DIGITAL ARCHIVES & PRESERVATION SYSTEMS

Part 1 Overview (part 1 of 7)

Kari R. Smith, MIT Institute Archives

Session Overview

- Digital archives and digital preservation systems.
- These open source tools are being developed and used in the digital archives and digital preservation communities to address the needs and gaps in systems for meeting standards, such as *OAIS, PAIMAS,* and *Trustworthy Digital Repositories.*
- They enable description, discovery, delivery, and preservation of digital collections.
- Some of these are currently being assessed and/or implemented at MIT Libraries.

Systems and Tools to be presented during this series

- **ArchivesSpace** – archives collection management and discovery
- **Archivematica** – digital preservation system (processes files chosen for preservation and creates Archival Information Packages and Dissemination Information Packages)
- **(atom) Access to Memory** – archives and special collections discovery and delivery system and collections management system
- **BitCurator and BitCurator Access** – digital forensics tools and processes for archival and special collections material
- **ePADD** – electronic mail processing, access, discovery system
- **Fixity** – a tool that enables the preservation actions of fixity checking of files in archival (or other) storage. Includes reporting on file attendance and file integrity.
- **Binder** – a tool that integrates with Archivematica to view administrative, technical, descriptive, and preservation metadata related to AIPs and the relationships among the metadata.

Lifecycle Environment

Storage Pathways

Digital Content Management Workflow

External Content Creators

Archives Collections Security Requirements

Storage Pathways

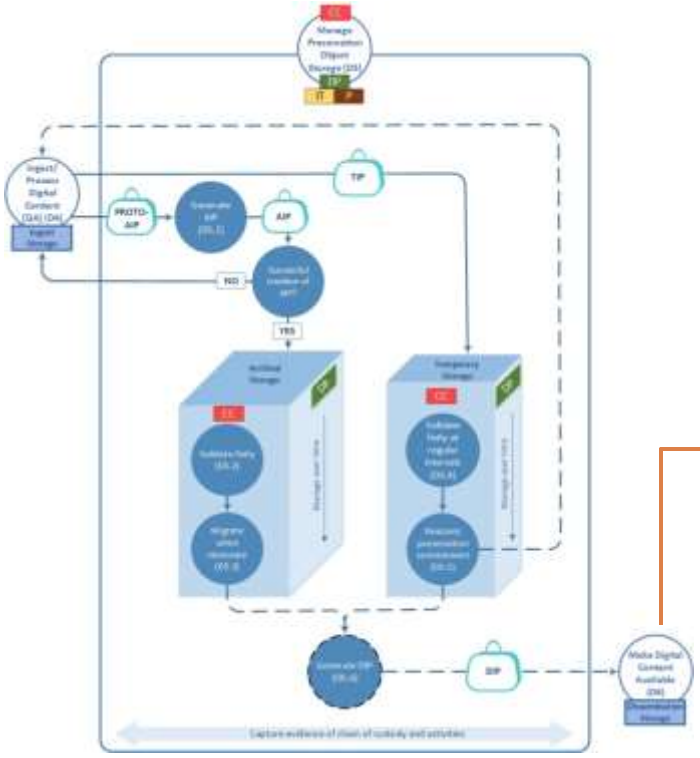
Steps in *Manage Digital* workflow showing Storage Areas and Activities

