Digital Preservation Progress at the Cornell University Library

Nancy McGovern

Digital Preservation Officer

Marcy E. Rosenkrantz

Director of Library Systems

Cornell University Library

CUL's History in Digital Preservation

- Earliest efforts ca 1988
- Early Digitization Efforts aimed at preserving access to print (e.g. Making of America)
- RLG DigiNews — Produced by the Preservation Department on behalf of RLG since 1997, this bi-monthly newsletter focuses on digital imaging and preservation issues that pertain to educational and cultural institutions. Anne R. Kenney and Nancy Y. McGovern are the co-editors.
- Anne R. Kenney and Oya Y. Rieger, Moving Theory into Practice: Digital Imaging for Libraries and Archives. Mountain View, CA: Research Libraries Group, Inc., 2000.
- Risk Management of Digital Information: A File Format Investigation by *Gregory W. Lawrence, William R. Kehoe, Oya Y. Rieger, William H. Walters, and Anne R. Kenney*. Washington, DC: CLIR, June 2000.
- Anne R. Kenney and Oya Y. Rieger, chairs, Establishing a Central Depository for Preserving Digital Image Collections - Part I: Responsibilities of Transferee, report of the Digital Preservation Policy Working Group, March 2001.
- Establishment of a Digital Preservation Officer, Aug. 2001
- Digitization now fully integrated with preservation efforts of Kroch Rare and Manuscript Collection (e.g. Samuel J. May Antislavery collection)

CUL's History in Digital Preservation, cont'd

- Moving Theory into Practice: Digital Imaging Tutorial

- Cornell's Digital Preservation Management Workshops and Tutorials, ongoing since 2003 http://www.library.cornell.edu/iris/tutorial/dpm/index.html

- Participation in numerous working groups and organizations

- Recipient of numerous grants for preservation activities

Cornell University Library

CUL Strategic Goal

In collaboration with CIT and external partners, establish within three years a fully functioning, administratively supported, and sustainable Open Archival Information System (OAIS) for managing, preserving, and providing ongoing access to Cornell University's digital assets over time, extending beyond the library to encompass university records of continuing value. This system must meet or exceed certification requirements currently being defined by an OAIS working group.

Cornell University Library

Team Charge

Develop an organizational context for the OAIS that will:

- Foster collaboration both within the university and beyond
- Seek to secure institutional commitment to this process that reaches the highest administrative levels of the university
- Work with national and international OAIS initiatives at other institutions whenever appropriate to define common goals and ensure system compatibility
- Develop, refine, and implement a three year plan to meet the above goals

Cornell University Library

Plan

- *Year 1*: Design and Plan the OAIS
- *Year 2*: Begin to Build the OAIS
- *Year 3*: Complete and populate the OAIS
- Years 4 and Beyond: sustain the system

