Outline

- OCLC Digital Archive overview
- Influences on development decisions
- Preservation planning
OCLC Web Archiving Services

OCLC’s Connexion Service

User Services
- Metadata
- Harvest
- Ingest
- View
- Disseminate

WorldCat

Admin Services
- Admin

Digital Archive

System Processes
- Auditor
- Annotation

Staff
Discovery Services

- Local Catalog
- Local Catalog Service
- WorldCat
- OCLC's FirstSearch Service
- Digital Archive
- Google
- Researcher
- Web Page
Influences on Development Decisions

- Institution developing the archive
- Local depositor community
- Global digital archiving community
Influence: Archive & Depositors

- **OCLC**
  - International member cooperative
  - Members were pilot depositors
  - Corporate commitment

- **Local depository community**
  - Guided requirements
  - Confirmed following OAIS model
  - Emphasis on pre- and post-ingest processes
  - Influenced preliminary content object formats and types
Influence: Global Digital Archiving Community

- Provided inspiration and practical experience
- Source of evolving standards
- Opportunity for hands-on involvement with peers around the world
  - PREMIS
  - Task Force on Certification
- Ongoing access to an international knowledge-base
Decisions Influenced

- Object types and formats
- Access
- Preservation metadata
- Rights management capabilities
- Preservation planning
Decision: Object Types & Formats

- Initial formats
  - Text and still image
  - PDF, HTML, TXT, TIFF, GIF, JPEG, BMP

- Initial object types
  - Born-digital documents
Decision: Access
Decision: Preservation Metadata

- Built from what existed at the time
- Considered metadata in context
  - Types of objects archived
  - Element scope and purpose
  - Availability of the metadata at ingest
  - Function of the metadata in archive
Decision: Tools

- **OCLC's Connexion Service**
- **User Services**
  - Metadata
  - Harvest
  - Ingest
  - View
  - Disseminate
- **Admin Services**
  - Admin
- **System Processes**
  - Auditor
  - Annotation
- **WorldCat**
- **Digital Archive**
- **Staff**
Decision: Rights Management

- Local community
  - Long experience working with them
  - Assume right to harvest and preserve
  - Knowledgeable about copyright

- Types of objects archived
  - Preserve government publications
  - Provide access now and in future
Decision: Preservation Planning

- Local depositor community
  - Use of objects long-term
  - Significant properties
  - Review of policy

- Global digital archiving community
  - Format registry
  - Review of policy
Preservation Policy

- Provides a general outline of OCLC’s approach to digital preservation
- Specifies the list of digital formats currently accepted and how that list can change over time
- Specifies the preservation service levels
- Specifies what services are implemented to ensure uninterrupted access
- Defines technology watch service
- Defines how the policy itself changes
- Defines roles and responsibilities of participants
Preservation Planning

- Designed a strategy behind digital preservation processes
- How to identify the digital object formats best suited for preservation
- How to determine what computing platform offers the longest life-span expectancy
- How to measure preservation decisions
- What classes of metadata are necessary
- How to determine preservation actions
Preservation Planning Tools

- Designed a set of metrics necessary to assess the viability of any solution
- Designed the set of viable preservation actions
- Designed formal models to specify and document the computing platform necessary to render digital objects
- Designed processes to assess and prove authenticity of the digital objects
Preservation Planning Tools

- Designed a risk-based assessment methodology
- It is designed to look at all facets of the digital format specification itself and its dependencies
- Considered environment dependencies such as other specifications, software and hardware
- Considered associated organizations such as vendors, standard organizations, content owners, beneficiaries
Preservation Planning Tools

- Adopted the risk exposure calculation scales developed by Gregory Lawrence (Cornell Univ.)
- Identified several classes of potential reviewers, including computer scientists, format experts, librarians, preservation specialists, hardware vendors, lawyers, digital archive developers
- Developed a methodology to collect, analyze and compare the assessments
- Developed methods to relate assessment results to preservation actions
Preservation Actions

- The proposed strategy is to identify preservation actions by monitoring and reducing risk.
- Choose open source, open license and open specification software, hardware and formats.
- Identify and test alternative open source software tools and computing environments.
- Specify and test one current era commercial alternative, respecting the investment in technology already made by beneficiaries.
- Migrate to new formats when risk is too great.
In Conclusion

Questions and Comments

Contact:
Andreas_Stanescu@oclc.org
Leah_Houser@oclc.org
Pam_Kircher@oclc.org