Digital Preservation Tools, Practices, and Policies in Islandora

Nick Ruest, Mark Jordan, and Donald Moses
OR 2014

Outline

- Tools
- Interest Group
- Use cases

Fixity and File Format Identification

- ## Islandora Checksum
  - Enables checksum generation on datastreams
  - Retrospective checksum generation

- ## Islandora FITS
  - Applied to canonical archival file (OBJ)
  - File Format Identification
  - Alternate Checksum Generation
  - Extraction of Technical Metadata stored in a separate datastream.

Data Integrity

- [Islandora Checksum Checker](#)
  - cron or user configured checksum verification
  - verification is recorded as a PREMIS-compatible event
  - alerts automatically set when mismatch discovered.

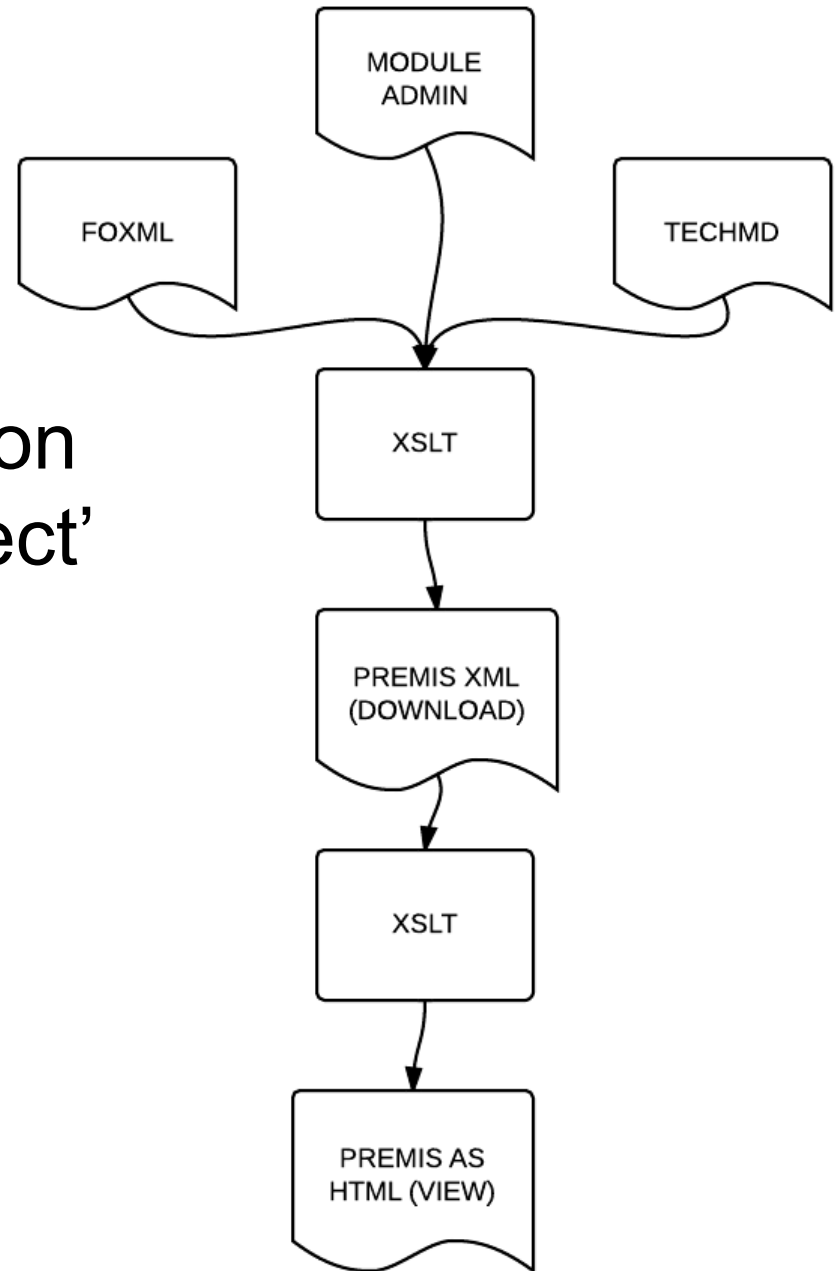# Preservation File Formats / Normalization

