Welcome to Archivematica Camp!

Dinner sign-up: http://bit.ly/york-dinner

Wifi: CityConnectWIFI **Twitter**: #AMCampYork

Camp materials: http://bit.ly/AMCampYork

Breakout room: Pod on the pond (out the front door, farthest to the right)

Agenda

9:00-10:00 - Welcome, introductions, etc.

10:00-12:00 - Introduction to Archivematica

12:00-13:00 - Lunch

13:00-17:00 - Hands-on Archivematica work, or
Introduction to AtoM

18:00-late - small sign-up dinners (http://bit.ly/york-dinner)

Who we are

- **Evelyn McLellan** - President, Artefactual Systems
- **Justin Simpson** - Archivematica Technical Services Director, Artefactual Systems
- **Helen Sherwood-Taylor** - Developer, Artefactual Systems
- **Kirsty Lee** - Digital Preservation Curator, University of Edinburgh
- **Sara Allain** - Systems Archivist, Artefactual Systems
- **Jenny Mitcham** - Digital Archivist, University of York

Who you are

1. Name and institution
2. Your experience with Archivematica, if any
3. What you're hoping to get out of the day/camp

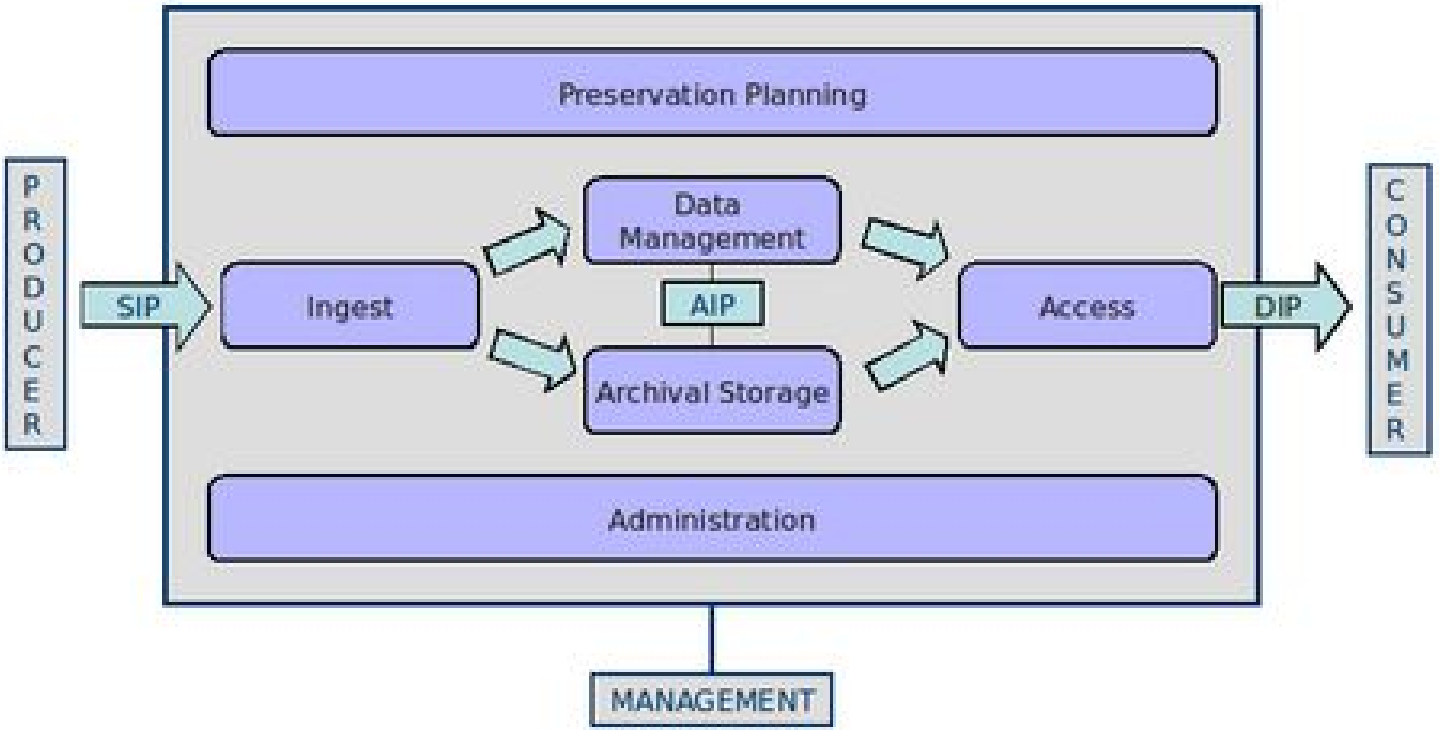
Intro to Archivematica

Sara Allain

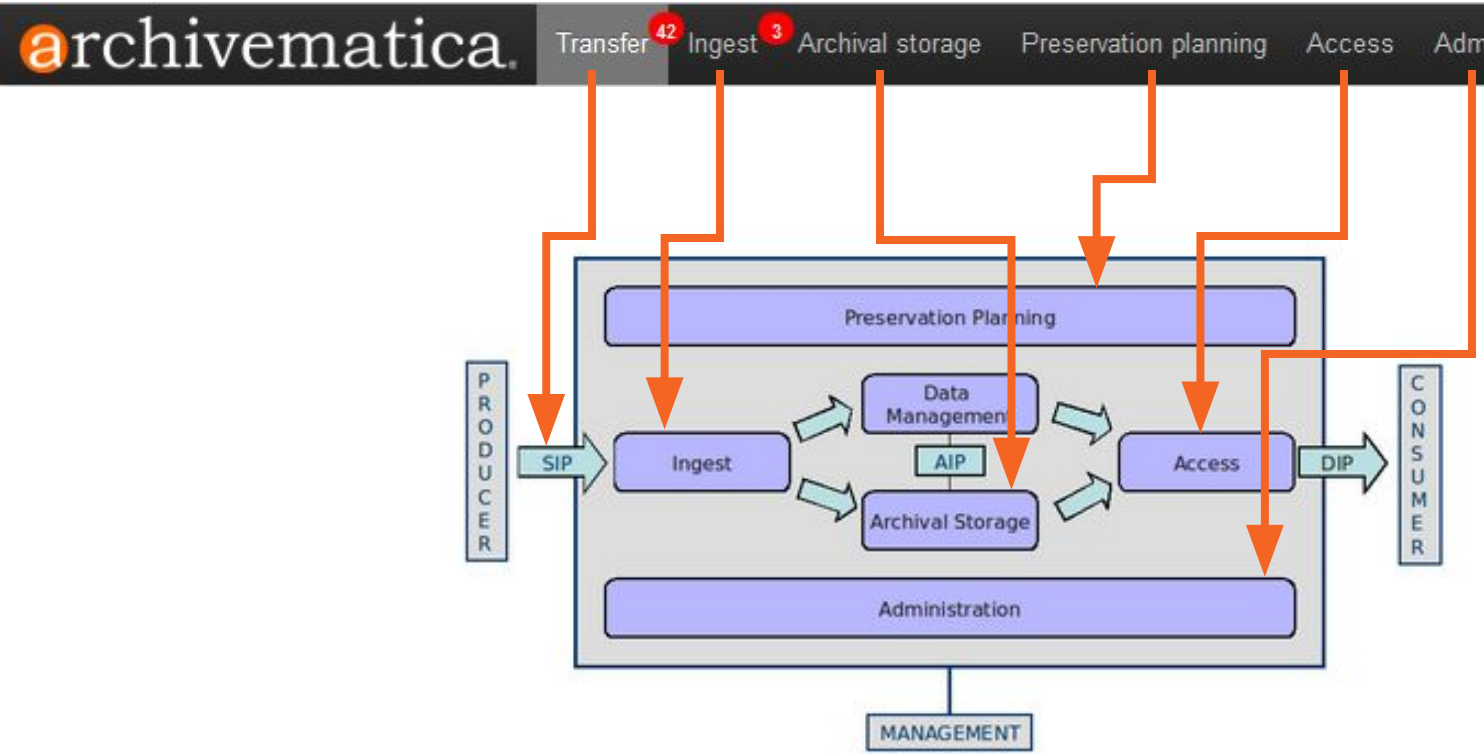
@archivematica

#AMCampYork

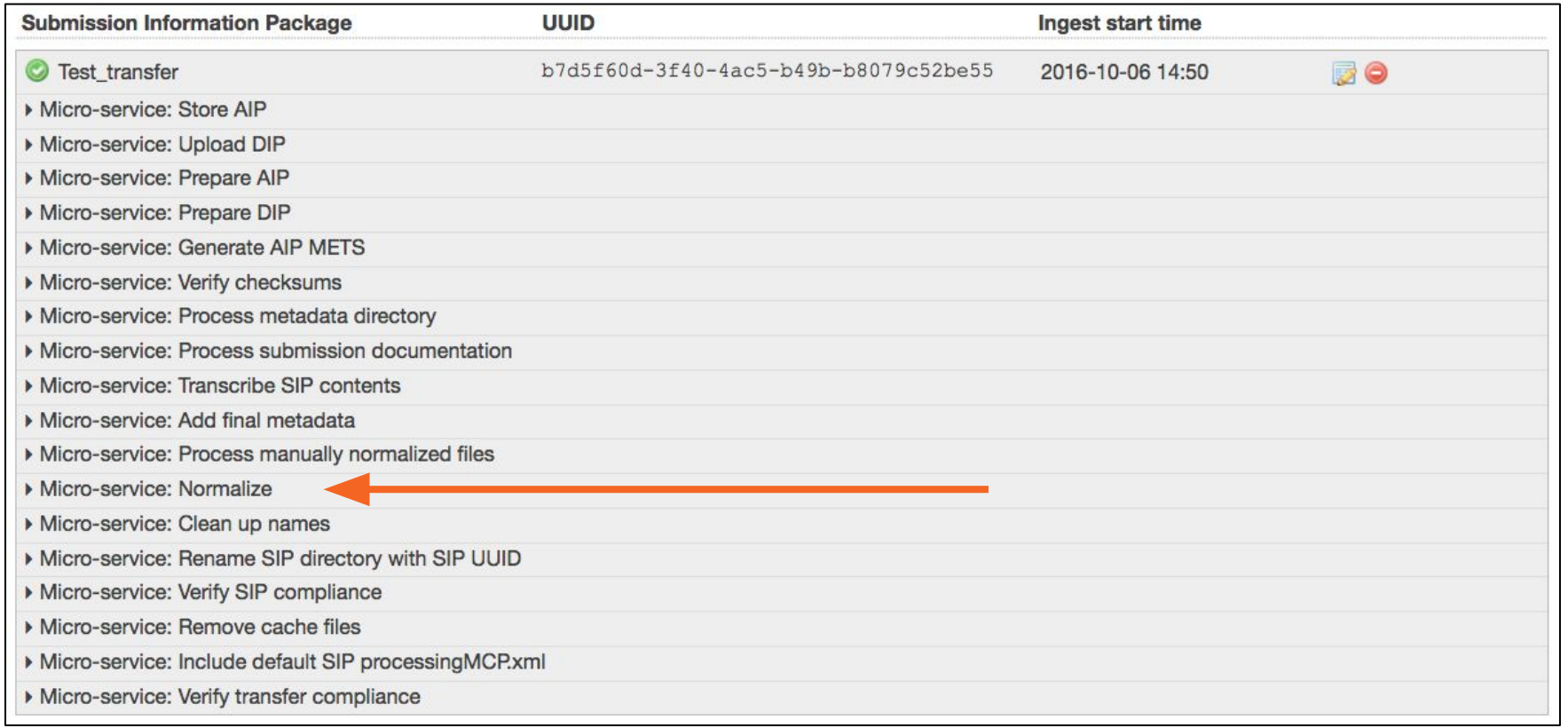
OAIS model


Preservation Planning
PRODUCER
SIP
Ingest
Data Management
AIP
Archival Storage
Access
DIP
CONSUMER
Administration
MANAGEMENT

Mapping the OAIS model



Micro-service architecture



Micro-service architecture

Micro-service: Normalize		
Set file permissions	Completed successfully	
Move to processing directory	Completed successfully	
Approve normalization [?]	Completed successfully	
Move to approve normalization directory	Completed successfully	
Remove files without linking information (failed normalization artifacts etc.)	Completed successfully	
Set file permissions	Completed successfully	
Normalize for preservation	Completed successfully	
Normalize for access	Completed successfully	
Normalize for thumbnails	Completed successfully	
Create thumbnails directory	Completed successfully	
Create DIP directory	Completed successfully	
Move to processing directory	Completed successfully	
Normalize [?]	Completed successfully	
Resume after normalization file identification tool selected.	Completed successfully	
Identify file format	Completed successfully	
Select pre-normalize file format identification command	Completed successfully	
Move to select file ID tool	Completed successfully	
Grant normalization options for no pre-existing DIP	Completed successfully	
Set remove preservation and access normalized files to renormalize link.	Completed successfully	
Check for Access directory	Completed successfully	
Check for Service directory	Completed successfully	
Identify manually normalized files	Completed successfully	