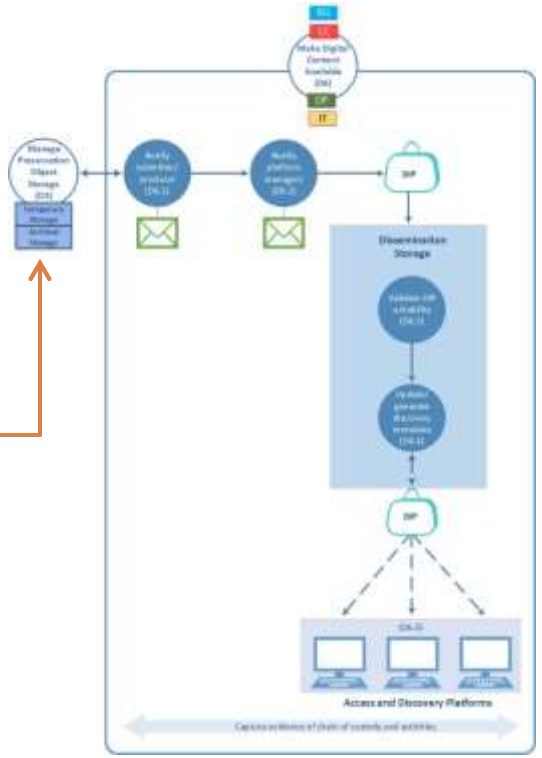
Submission Storage

Ingest Storage

Archival Storage

Dissemination Storage

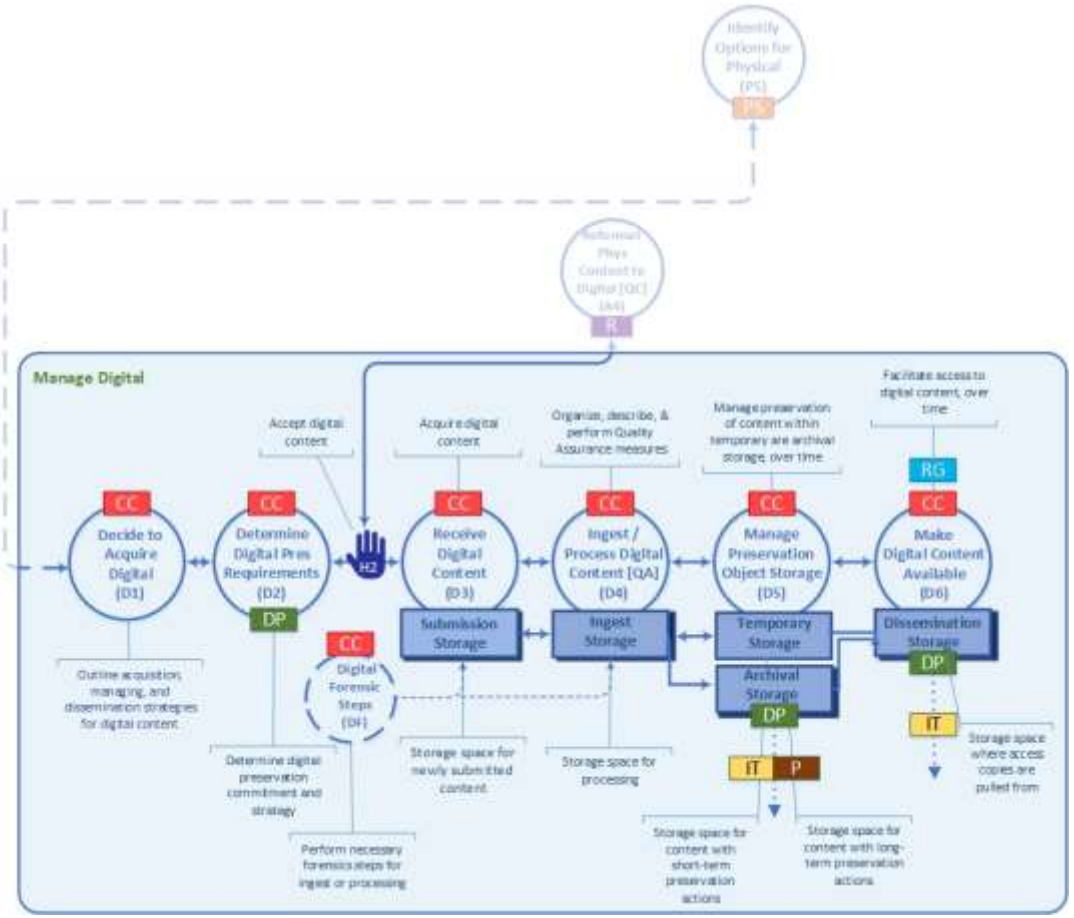


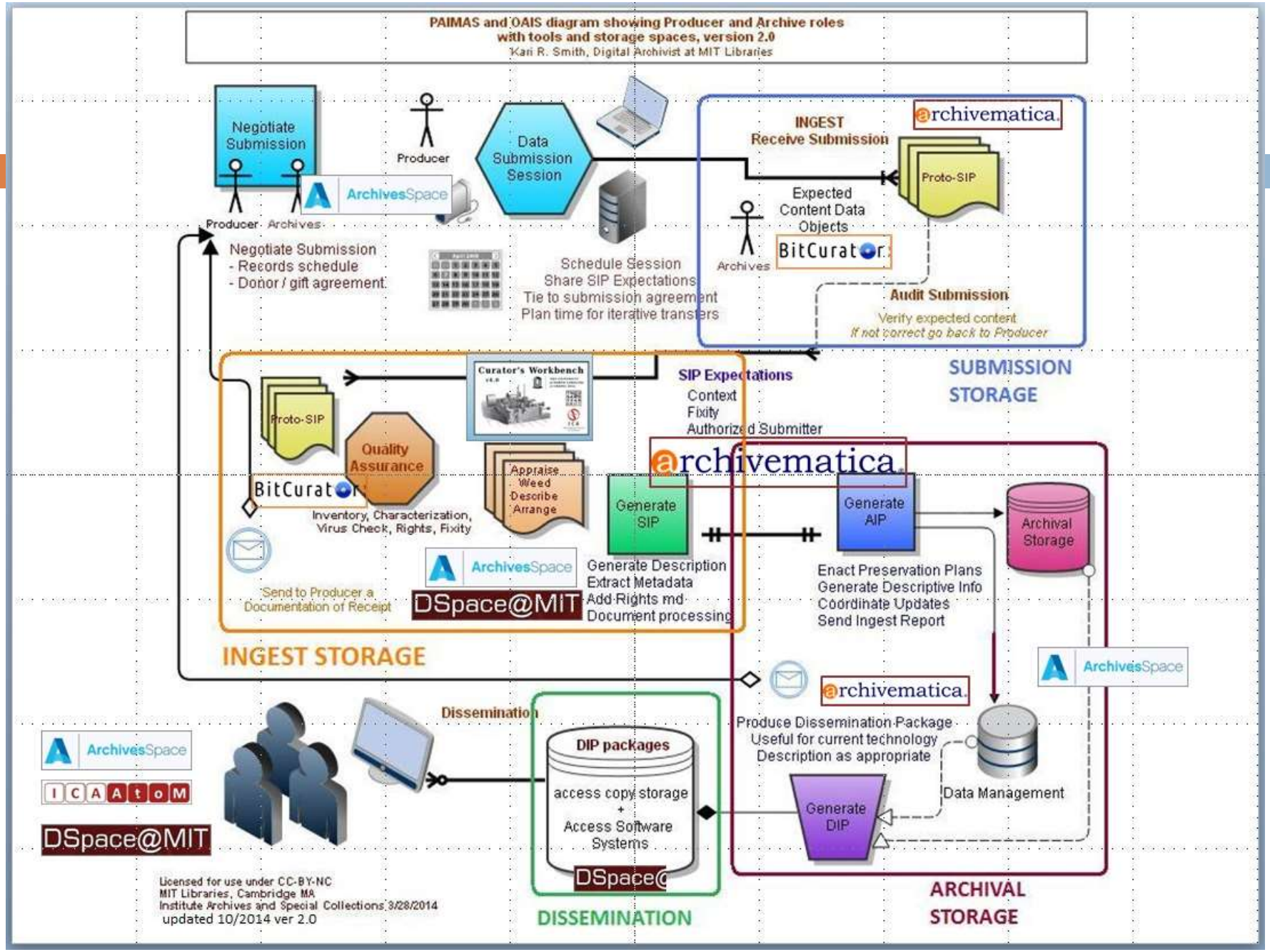


Manage Digital Workflow

Diagram showing steps in the Manage Digital Workflow and the Storage Areas.

Digital Content Management – Lifecycle Workflow [on the DCM wiki]




PAIMAS and OAIS diagram showing Producer and Archive roles
with tools and storage spaces, version 2.0
Kari R. Smith, Digital Archivist at MIT Libraries
Negotiate Submission
Producer
Archives
Data Submission Session
ArchivesSpace
Negotiate Submission
- Records schedule
- Donor / gift agreement
Schedule Session
Share SIP Expectations
Tie to submission agreement
Plan time for iterative transfers
INGEST
Receive Submission
archivematica
Proto-SIP
Expected Content Data Objects
BitCurator
Audit Submission
Verify expected content
If not correct go back to Producer
SUBMISSION STORAGE
Curator's Workbench
SIP Expectations
Context
Fixity
Authorized Submitter
Proto-SIP
Quality Assurance
BitCurator
Inventory, Characterization, Virus Check, Rights, Fixity
Appraise
Weed
Describe
Arrange
Generate SIP
archivematica
Generate AIP
Archival Storage
Send to Producer a Documentation of Receipt
ArchivesSpace
DSpace@MIT
Generate Description
Extract Metadata
Add Rights md
Document processing
Enact Preservation Plans
Generate Descriptive Info
Coordinate Updates
Send Ingest Report
INGEST STORAGE
ArchivesSpace
archivematica
Produce Dissemination Package
Useful for current technology
Description as appropriate
Data Management
Dissemination
DIP packages
access copy storage
+
Access Software Systems
DSpace@
Generate DIP
ArchivesSpace
ICAAtoM
DSpace@MIT
ARCHIVAL STORAGE
DISSEMINATION
Licensed for use under CC-BY-NC
MIT Libraries, Cambridge MA
Institute Archives and Special Collections 3/28/2014