- Solution Packs follow preservation best practices
    - Archival / Preservation Formats
    - Ingested files can be normalized to a preferred archival format
        - Typical workflow would be to convert a PDF to PDF/a, a .docx to .odt, etc.
    - Normalize to preservation and access formats

Bitstream/Object Replication

- [BagIt Module](#) exports Bags for individual objects or entire collections of objects
  - Bitstream/Object Copying
  - manifest includes checksums
  - Highly configurable
  - Can include PREMIS XML
  - Can produce Archivematica transfer Bags
- [Islandora Vault](#) allows Islandora to communicate with Fedora Cloudsync and Duracloud

Preservation Metadata

- Islandora PREMIS
- Generates PREMIS XML on demand from Fedora Object's Audit Log, extracted technical, descriptive, and rights metadata
- Includes Agent, Event, Object and Rights entities

MODULE ADMIN

FOXML

TECHMD

XSLT

PREMIS XML (DOWNLOAD)

XSLT

PREMIS AS HTML (VIEW)

"City of Dover" : bought by Penetang group

View **Manage** **MARCXML** **PREMIS**

Download PREMIS

PREMIS Object

Field	Value
objectIdentifierType	Fedora Commons datastreamVersion ID
objectIdentifierValue	AUDIT.0
compositionLevel	0
messageDigestAlgorithm	
messageDigest	
formatName	text/xml
contentLocationType	Fedora Commons contentLocation REF value
objectIdentifierType	Fedora Commons datastreamVersion ID
objectIdentifierValue	RELS-EXT.0
compositionLevel	0
messageDigestAlgorithm	SHA-1
messageDigest	8acd007d964a3bf29e44d0978c1369051a6abbd1
size	544
formatName	application/rdf+xml
contentLocationType	Fedora Commons contentLocation REF value
objectIdentifierType	Fedora Commons datastreamVersion ID
objectIdentifierValue	MODS.0
compositionLevel	0
messageDigestAlgorithm	SHA-1
messageDigest	a94e53eb3f379cdd43594ae55047583400a08bbb
size	2387
formatName	text/xml
contentLocationType	Fedora Commons contentLocation REF value
objectIdentifierType	Fedora Commons datastreamVersion ID
objectIdentifierValue	DC.0
compositionLevel	0
messageDigestAlgorithm	SHA-1
messageDigest	ad931b32519134be6074e22da8f332f79f268585
size	1276
formatName	text/xml
contentLocationType	Fedora Commons contentLocation REF value
objectIdentifierType	Fedora Commons datastreamVersion ID
objectIdentifierValue	OBJ.0
compositionLevel	0
messageDigestAlgorithm	SHA-1
messageDigest	ca62dacfbe4ec9ea85c140295102cefabbb72b4d
size	16129943
formatName	image/tiff
contentLocationType	Fedora Commons contentLocation REF value

Preservation Documentation

- Develop a standard set of preservation policies, action plans and related documentation relevant to an Islandora install.
- Ideally a module that an Islandora Repository administrator could install and then customize the resulting documentation as needed.

https://github.com/yorkulibraries/preservation_documentation/

https://github.com/dmoses/preservation_documentation/

Islandora Preservation Interest Group

"The purpose of the [Islandora Preservation Interest Group](#) is to develop and communicate a common approach to address preservation within the Islandora framework."

- [Repository](#) (Terms of Reference, Agendas, [Meeting minutes](#))
- [Members](#)

It is open and new members welcomed!

Islandora and preservation use cases

- Standalone preservation platform
- Content provider for Archivematica
- Other?