries
nathan.tallman@uc.edu
[@tallman1029](#)