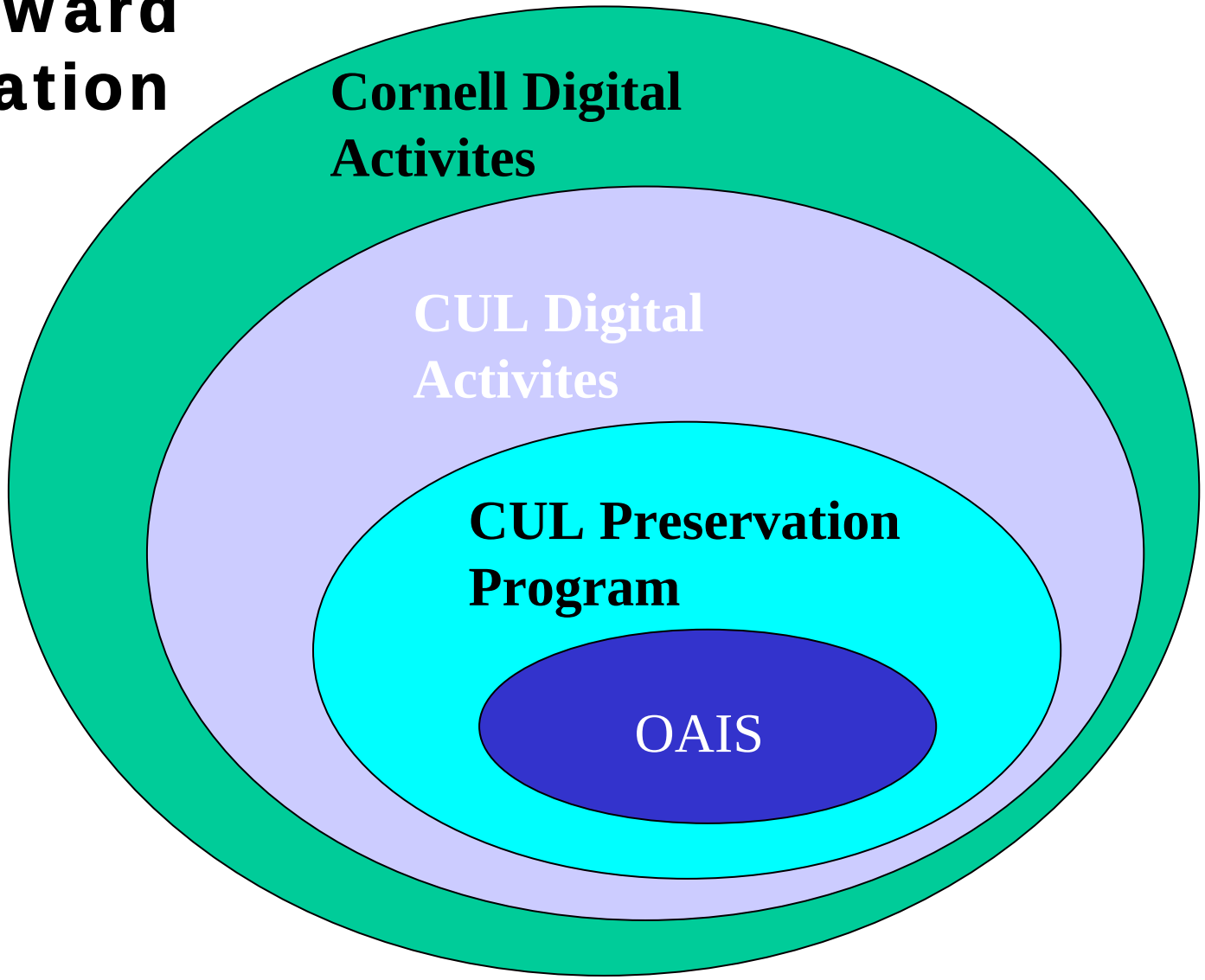
Cornell University Library

Progress—end of 2nd qtr

- Digital Preservation Policy Framework
- OAIS requirements review
  - CUL requirements
- Draft Selection Policy and procedures
- “Elevator Pitch”
- Gap analyses, in progress

Cornell University Library

Steps toward Preservation



Math Arc

Ensuring Access to Mathematics Over Time:

– Cooperative Management of Distributed Digital Archives

– Funded by NSF and DFG

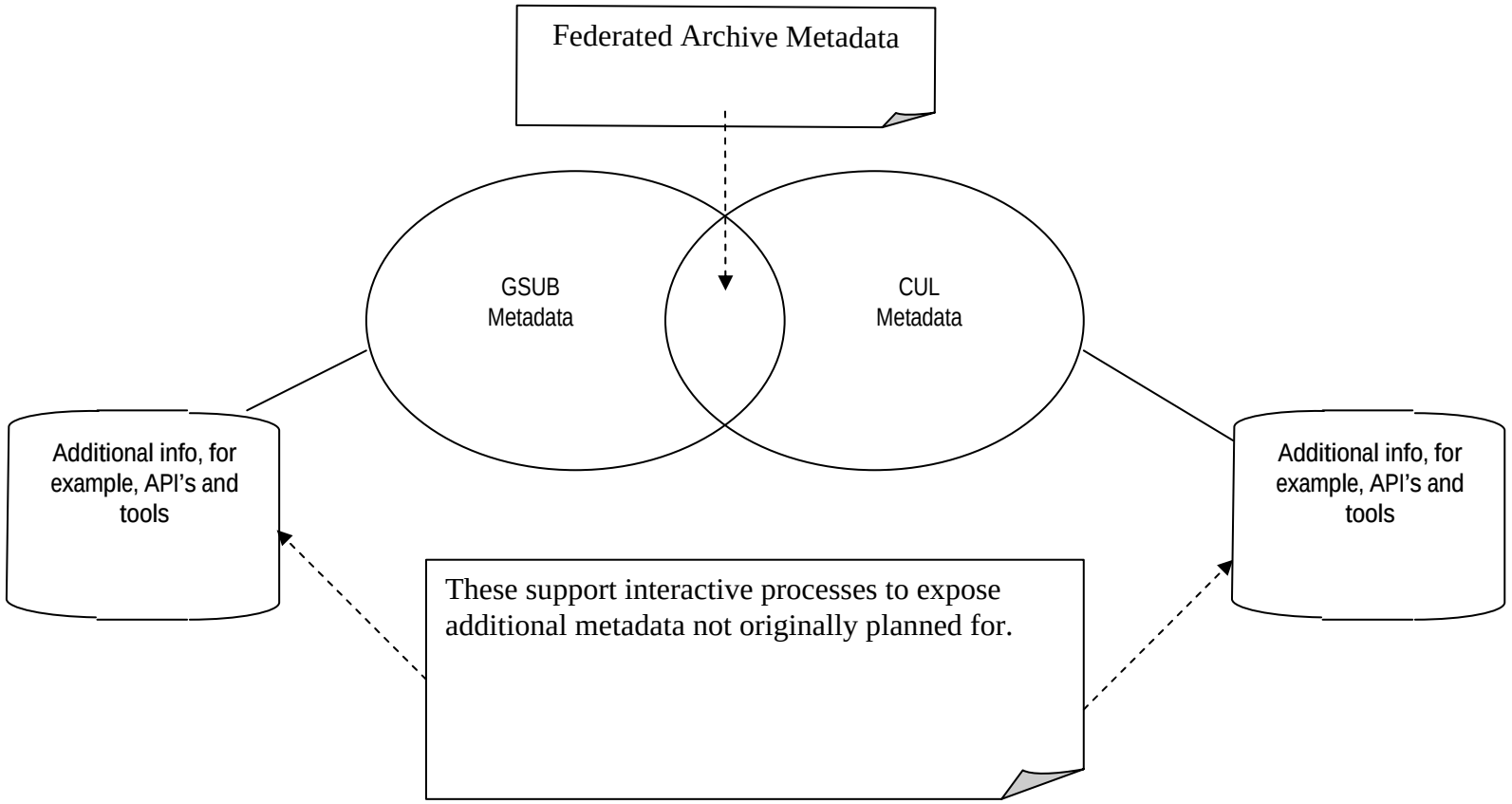
Math Arc

- develop an archive of serial mathematics literature that will be available to libraries worldwide and at the same time serve as a model for similar efforts in other disciplines within the library and publishing communities.

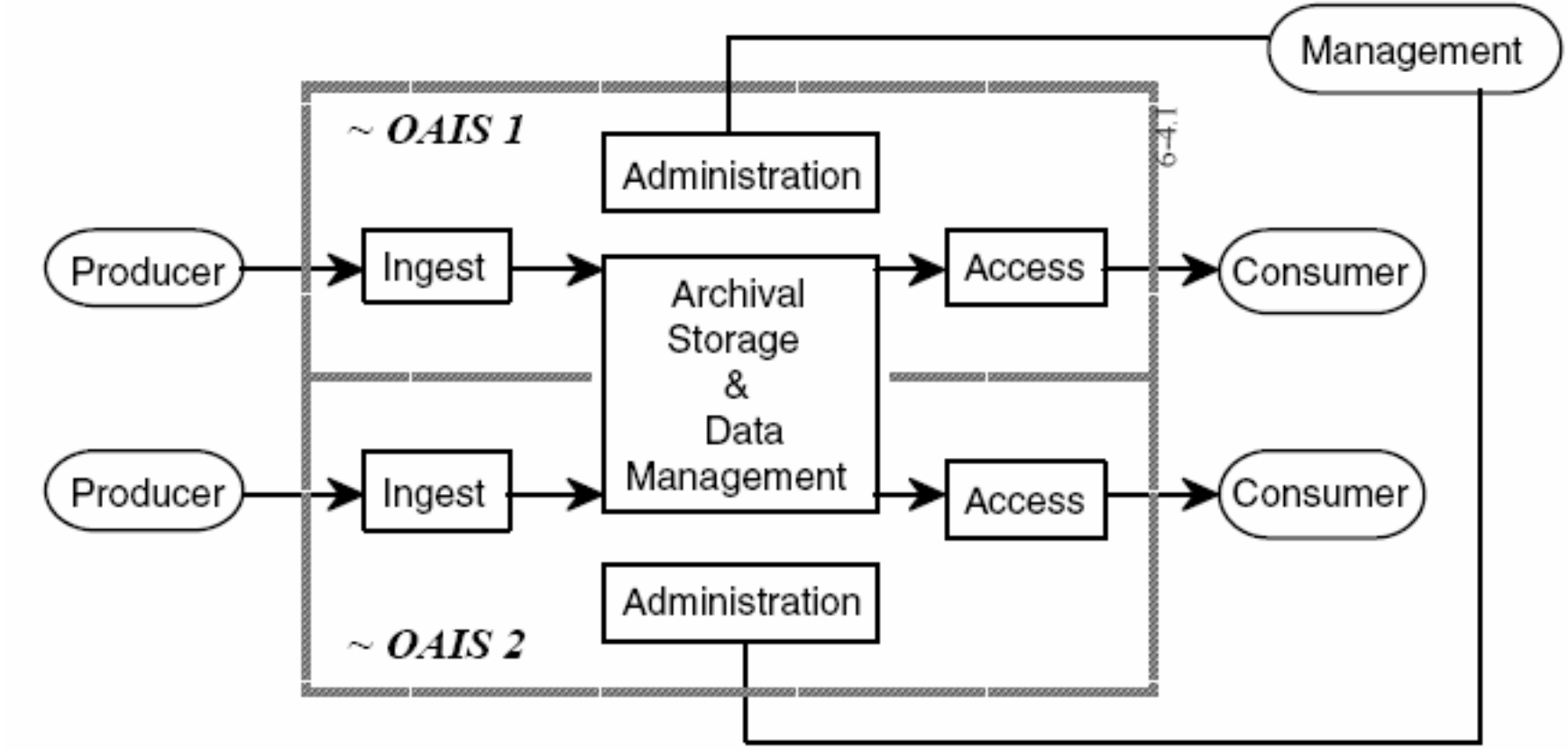
Math Arc
Progress to Date

- Stories – a page from Extreme Programming
- Message passing
- Metadata
- Journals, permissions, and ingest formats
- Agreements, principles, scenarios
- Asset Files, Registry