updated 10/2014 ver 2.0

ArchivesSpace

- Archives Management System

  "Software that provides integrated support for the archival workflow, including *appraisal*, *accessioning*, *description*, *arrangement*, *publication of finding aids*, *collection management*, and *preservation*."

http://test.archivesspace.org:8080/ http://archivesspace.vmi.edu/

Access to Memory (atom)

- Archives Management System

  Software that provides integrated support for the archival workflow, including appraisal, accessioning, description, arrangement, publication of finding aids, collection management, and preservation.

- Digital Content Delivery system

- http://demo.accesstomemory.org/

BitCurator

Tools for Digital Forensics Methods and Workflows in Real-World Collecting Institutions

- Adapt digital forensics tools for archives, libraries, other collecting institutions
- Popular Use Case – complete migration of data off media carriers
- Use forensic tools for data triage and understanding of content at disk level
- Process and export data from *disk image*
- *Open Source*

Archivematica

- Process your files for Preservation and Delivery
- After selection and description
- Workflow engine that
  - Sends your files through a set of tools
  - You make decisions during the workflow
  - You set the locations for the files to live
  - Archivematica collects and combines the metadata about all of the tasks, processes, and results into a METS file
  - Archivematica produces Bags (Bag-It) for storage

Integrations

ArchivesSpace – Archivematica – ArchivesSpace

ArchivesSpace - atom – Archivematica – atom

BitCurator – ArchivesSpace – Archivematica - ArchivesSpace

Other Tools…

Select Your Content

- ePADD
- QuickViewPlus
- Webrecorder.io

Monitor Your Content

- Fixity
- Binder

ePADD
[email: process, appraise, discover, deliver]

