A Tale of Two
Michigan State University
Digital Preservation Projects

Best Practices Exchange
September 29, 2010
Lisa Schmidt
lschmidt@ais.msu.edu
Michigan State University

- Est. 1855 by act of Michigan Legislature
- Nation’s pioneer land grant college
- Tier one research university with significant national and global impact
  - National Superconducting Cyclotron Lab
  - Center for Advancing Microbial Risk
  - Julian Samora Research Institute
- Leader in innovation and technology
Two Digital Preservation Projects

- H-Net E-Mail List Archives Preservation Project
- “Spartan Archive” Project
- Both funded by NHPRC
Overview

- H-Net E-Mail List Archives Preservation Project
- “Spartan Archive” Project
- Three points of comparison
  - Repository configuration
  - Digital preservation and curation tools
  - Administrative and technical support
- Conclusions
H-Net E-Mail Lists

- Part of H-Net: Humanities and Social Sciences Online scholarly consortium
- More than 180 networks (e-mail lists)
- Well-used
  - More than 130,000 unique subscribers
  - More than 1 million e-mail messages
- Hosted by MATRIX
H-Net E-Mail List Preservation Project

- NHPRC Electronic Records/Technologies Projects grant
- Awarded to MATRIX
- Grant award announced late 2006
- Two year project, 2007-2009
H-Net E-Mail List Preservation Project

- Conducted assessment of existing H-Net preservation policies and practices
- Applied OCLC/CRL TRAC checklist
- Developed and implemented an improved long-term preservation plan
  - Technical and administrative changes
  - Digital preservation policies
“Spartan Archive” Project

- “Spartan Archive: An Electronic Records Archive at Michigan State University”
- NHPRC Collaborative Electronic Records Projects grant
- Awarded to Michigan State University Archives & Historical Collections
- Grant award announced late 2009
- Three year project, 2010-2013
“Spartan Archive” Project

- Develop electronic records archives for university’s born-digital records
- Use proof-of-concept approach to test sustainability of archival solution for long-term access and preservation
“Spartan Archive” Project

- Test four electronic records series produced by Office of Registrar
  - Catalog of Academic Programs
  - Description of Courses
  - Annual Student Directory
  - Schedule of Courses
  - 71,200 logical data records
“Spartan Archive” Project

- Develop repository policies and procedures
  - Electronic records processing manual
  - Metadata guide
- Demonstrate accessibility of records using iRODS
- Extend to include other records of enduring value
Three Points of Comparison

- Repository configuration
- Digital preservation and curation tools
- Administrative and technical support
Repository: H-Net E-Mail List Archives

- “Repository” of LISTSERV-based e-mail lists stored on server
- “Living” archives
  - No separate preservation environment
  - Integrity checks, redundant backups, periodic archiving to tapes
- Not built as OAIS-compliant system
  - E-mail posting, storage, and access processes map to OAIS model
H-Net Ingest, Storage, and Retrieval
Repository: Spartan Archive

- Build from ground up as archival repository
- Will include preservation environment
  - Built on iRODS
- Designed as OAIS-compliant system
Tools: H-Net E-Mail List Archives

- Trusted Repositories Audit & Certification (TRAC): Criteria and Checklist
- Digital Preservation Policy Framework
TRAC

- OCLC/RLG published in February 2007
- Based on *Trusted Digital Repositories: Attributes and Responsibilities*
- Third-party certification or self assessment
- Three sections
  - Organizational Infrastructure
  - Digital Object Management
  - Technologies, Technical Infrastructure, & Security
- 84 criteria
TRAC

- “Gap Analysis”: Compare core audit criteria to local capabilities
- Formulate strategies to narrow gap and improve trustworthiness of repository
TRAC: H-Net E-Mail List Archives

- First analysis: February 2008
- Consulted with systems administrator, other MATRIX and H-Net staff
- Recommended technical and administrative improvements
  - Integrity checks
  - Documented policies, evidence of compliance
- Second analysis: July 2009
Digital Preservation Policy Framework