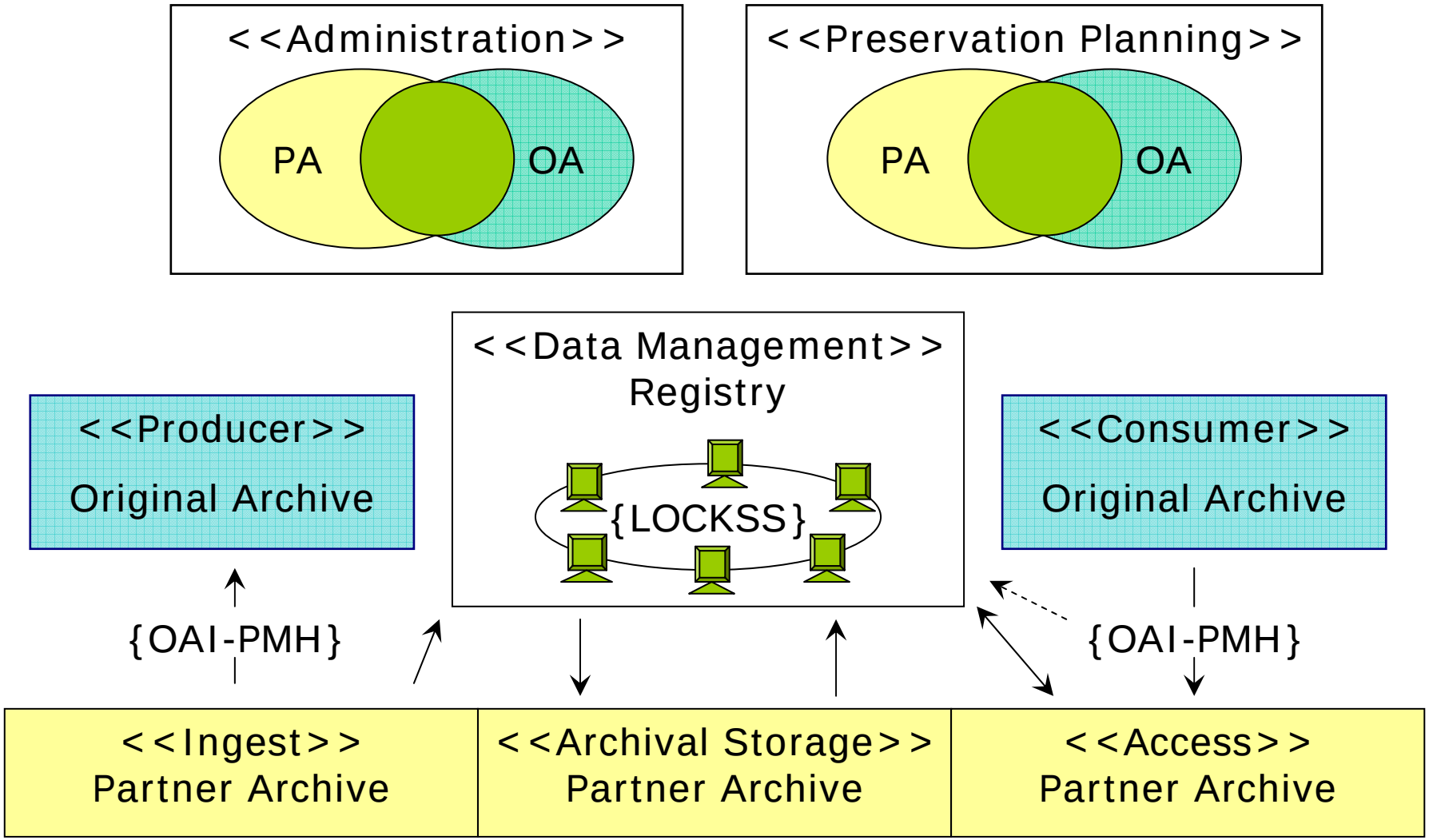
Cornell University Library



Frank Klaproth's conceptual diagram



Cornell University Library



Three Stories

1. Develop Packaging Designs and Migration Plans
2. Establish Standards and Policies
3. Develop Preservation Strategies and Standards
4. Administer Database
5. Co-ordinate Access Activities
6. Generate DIP

1. Receive Data
2. Co-ordinate Updates
3. Error Checking
4. Manage Storage Hierarchy

1. Receive Submission
2. Generate AIP
3. Quality Assurance
4. Archival Information Update
5. Audit Submission

Scenarios

- A new asset is added to one partner's site
- An Asset is migrated in the original archive (result is a new asset)
- An Asset is deleted intentionally in the original archive
- Update metadata for an asset in the original archive
- Update metadata for one or more content files in an asset in the original archive
- A partner drops out
- A new partner joins

Cornell University Library

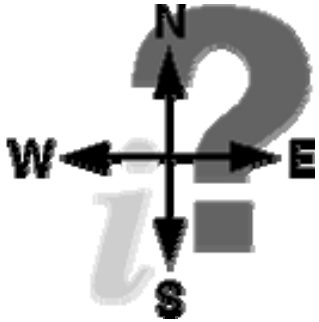
Scenarios, continued

- Disseminating content or assets
- Checking for inconsistencies on partner sites
- Checking for inconsistencies on own site
- Repairing inconsistencies
- New migration process becomes available
- Preservation plan/policy at any site is updated
- Request for resources cannot be met (temporarily or permanently)
- Rights or permissions change on any side