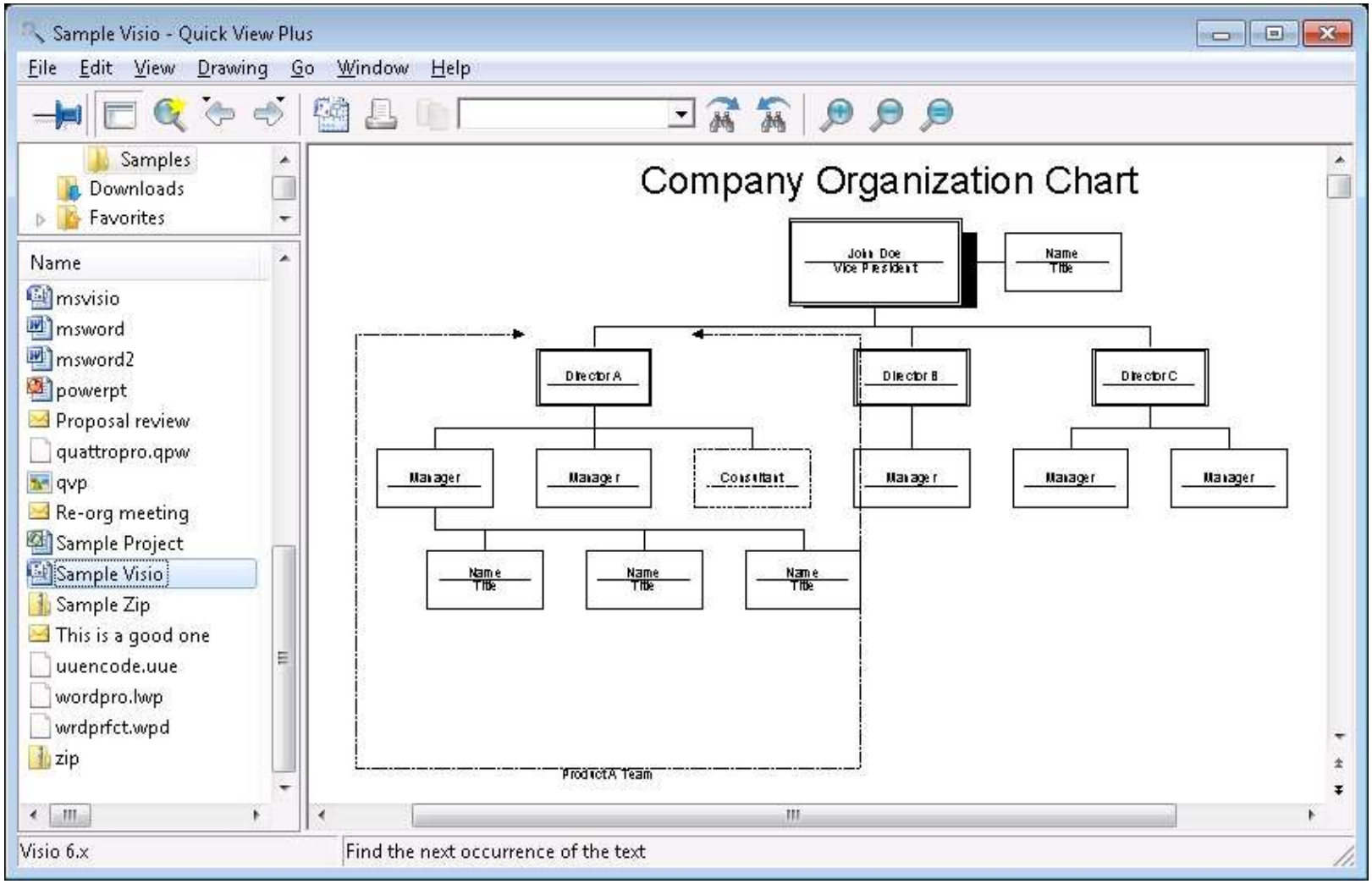
- An open-source tool that will allow archivists and other individuals to interact with email archives before and after they have been transferred to an Archival repository
- Multiple partners involved
- Four modules based on a functional activities:
  - Processing (arrangement and description)
  - Appraisal (collection development)
  - Discovery (online via the web)
  - Delivery (access)
  - http://epadd.stanford.edu/muse/archives/

QuickView Plus

- Current files are easy to render / use
- Older file formats require either transformations or old software
  - May require licenses ($)
  - May require older operating systems

- QuickView Plus can render the files so you can appraise them and make decisions
- QuickView Plus can be used to make your files available to patrons also

Simple Interface



Webrecorder.io

- Manual capture
- Creates WARC file
- Viewers
  - Off-line
  - On-line

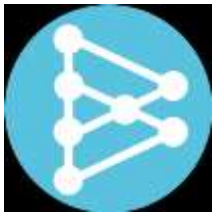
Monitor digital files

Fixity

- A utility for the documentation and regular review of stored files.
  - Fixity scans a folder or directory, creating a **manifest** of the files including their file paths and their checksums, against which a **regular comparative analysis** can be run.
  - Fixity **monitors file integrity** through generation and validation of checksums, and **file attendance** through monitoring and reporting on new, missing, moved and renamed files.
  - Fixity **emails a report** to the user documenting flagged items along with the reason for a flag, such as that a file has been moved to a new location in the directory, has been edited, or has failed a checksum comparison for other reasons.

