A Place Where You Process:
An Introduction to Archivematica Workflows

archivematica®

Ethan Gates
Talking Tech for Advanced Topics in Preservation Studies
10/19/2017

Itinerary

1. Overview: Why Are We Here? (15 min.)

2. Explore Storage Service and Dashboard interfaces (30 min.)

3. Sample Ingests (45 min.)

4. Q&A/Mess Around Activities (30 min.)

5. Resources and Documentation

What Is Archivematica?

- a networked application/platform

- a suite of integrated open-source preservation tools

- an extensible, flexible, customizable series of micro-services to aid the ingest

  and transfer of digital objects to and from storage

- a practical example of OAIS in action

List of MicroServices and Tools used by Archivematica

- Receive SIP
  - verifyChecksum
- Review SIP   EXT3, Thunar, incron, flock
  - extractPackage
  - assignIdentifier
  - parseManifest
  - clean Filename
- Quarantine SIP   UUID, Detox, Easy Extract, ClamAV
  - lockAccess
  - virusCheck
- Appraise SIP   FITS, JHove, DROID, NLNZ Extractor
  - identifyFormat
  - validateFormat
  - extractMetadata
  - decidePreservationAction
- Prepare AIP   FFident, Unoconv, Ffmpeg, OpenOffice
  - gatherMetadata
  - normalizeFiles
  - createPackage
- Review AIP   ImageMagick, Inkscape, Xena
  - decideStorageAction
- Store AIP   Bagit, SAMBA, NFS-common, Poster
  - writePackage
  - replicatePackage
  - auditfixity
  - readPackage
  - updatePackage
- Provide DIP   ICA-AtoM, DCB Dashboard
  - uploadPackage
  - updateMetadata
- Monitor Preservation
  - checkFormatRegistry
  - migrateFormat
  - synchronizeAIPsandDIPs

What It's <u>Not</u>

A repository

A catalog

Installation Requirements

Operating System:

- Ubuntu 14.04 or 16.04 (beta)
- CentOS 7

Hardware minimums:

- Testing: dual-core CPU, 2 GB RAM, 7 GB disk space (+ 2-3x the size of collection materials)
- Full-scale processing: dual-core CPU, 4 GB RAM, 20 GB disk space (+ 3-4x size of collection materials)
 - likely need more power if working with A/V

See:
https://www.archivematica.org/en/docs/archivematica-1.6/admin-manual/installation/installation/ (last accessed October 2017)

When/Where/Why Would You Use It?

- Lots and lots of files…

- …in a digital repository/on a server (or you want to set up one)

- Regular/frequent acquisition

- Already using integrated front-facing content management/access software:

  - Archivist's Toolkit/ArchivesSpace

  - AtoM

  - DSpace

Let's Look At This Diagram Again
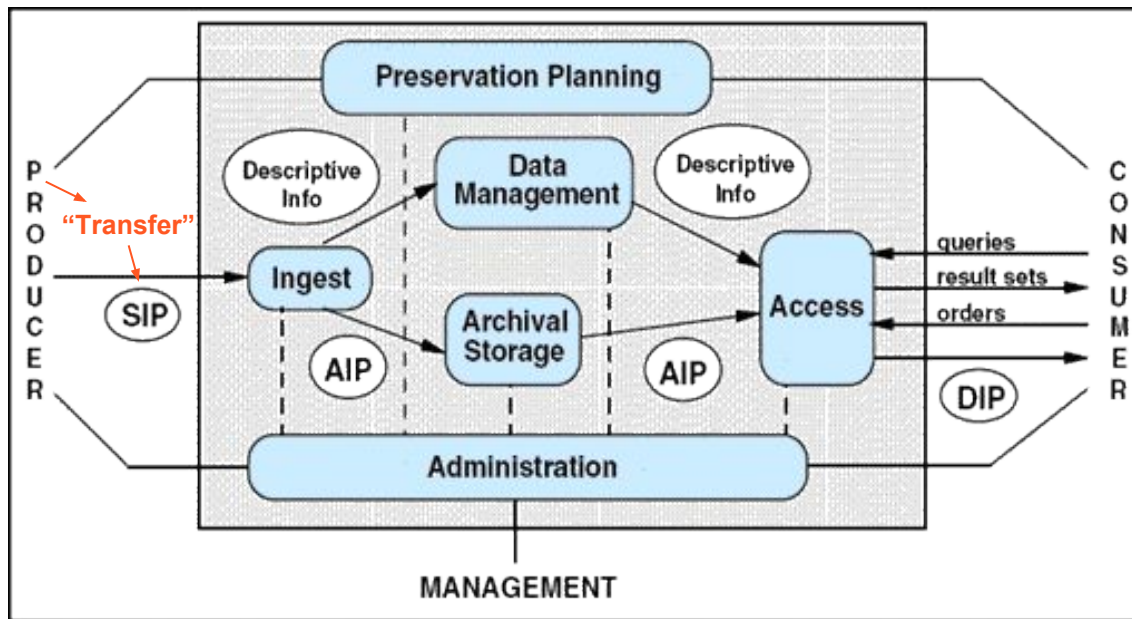
Some advice:

- Take your time (play in a Vagrant test machine or the [Sandbox](#))

- Think micro, not macro

- Don't get too caught up in SIP/AIP/DIP

Dave Rice
@dericed

Follow

I'm officially announcing my defection from OAIS to pioneer the OHAI model with 3 electronic-resource-packages: HERP, DERP & BERP.