- Disseminating content or assets
- Checking for inconsistencies on partner sites
- Checking for inconsistencies on own site
- Repairing inconsistencies
- New migration process becomes available
- Preservation plan/policy at any site is updated
- Request for resources cannot be met (temporarily or permanently)
- Rights or permissions change on any side

Cornell University Library

Questions



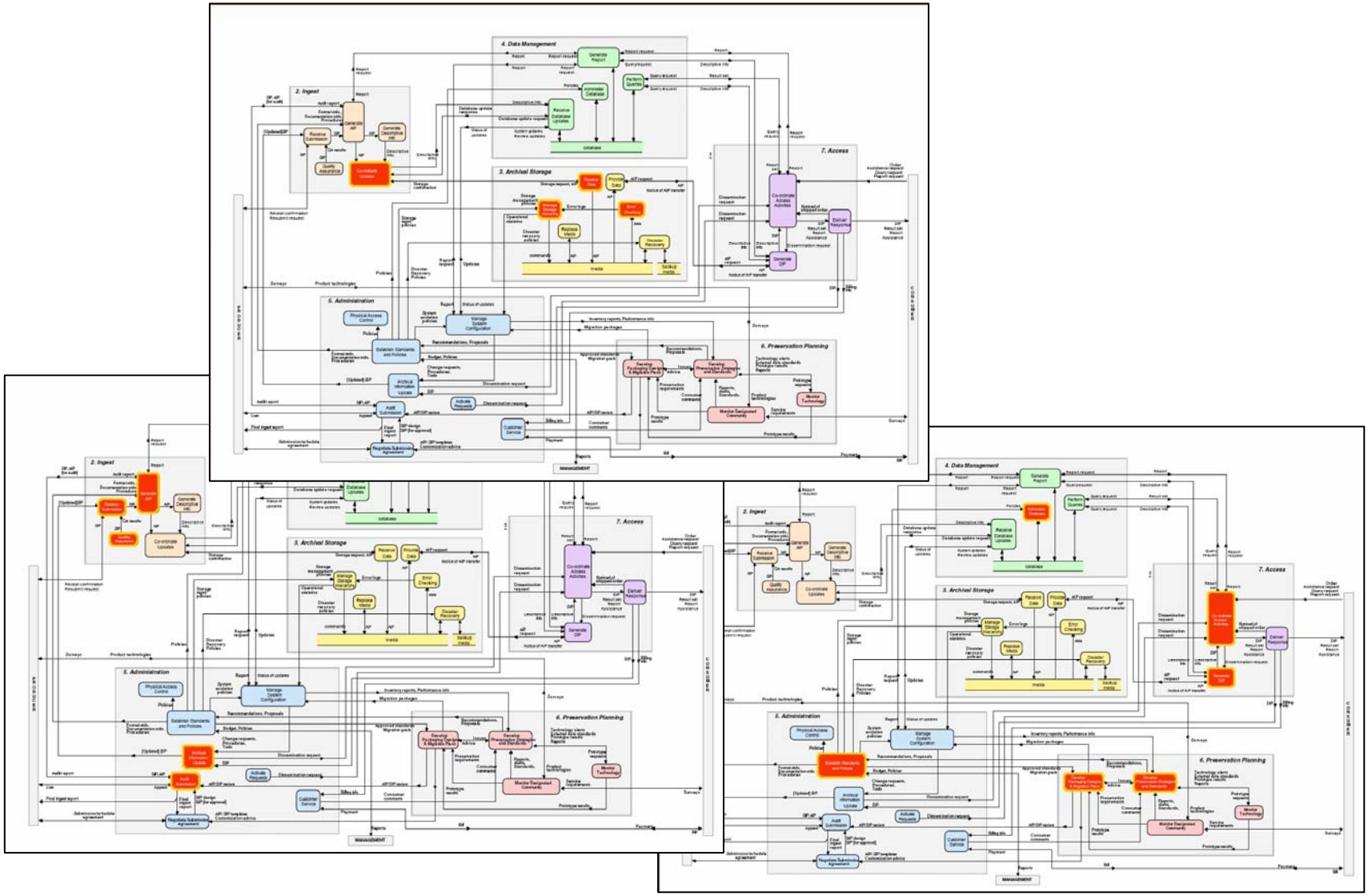
Thanks to Bill Kehoe

http://www.library.cornell.edu/dlit/MathArc

http://commondep.library.cornell.edu

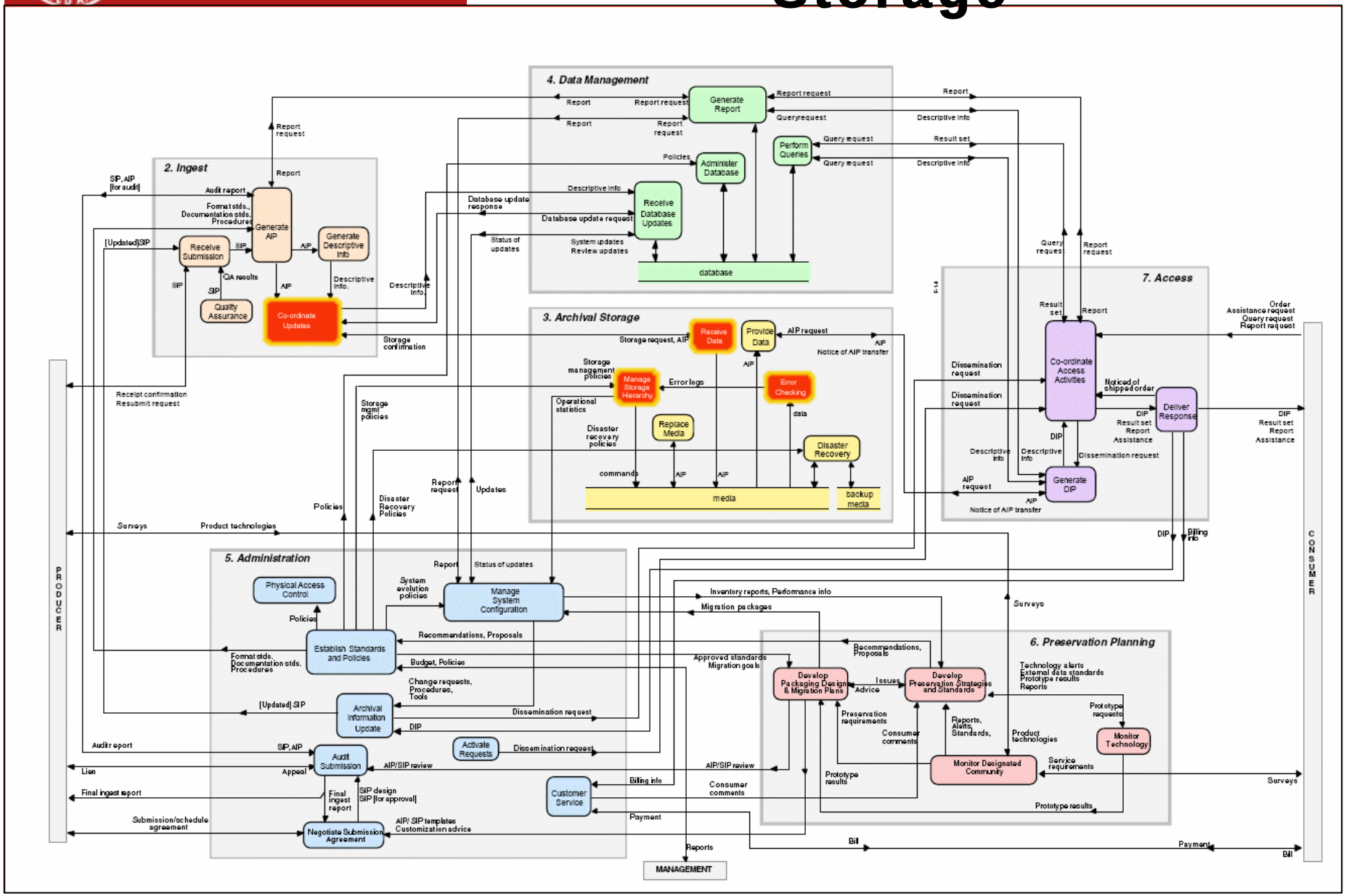
Cornell University Library

Three “Stories”