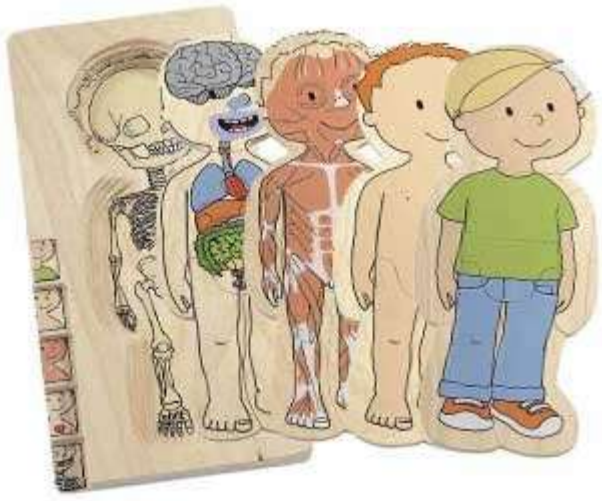
Binder

- The app provides a **central interface** through which they can access, view and manage the **rich technical metadata** contained in Archival Information Packages (**AIP**s) held by the repository.
- It also allows for managing and describing the relationships between the components of a collections object, its constituent digital objects, and the various external dependencies required to preserve and display the collection over the long-term.
- Binder gathers together all of this information required to **make long-term preservation and assessment decisions** in a single user-friendly interface.



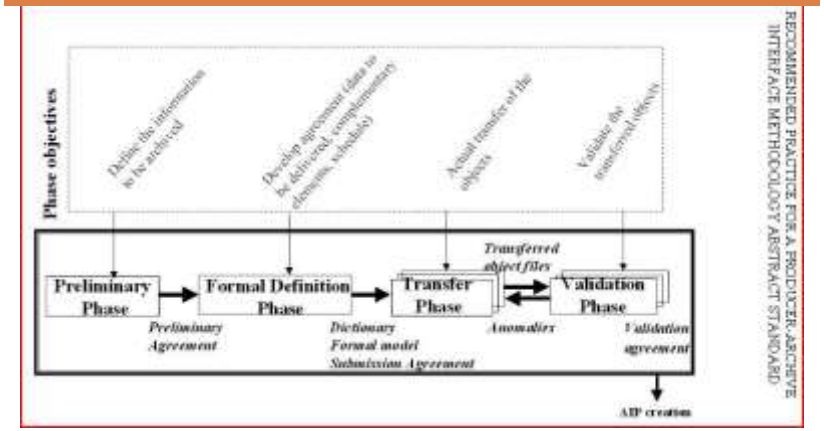
Breaking down the Ecosystem

Remember those
multi-layered anatomy books?