RETWEETS 11   LIKES 29

2:52 PM - 10 Dec 2015

11   29
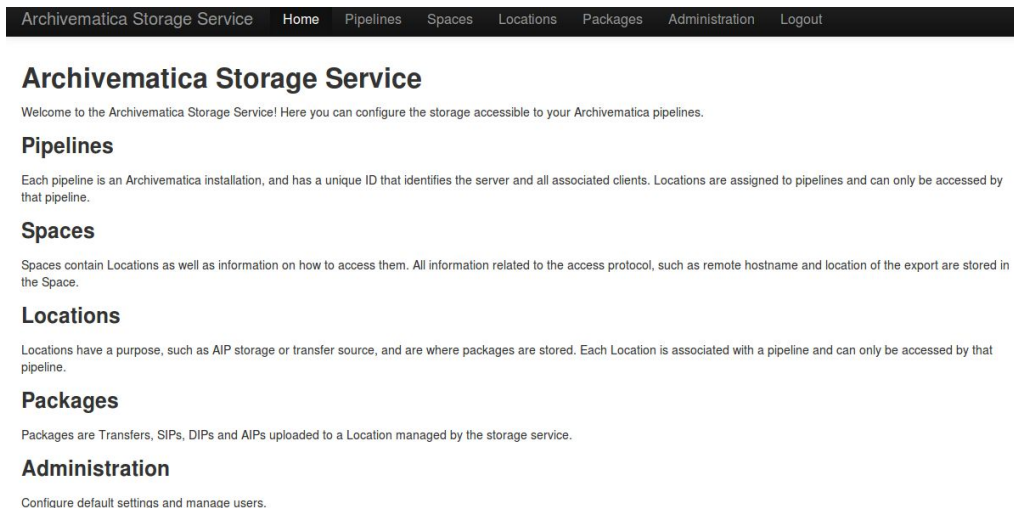
Before We Start!

Please run the following:

- Open Terminal

- $ cd archivematica-vagrant

- $ vagrant up

Archivematica Storage Service

- Navigate in web browser to "http://10.10.10.20:8000"

- Log in with "test/test"

Archivematica Storage Service

Pipeline: a single installation of the Archivematica dashboard software

UUID	Description	Enabled	Edit
7f20eb03-a0de-435e-a34b-59e212bb3d91	Archivematica on miap-VirtualBox	Enabled	Edit | Disable | Delete
Showing 1 to 1 of 1 entries			Previous Next

Archivematica Storage Service

Space: a specific storage device (local hard drive, server accessed via network or remote protocol)
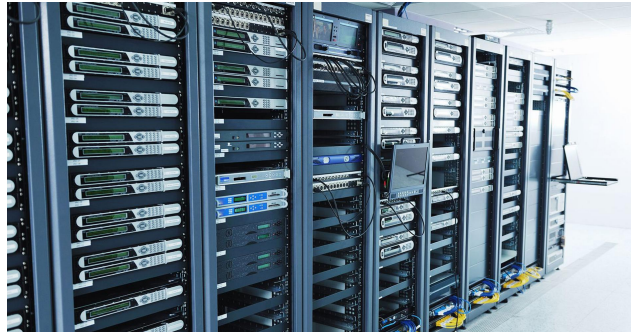
Archivematica Storage Service

Location: sub-division of a Space, used for particular purpose during an Archivematica ingest (AIP Storage, DIP Storage, etc.)
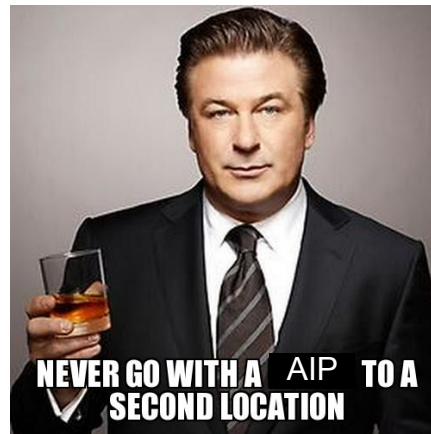
Archivematica Storage Service

Package: a bundle of files, stored in a Location

Translate

Turn off instant translation

| Russian | English | Spanish | **Archivematica** |

I used a local pipeline to transfer and ingest a package to a location on a space

Type text or a website address or translate a document.

| **English** | Russian | Spanish |

Translate

I used Archivematica, installed on my laptop, to move files from an external hard drive to a directory on a server

Archivematica Dashboard

- Open new tab, navigate in browser to "http://10.10.10.20"

- Log in with "miap/miap"
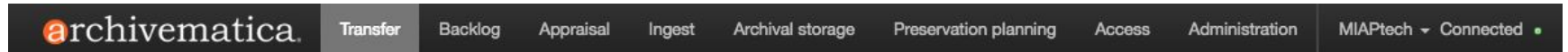
Exercises to Try

- Using SampleTransfers > Multimedia, create a DIP that normalizes .wmv (Windows Media Video) into .mkv (Matroska)

- Activate the micro-service to transcribe incoming files using OCR, then use on SampleTransfers > OCRImage to create an AIP and/or DIP that contains a transcribed .txt file

- Maximize automation as much as possible - what's the fewest steps you can make it through a full transfer/ingest?

Further Resources

Archivematica Wiki: https://wiki.archivematica.org

User Forum (Google Group): https://groups.google.com/d/forum/archivematica

Sarah Romkey (@ArchivesSarah) - Program Manager for Archivematica at Artefactual

"Adding Metadata and Ingesting Large Born-Digital Archives with Archivematica", Dina Sokolova and Jane Gorjevsky, presentation for SAA Forum, 2014
https://academiccommons.columbia.edu/catalog/ac%3A180720