

Towards Interoperable Preservation Repositories

TIPR

JTC-1 Study Group on Digital Content Management and Protection

August 2010

Joseph Pawletko (NYU)



Presentation Overview

Motivation & Target Audience

Other Approaches

TIPR Approach

TIPR Partners: Heterogeneous Repositories & AIPs

Repository Exchange Package (RXP)

Testing, Issues, & Lessons Learned to Date

Future Plans

Questions

Motivation & Target Audience

Motivation:

Facilitate AIP exchange between heterogeneous preservation repositories

AIP exchange important for:

- redundancy

- software migration

- specialized AIP processing

- succession planning

Target Audience:

those interested in exchanging AIPs between heterogeneous preservation repositories

3

Other Approaches

Other Approaches to AIP exchange between heterogeneous repositories

NDIIPP ECHO DEPOSITORY Hub and Spoke (HandS)

Lead Partner: University of Illinois at Urbana-Champaign

Hub adds preservation metadata during exchange

<http://www.ndiipp.illinois.edu/>

Open Archives Initiative's Object Reuse and Exchange (OAI-ORE)

Fedora/EPrints demonstration project by Tarrant, et. al.

Not formally about AIP exchange, but may be applicable in preservation context

<http://journal.code4lib.org/articles/1062>

4

TIPR Approach

Define a common exchange package format:
the Repository Exchange Package (RXP)

No reliance on transport protocol

RXP accommodates heterogeneous AIP structures and
heterogeneous repository implementations

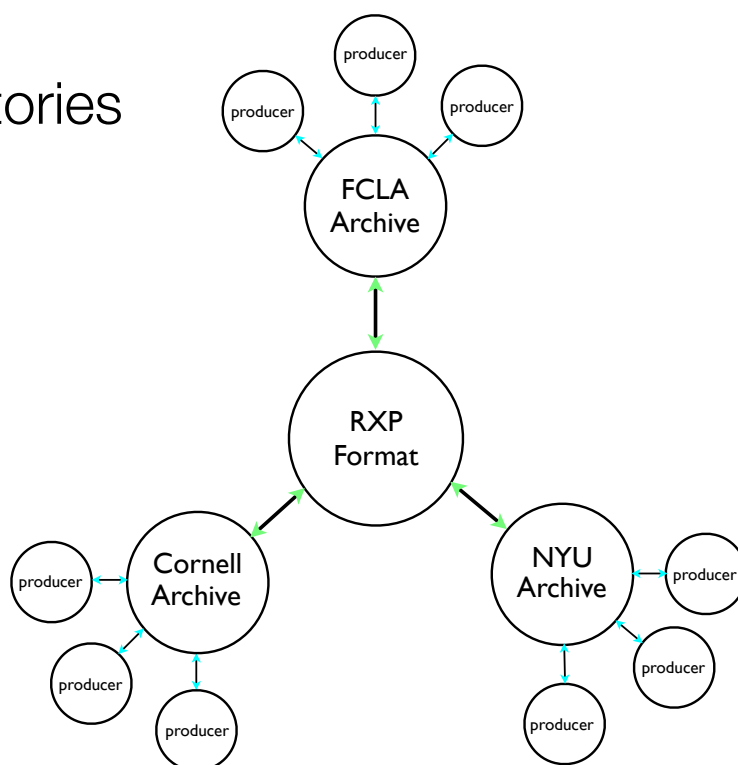
Exchanging Repositories can Ingest and Disseminate
RXPs