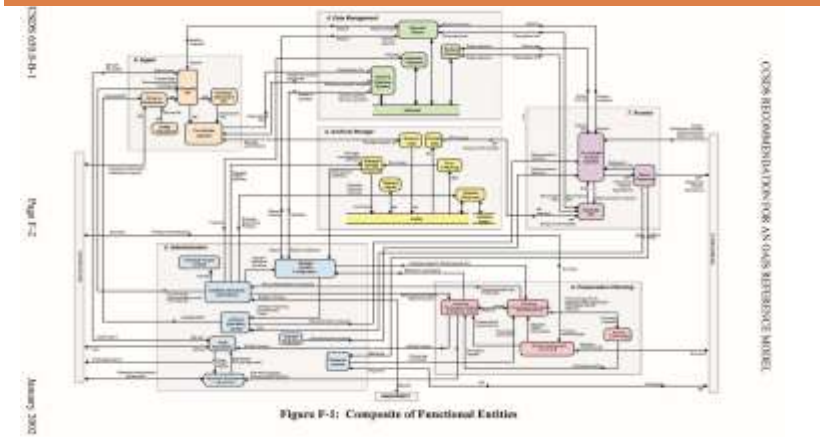


Mapping Standards to Local

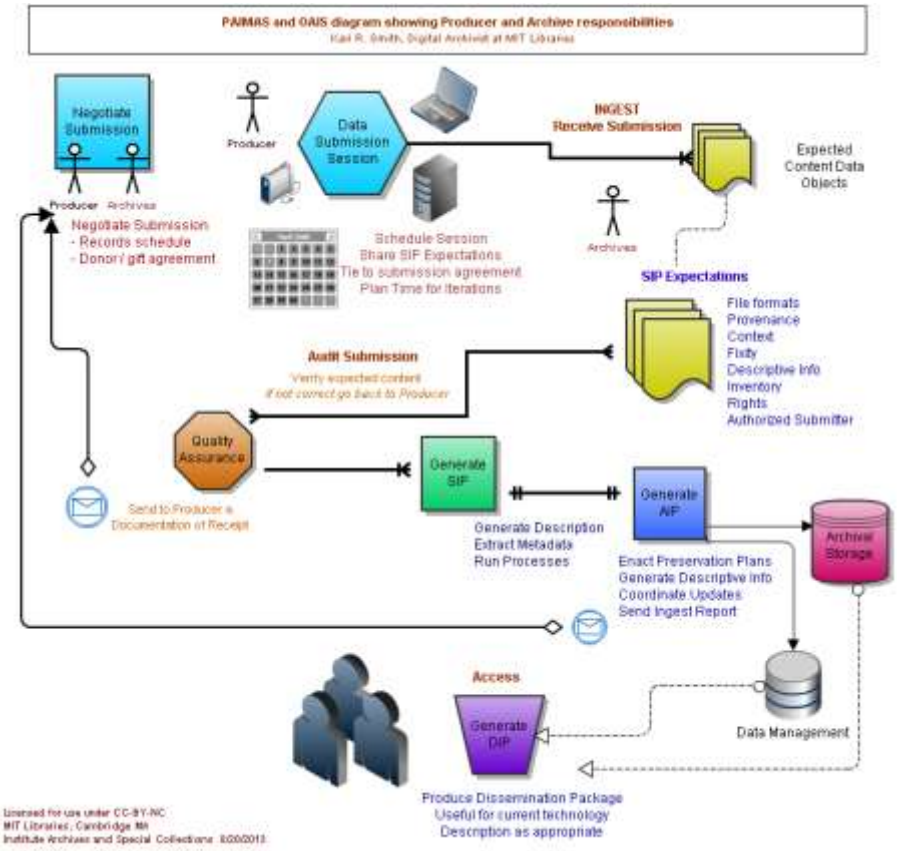
PAIMAS: Producer – Archive Interface Methodology Abstract Standard