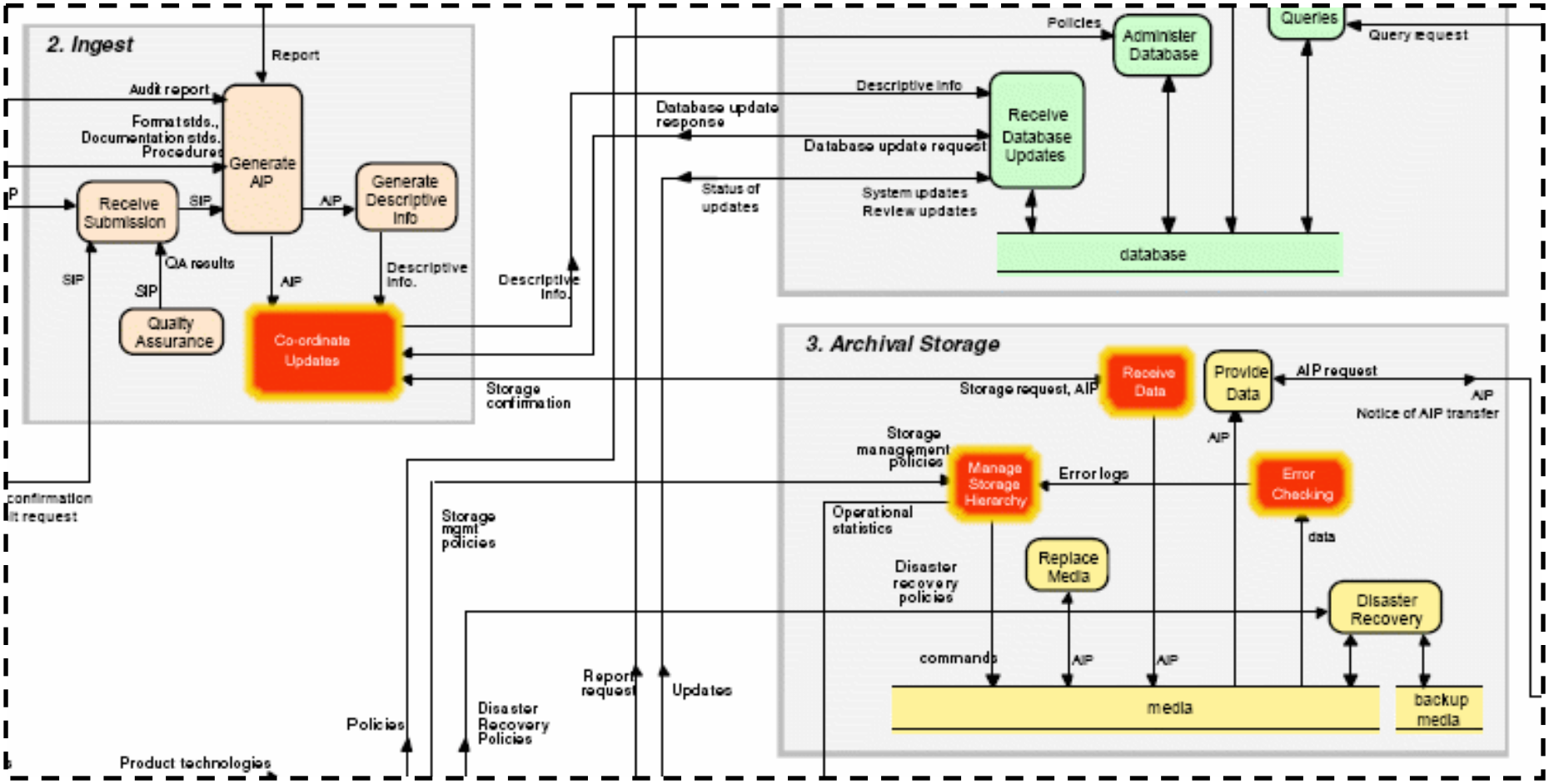
Cornell University Library

Story 2 – "Update Storage"