Interchanging repositories agree on exchange
parameters

Peer-to-peer repository exchanges

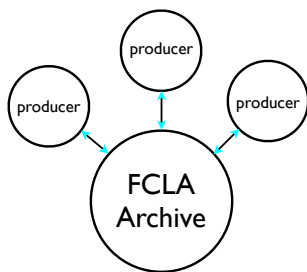
5

TIPR Repositories



6

TIPR
Repositories
FCLA



Florida Center for Library Automation (**FCLA**)

runs the **DAITSS** Preservation Repository

Team: Priscilla Caplan (PI), Franco Lazzarino, Marly Wilson, Randy Fischer

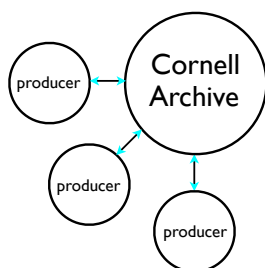
TIPR
Repositories
CUL

Cornell University Libraries (**CUL**)

runs an **aDORe**-based repository

migrating to **Fedora**

Team: Oya Rieger, Bill Kehoe, Rick Silterra, Adam Smith

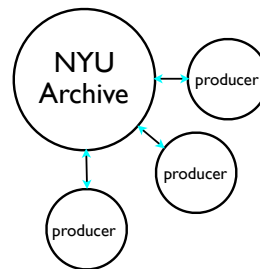


TIPR
Repositories
NYU

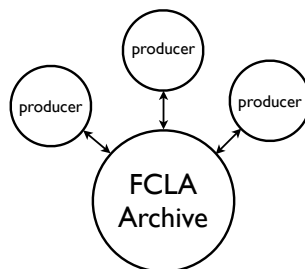
New York University Libraries (**NYU**)

runs a **DSpace**-based repository

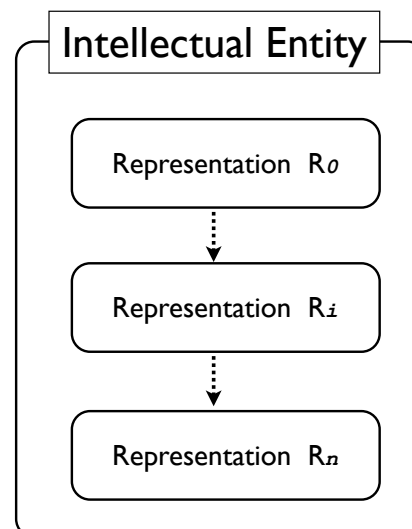
Team: Dr. Michael Stoller, Joseph Pawletko, Rasan Rasch



TIPR
AIP Structures
FCLA

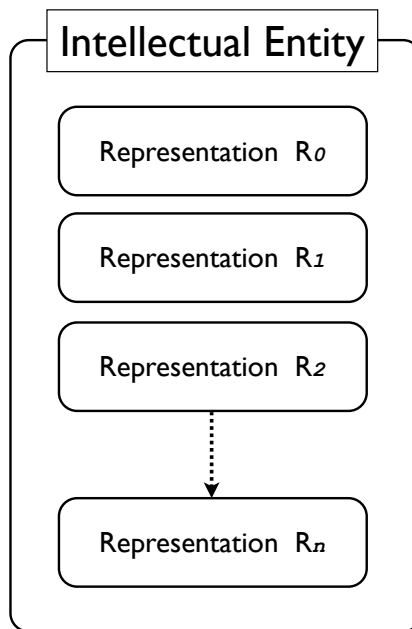
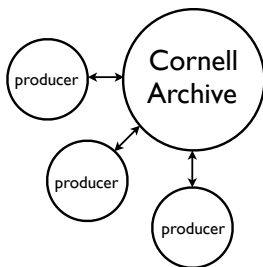


- One AIP per Intellectual Entity
- Retain First and Latest Representations
- Discard Intermediate Representations