OAIS

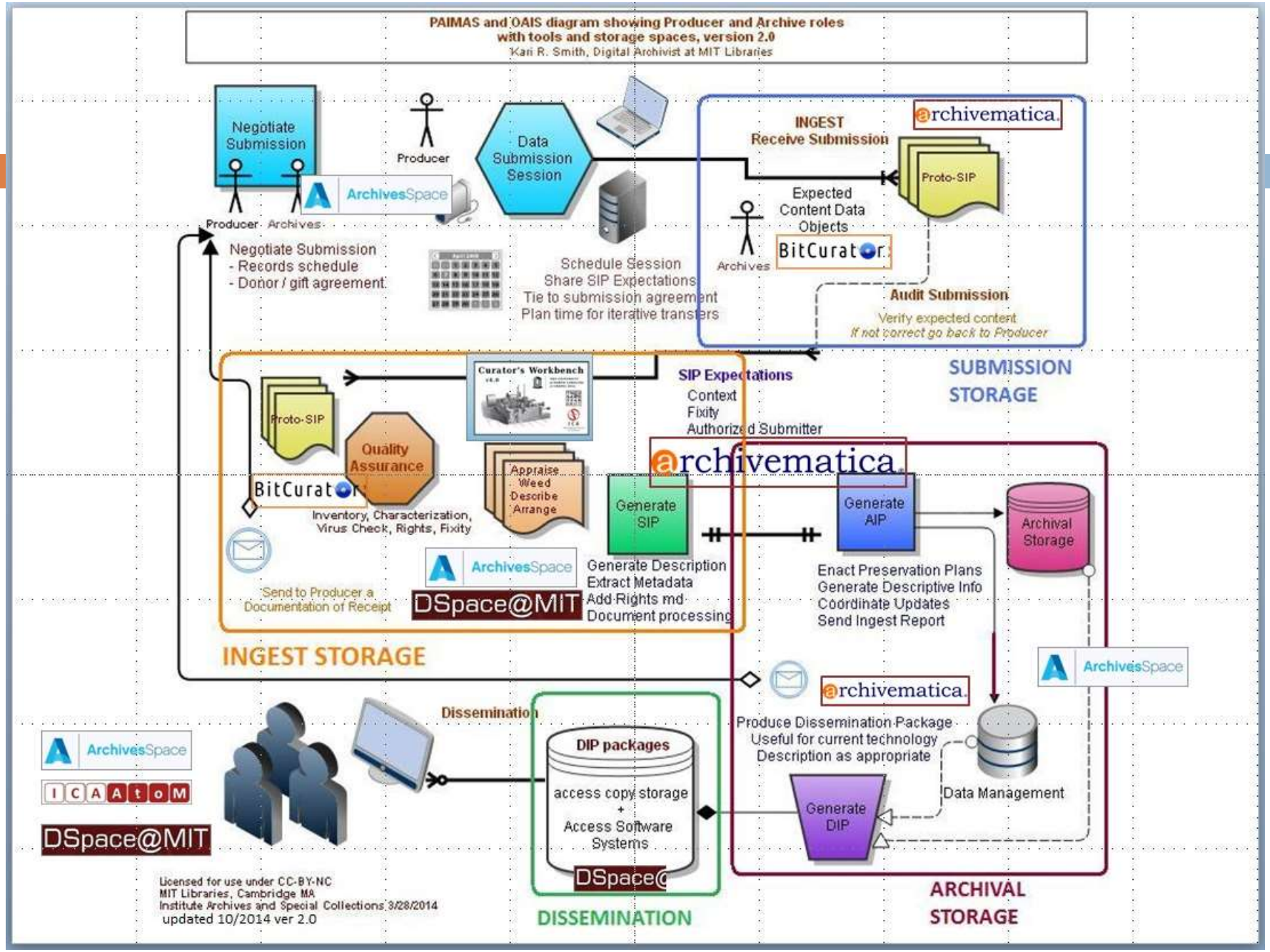


PAIMAS and OAIS diagram showing Producer and Archives responsibilities



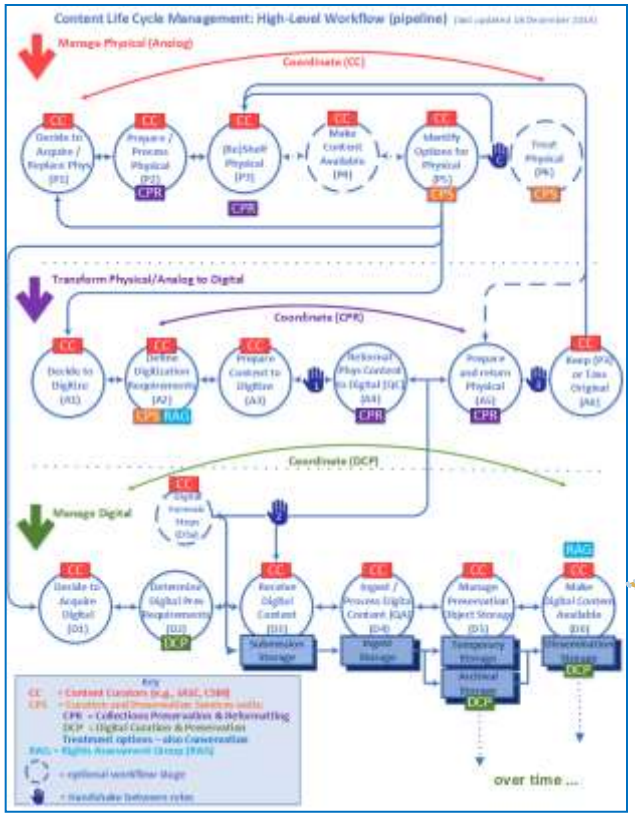
Mission, Tasks, Components

- What is our mission and goals?
- Break them into tasks to accomplish
- Determine where they happen during workflow
- Identify
  - Information inputs
  - Digital content inputs
  - Intended output
- May these onto the lifecycle storage areas
- Select tools that assist us accomplish our goals
  - **Monitor, Adjust, Repeat**