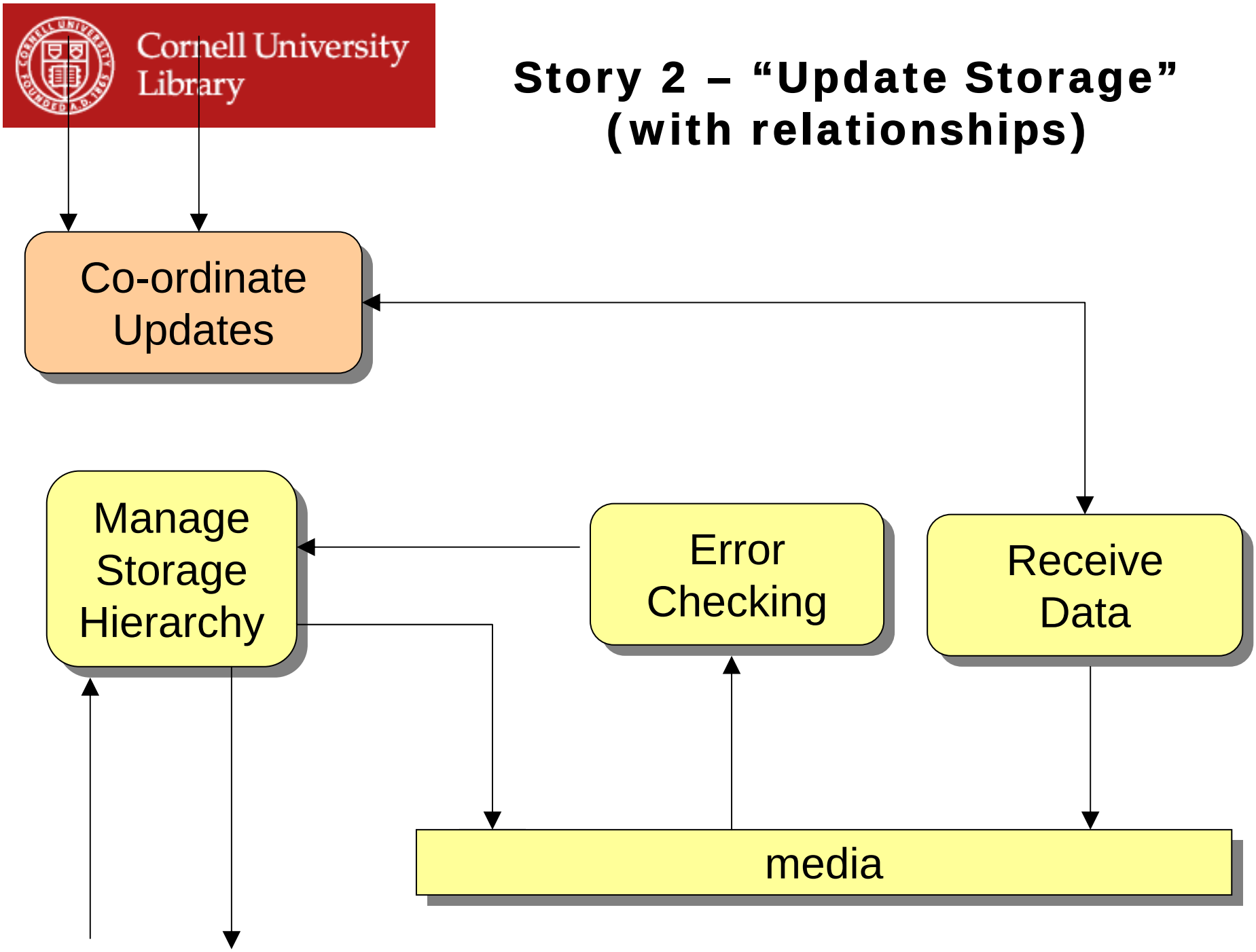
Cornell University Library

Story 2 – "Update Storage" (detail)



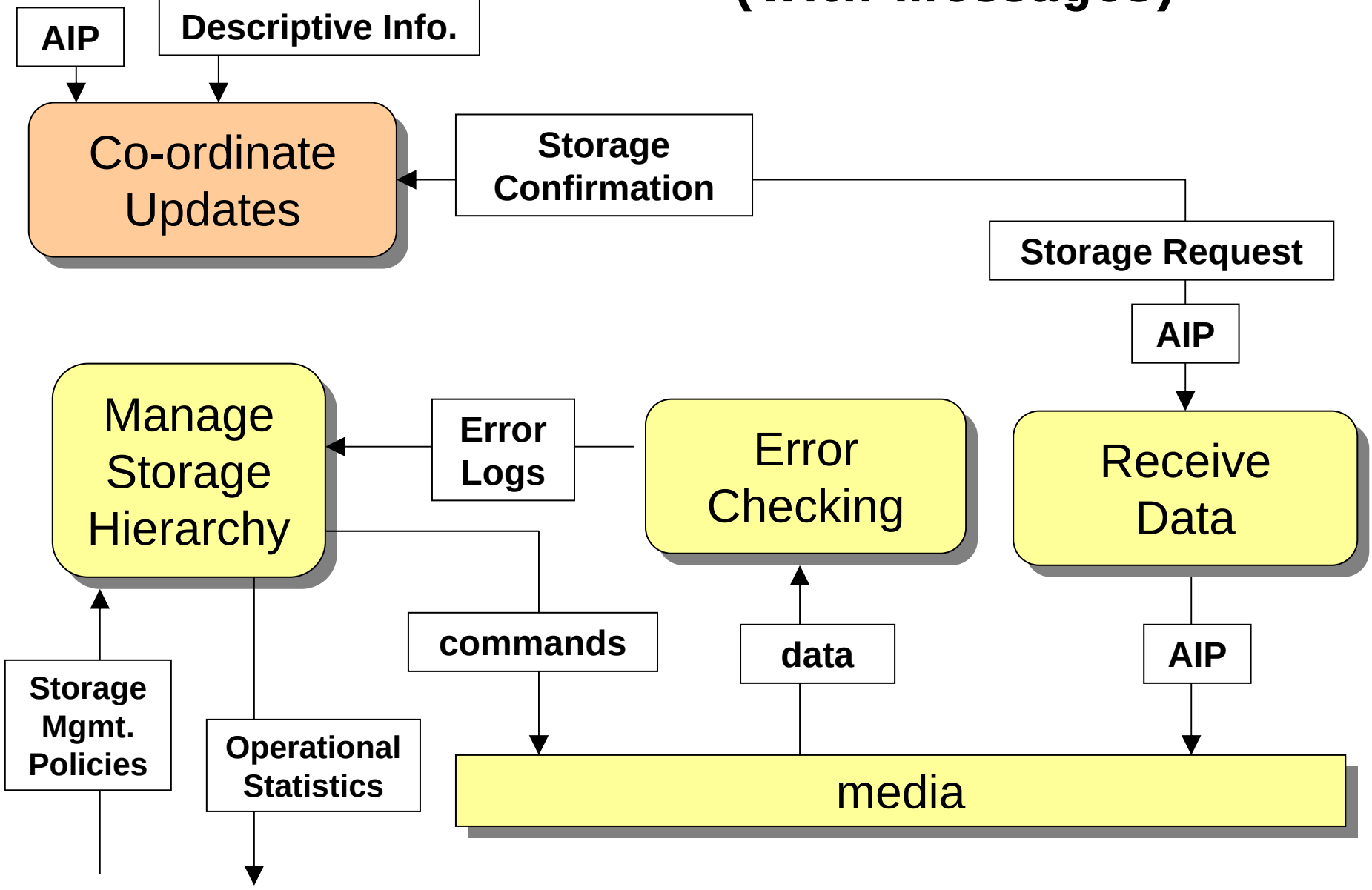
Cornell University Library

Story 2 – "Update Storage" (with relationships)



Cornell University Library

Story 2 – "Update Storage" (with messages)



Cornell University Library

Story 2 – "Update Storage" (with categorized messages)