TIPR AIP Structures CUL

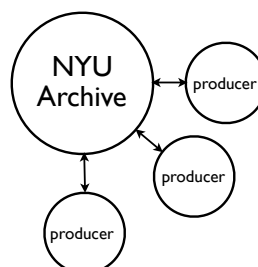
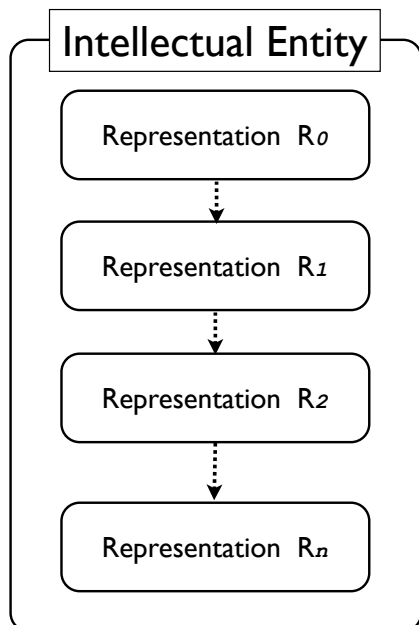
- One AIP per Representation
- Retain All Representation AIPs (including “intermediates”)



11

TIPR AIP Structures NYU

- One AIP per Intellectual Entity
- Retain all Representations (including “intermediates”)



12

Repository / AIP Heterogeneity:

	FCLA	CUL	NYU
Repository Implementation	DAITSS	aDORe -> Fedora	DSpace-based
# of AIPs per n representations	1	n	1
retain intermediate representations?	no	yes	yes

13

Repository Exchange Package (RXP)

RXP design goals:

use standards familiar to Preservation community:

METS and PREMIS

be able to accommodate different AIP structures

contain sufficient data for receiving repository at:

RXP level

Representation level

identify data that receiving repository must understand

14

RXP: Minimal Structure

`./rxp.xml`

`./rxp-digiprov.xml`

`./rxp-rep-1.xml`

`./rxp-rep-1-digiprov.xml`

`./files/`

15

RXP: Minimal Structure

`./rxp.xml`

`./rxp-digiprov.xml`

`./rxp-rep-1.xml`

`./rxp-rep-1-digiprov.xml`

`./files/`

RXP-level
information

16

RXP: Minimal Structure

`./rxp.xml`

`./rxp-digiprov.xml`

`./rxp-rep-1.xml`

`./rxp-rep-1-digiprov.xml`

`./files/`

Representation-
level information

17

RXP: Minimal Structure

`./rxp.xml`

`./rxp-digiprov.xml`

`./rxp-rep-1.xml`

`./rxp-rep-1-digiprov.xml`

`./files/`

METS document
containing
Source
Repository info.
References RXP
provenance,
(optional) rights,
and
representations

18

RXP: Minimal Structure

`./rxp.xml`

`./rxp-digiprov.xml`

`./rxp-rep-1.xml`

`./rxp-rep-1-digiprov.xml`

`./files/`

PREMIS
document
containing
RXP-level
digital
provenance

19

RXP: Minimal Structure

`./rxp.xml`

`./rxp-digiprov.xml`

`./rxp-rep-1.xml`

`./rxp-rep-1-digiprov.xml`

`./files/`

METS document
describing
representation 1

20

RXP: Minimal Structure

`./rxp.xml`

`./rxp-digiprov.xml`

`./rxp-rep-1.xml`

`./rxp-rep-1-digiprov.xml`

`./files/`

PREMIS
document
containing
representation 1
digital provenance

21

RXP: Minimal Structure

`./rxp.xml`

`./rxp-digiprov.xml`

`./rxp-rep-1.xml`

`./rxp-rep-1-digiprov.xml`

`./files/`

directory
containing the
representation
files

22

RXP: Optional Files

`./rxp-rights.xml`

PREMIS document containing
RXP-level rights information

`./rxp.xml.sig`

a stand-alone digital signature in
OpenPGP format generated using
sender's private key, and `rxp.xml`

23

RXP: Optional Files

`./rxp-rep-2.xml`

`./rxp-rep-2-digiprov.xml`

⋮

`./rxp-rep-n.xml`

`./rxp-rep-n-digiprov.xml`

information for
additional
Representations

Each `rxp-rep-n.xml` must be
accompanied by a corresponding
`rxp-rep-n-digiprov.xml`.