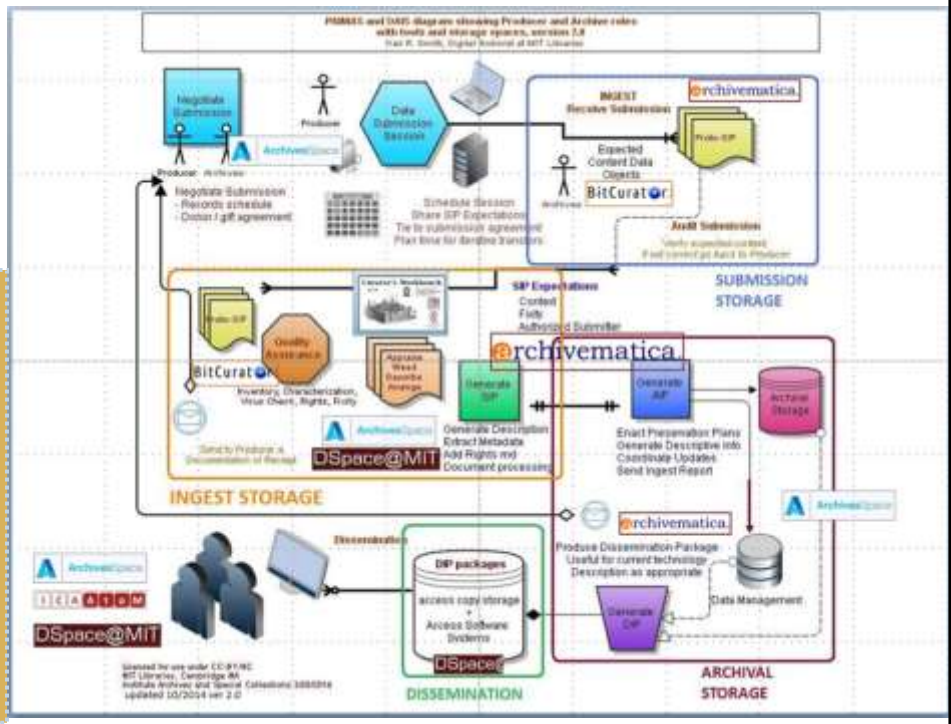

PAIMAS and OAIS diagram showing Producer and Archive roles
with tools and storage spaces, version 2.0
Kari R. Smith, Digital Archivist at MIT Libraries
Negotiate Submission
Producer
Archives
Data Submission Session
ArchivesSpace
Negotiate Submission
- Records schedule
- Donor / gift agreement
Schedule Session
Share SIP Expectations
Tie to submission agreement
Plan time for iterative transfers
INGEST
Receive Submission
archivematica
Proto-SIP
Expected Content Data Objects
BitCurator
Audit Submission
Verify expected content
If not correct go back to Producer
SUBMISSION STORAGE
Curator's Workbench
SIP Expectations
Context
Fixity
Authorized Submitter
Proto-SIP
Quality Assurance
BitCurator
Inventory, Characterization, Virus Check, Rights, Fixity
Appraise
Weed
Describe
Arrange
Generate SIP
archivematica
Generate AIP
Archival Storage
Send to Producer a Documentation of Receipt
ArchivesSpace
DSpace@MIT
Generate Description
Extract Metadata
Add Rights md
Document processing
Enact Preservation Plans
Generate Descriptive Info
Coordinate Updates
Send Ingest Report
INGEST STORAGE
ArchivesSpace
archivematica
Produce Dissemination Package
Useful for current technology
Description as appropriate
Data Management
Dissemination
DIP packages
access copy storage
+
Access Software Systems
DSpace@
Generate DIP
ArchivesSpace
ICAAtoM
DSpace@MIT
ARCHIVAL STORAGE
DISSEMINATION
Licensed for use under CC-BY-NC
MIT Libraries, Cambridge MA
Institute Archives and Special Collections 3/28/2014
updated 10/2014 ver 2.0

Placing the Ecosystem in Context

Content Lifecycle Management



Manage Digital = IASC Ecosystem



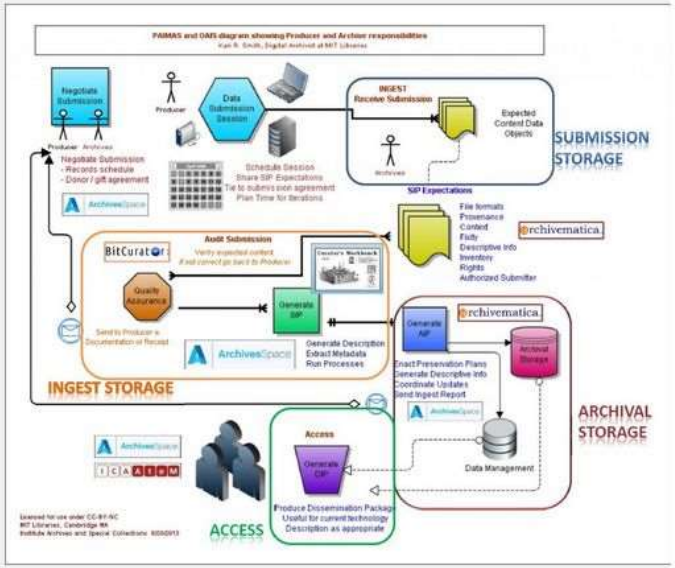
Engineering the Future of the Past

Blog about MIT Libraries' Digital Archives Work

Integrating Tools into our Process and Workflows

Posted on November 7, 2013 by Kari R. Smith

In the last blog post, I described the process and data flow for digital archives using the PAIMAS and OAIS roles, responsibilities, and a few of the functions. Since that diagram, I have received feedback from MIT colleagues and have over-laid onto it two important considerations. First, how the storage spaces for metadata and digital objects fit into the process, and second where some of the software tools we are considering will fit into the processes.



Proposed workflow showing human, tool, and storage interactions.

Search

About this blog

Engineering the Future of the Past shares information on what the MIT Institute Archives is doing about acquiring and managing born-digital and digitized collections.
Kari R. Smith, Digital Archivist [smithkr [at] mit.edu] is the primary Blog author.

Links

- MIT Institute Archives

Recent Posts

- Integrating Tools into our Process and Workflows
- Visualizing PAIMAS and OAIS
- Resources we've been reading

Previous Posts

Select Month

Categories

- All (12)
- Events (3)
- Gift agreements (2)
- Legal issues (4)
- life-cycle experiments (1)
- metadata (3)
- personal archives (6)
- roles and responsibilities (4)
- Tools (8)
- workflows (2)

Posts

diagram digital archives documentation finding IAP metadata MIT OAIS PAIMAS presentation role tagging template

THANK YOU FOR YOUR ATTENTION!

JUNE 11 (1PM): ARCHIVESSPACE
JUNE 24 (NOON): BITCURATOR
JULY 14 (11AM): ARCHIVEMATICA AND STORAGE SERVICE
JULY 16 (10AM): ACCESS TO MEMORY (ATOM)
JULY 21 (9AM): EPADD
JULY 23 (9:30AM): FIXITY

smithkr@mit.edu