- Developed at Cornell University and ICPSR, 2007
- Combines organizational and technological perspectives for developing appropriate responses to challenges of digital preservation
Digital Preservation Policy Framework

- Addresses attributes of trusted digital repository, plus OAIS compliance
  - Administrative Responsibility
  - Organizational Viability
  - Financial Sustainability
  - Technological and Procedural Suitability
  - System Security
  - Procedural Accountability
Digital Preservation Policy Framework

- Presents high-level perspective of organization’s digital preservation program
- Reflects current capabilities
- Provides links to lower-level policies and procedures
- Points to near-term preservation plan
- Documents policy approval/maintenance process
Policies: H-Net E-Mail List Archives

- Documented and developed written digital preservation policies based on framework
  - Including “lower-level” policies and procedures
- Incorporated existing policies regarding H-Net and use of e-mail lists
  - Submission guidelines
  - Copyright
Tools: Spartan Archive

- TRAC
- Digital Preservation Policy Framework
- Archivists’ Toolkit
- Ingest tools
- Repository software
- Other digital preservation tools
TRAC: Spartan Archive

- Develop with TRAC compliance in mind
- Self-audit planned for third year of project
Policies: Spartan Archive

- Processes and procedures will be documented during course of project
- Policies will be developed with Digital Preservation Policy Framework in mind
Archivists’ Toolkit: Spartan Archive

- Collection management system chosen by University Archives
- Grant includes AT implementation
- Common web interface with Spartan Archive
Ingest Tools: Spartan Archive

- Tufts Accessioning Program for Electronic Records (TAPER)
  - Automated accessioning
  - Submission agreements
- Software-Independent Archiving of Relational Databases (SIARD)
- BagIt
Repository Software: Spartan Archive

- Fedora
  - Hydra content management framework
- DSpace
  - Ingest workbenches
- Archivematica
Other Tools: Spartan Archive

- iRODS
- Metadata extraction
- Integrity checking
- Web interface
Administrative & Technical Support

- Project support
- Ongoing support beyond project
Support: H-Net E-Mail List Archives

- Outside of MATRIX’s main area of focus
- Different leadership
- Issues of institutional inertia
  - Mature system working well
- Challenge of imparting value
- Question of ongoing attention to preservation
Support: Spartan Archive

- Addresses need of university
- Falls under MSU Archives mandate
- Top level leadership support
- Build from the ground up
  - No old system inertia
- Dedicated technical resources
- Will receive ongoing support
Conclusions: H-Net E-Mail Lists

- **Opportunities**
  - Tried out digital preservation tools
  - Creative approaches and applications

- **Challenges**
  - Grafting of preservation onto existing system
  - Scattered support
Conclusions: Spartan Archive

- **Opportunities**
  - Building from the ground up
  - Solid support

- **Challenges**
  - Lots of choices
  - High importance, high visibility, higher stakes
Michigan State University Archives

“Belle Sarcastic”
Discussion

- Experience with tools
  - TRAC, Digital Preservation Policy Framework, TAPER, SIARD, others

- How tools were chosen

- Experience grafting preservation onto existing system
  - Especially “living” archive

- Support issues
References

References

References

- Integrated Rules-Oriented Data System (iRODS), https://www.irods.org/
- Michigan State University Archives & Historical Collections, http://www.archives.msu.edu
References

- Software-Independent Archiving of Relational Databases (SIARD),
  http://www.planets-project.eu/docs/presentations/ICA2008_Comment_SIARD.pdf

- Spartan Archive,
  http://www.archives.msu.edu/about/spartan_archive.php

- Trusted Digital Repositories: Attributes and Responsibilities,
References

- Tufts Accessioning Program for Electronic Records (TAPER), [http://dca.tufts.edu/?pid=49](http://dca.tufts.edu/?pid=49)