Standards-based architecture

BagIt

- Standard for packaging multilevel, hierarchical content, developed by the Library of Congress (USA)

METS

- XML schema for encoding descriptive, administrative, and technical metadata, also developed by the Library of Congress

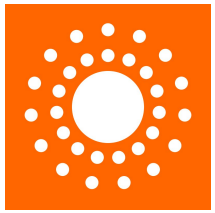
Standards-based architecture

PREMIS

- Standard for defining preservation metadata, such as formats, implementation, hardware requirements, agents, and rights, developed by the Library of Congress

Dublin Core (ISO 15836:2009)

- Standard for capturing descriptive metadata, developed by the Dublin Core Metadata Initiative



Standards-based architecture

PRONOM

- Technical registry providing impartial and definitive information about file formats, software products and other technical components required to support long-term access to electronic records, developed and maintained by the National Archives of the UK.

More external tools

- AtoM
- BagIt
- Bulk_extractor
- ClamAV
- Django
- ElasticSearch
- ExifTool
- FFmpeg
- fido
- FITS
- Gearman
- Imagemagick
- Inkscape
- JHOVE
- md5deep
- MediaInfo
- NFS-common
- p7zip
- Python-lxml
- Seigfried
- Sleuthkit
- Tesseract
- Ubuntu Linux
- UUID
- Unar
- Zip

Authenticity

METS.xml (including PREMIS-in-METS)

- The METS.xml is the statement of record for an Archivematica AIP
- It describes the initial state of the transfer, the changes that took place while Archivematica was running, and the final state of the transfer
- Everything that happens to a file is recorded in the METS.XML

Authenticity

Checksums

- A checksum is a calculation of each unit of data in a file
- Archivematica generates checksums early in the transfer process, ensuring that file integrity is captured at the beginning of the workflow
- If you created checksums prior to transferring your material into Archivematica, you can add them to a transfer. Archivematica will verify them and use them going forward.

Authenticity

UUIDs

- Unique universal identifiers ensure that every file transferred into Archivematica is identifiable
- UUIDs are applied to files, directories, and packages

Relationships

METS.xml structmap

- The METS structmap describes the arrangement of the SIP and the AIP, preserving contextual relationships even if the files are moved

Metadata

- Information included in the Dublin Core metadata dc.relations field will be written to the METS as well

Format agnostic packages

- Archivematica creates content and format agnostic AIPs, meaning that you do not require a particular system to store and read AIPs in the future
- AIPs can be stored in any file system that permits packaged formats (.tar files, .zip files)
- You can migrate AIPs between systems just like any other type of file or package

Hosted and local storage

- The Archivematica team has completed storage integrations with Amazon Web Services (S3 and Glacier), Microsoft Azure
- There are full integrations with systems like Arkivum and DuraCloud

Store files where it makes sense

Arkivum

DURACLOUD™

- Whatever storage system you choose to use, you can implement tools to run checksum validation, virus scans, and other authentication tools.
- Systems like Arkivum and DuraCloud are intended for long-term storage and have built-in authentication tools.
- Ensuring that your data is regularly backed up and verified happens outside of Archivematica, but is critical for meeting this requirement.

Providing access

- Archivematica has integrations with AtoM, ArchivesSpace (and Archivists' Toolkit), and DSpace - you can automatically upload DIPs to these systems to make them accessible to the public.
- DIPs are also system-agnostic, so developing integrations with other open-source platforms is easy - or the contents can be uploaded manually.

Archivematica is one tool of many in your toolbox.

It must be combined with other tools to create a holistic preservation system.

Digital preservation
does not have a
one-click solution.

Building systems using the best standards, the best tools, and the best practices...

And ensuring that the system is content agnostic, open source, and interoperable...

Gives us the best opportunity to preserve our digital content.

Archivematica

With Archivematica, we've tried to build a tool that is robust enough that takes care of a lot of the more difficult aspects of digital preservation, like checksum creation and validation, format identification, and file normalization.

But we've also tried to build a tool that is customizable and extensible enough to work with use cases in many different kinds of institutions, all over the world.

Archivematica

Adding new tools, removing obsolete tools, and adapting the system to reflect best practices in digital preservation is key to making sure that Archivematica continues to be a leader in digital preservation.

Archivematica is built for and by a community of dedicated users and digital preservation experts.