24

Transfer Tests

Two different transfer scenarios

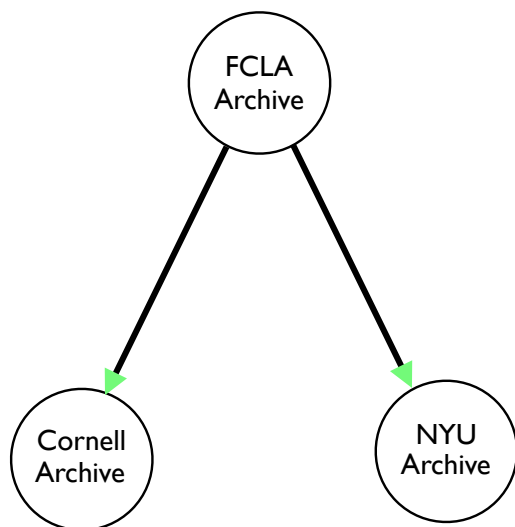
Broadcast Transfer

Ring Transfer

Analyzing test results against expectations
using results to improve RXP structure

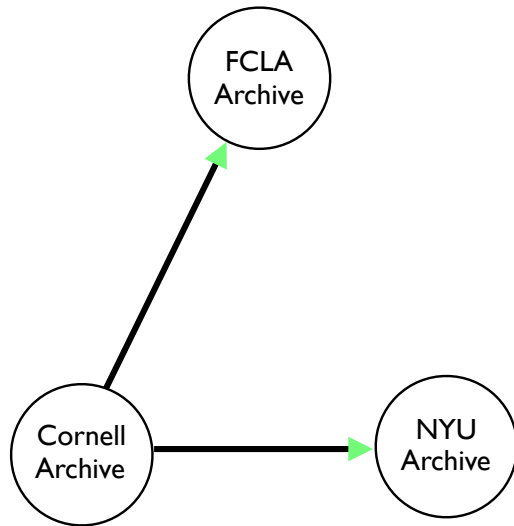
25

TIPR
Testing:
Broadcast
Transfer

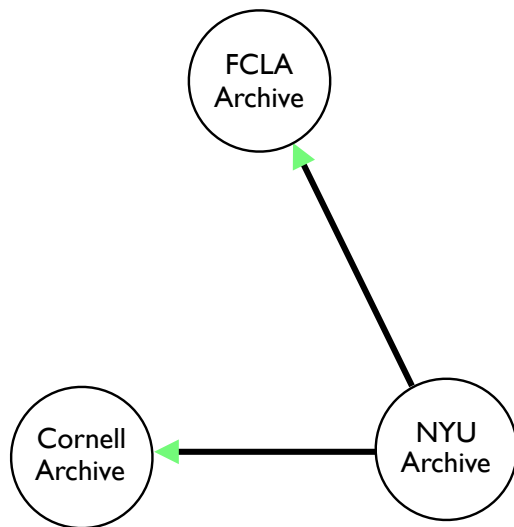


26

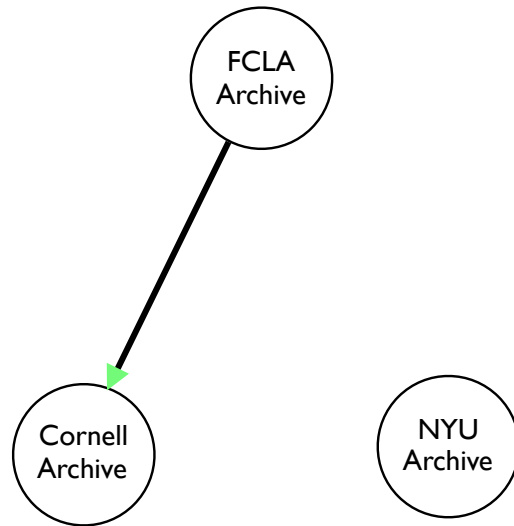
TIPR
Testing:
Broadcast
Transfer



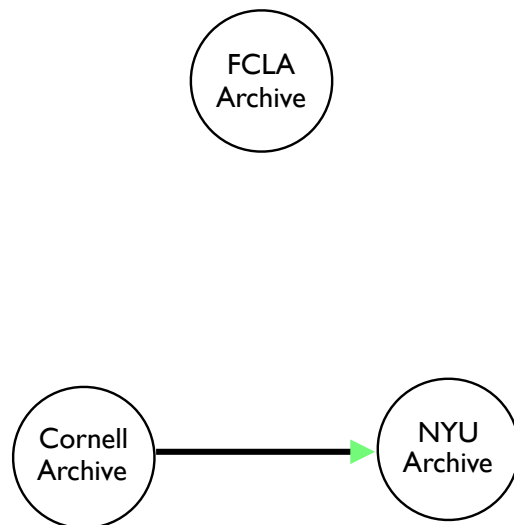
TIPR
Testing:
Broadcast
Transfer



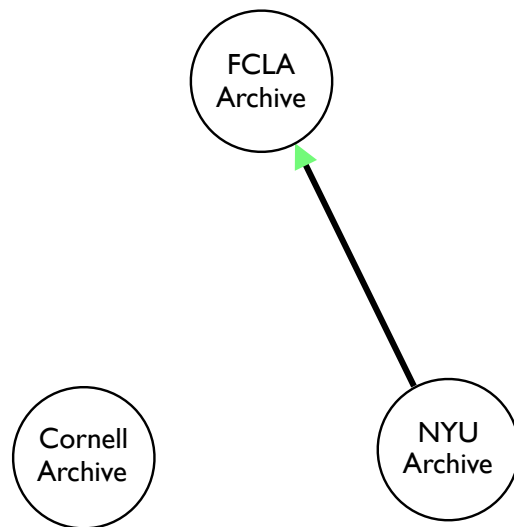
TIPR
Testing:
Ring
Transfer



TIPR
Testing:
Ring
Transfer



TIPR Testing: Ring Transfer



31

Issues:

PREMIS:

Repositories need rights and digiprov at RXP-level

Highest unit of description in PREMIS is **representation**

RXP can contain multiple representations

RXP more like an Intellectual Entity

Asked PREMIS Editorial Committee to consider allowing PREMIS elements to describe Intellectual Entities when applicable

TIPR project still using PREMIS for RXPs

32

Issues:

METS Profiles:

RXP METS profiles are registered and available

<http://www.loc.gov/standards/mets/profiles/00000032.xml>

<http://www.loc.gov/standards/mets/profiles/00000033.xml>

<http://www.loc.gov/standards/mets/profiles/00000034.xml>

33

Issues:

Identified need for Inter-repository Service Agreements (SA):

SA documents inter-repository relationship:

RXP composition vis-à-vis optional files

RXP transfer logistics

target repository actions upon RXP receipt

inter-repository rights and permissions agreements

archiving & preservation treatment at target repository

financial arrangements between source and target repositories

legal aspects of source and target relationship

34

Lessons Learned to Date

Effort required to generate RXPs is reasonable

Ingesting foreign RXPs more difficult, but not prohibitive

Maintaining cross-repository provenance is tricky

Transfer format is only one part of solution

Inter-repository agreements are important

35

Future Plans

Finish grant-related work

Continue to talk and write about TIPR

Respond to feedback from the community on RXP spec

36

Acknowledgements

TIPR is funded by a generous grant from the
Institute of Museum and Library Services (IMLS)

TIPR Towards Interoperable
Preservation Repositories



Cornell University



NEW YORK UNIVERSITY



The Florida Center
for Library Automation

Questions?

Thank you for your time...

<http://wiki.fcla.edu:8000/tipr/>

TIPR Towards Interoperable
Preservation Repositories



Cornell University



NEW YORK UNIVERSITY



The Florida Center
for Library Automation