Repository Institutionalization:
What makes it worth becoming infrastructure?

Patricia Galloway

School of Information

University of Texas at Austin

Social engineering: seed--evolve--reseed

- Gerhard Fischer, U. Of Colorado
- Collaborative system modeling concept, to support social creativity
  - Prototype
  - Practice
  - Revision
- Applicability to our case
  - Digital archiving as an emergent requirement: applications and standards in "continuous beta"
  - Departmental needs for secure support for individual and collaborative creativity

Seed--evolve--reseed in action

- Students' early class work: developing proposals and policies (2003-2004)
- First year of repository: experiments with file types, metadata standards, digital archaeology (spring 2005)
- Evolution: software upgrade; policy improvement (summer/fall 2005)
- Second year of repository: extend (and restructure) some projects, begin new ones (2006)
- Evolution: software upgrade; metadata refinement (fall 2006)

Stakeholder environment

- ◆ Faculty
  - Need secure storage
  - Want prompt exposure of work
- ◆ Administration/Staff
  - Need flexibility of archival storage of official papers
- ◆ Students
  - Want exposure of professional work
- ◆ IT infrastructure
  - Need ability to archive tools
  - Need versioning repository

Filling explicit actual needs

- Obeying records schedules
- Providing faculty work exposure and versioning store
- Providing students with the same
- Providing secure storage for learning objects
- Providing school's IT with secure archive for saved objects
- Providing an environment for archives students to put theory into practice
- Providing archival preservation for materials of historical value

Obeying records schedules

- University of Texas is public university
- Recordkeeping subject to the laws of the state of Texas and other applicable laws
- Recordkeeping practices and schedules mandated by UT System
- Some permanent record material now exists only in digital form (e.g. on School of Information website)
- And besides: we teach records management!

Providing faculty services

- Preprints/postprints exposure to maximize citation
- Secure storage of materials that document research process
- Permanent storage of materials that document faculty careers
- Permanent storage of data collected in the course of research (UT mandate)

Providing student services

◆ Publications (e.g. *The Cochineal*)

◆ Portfolios and other online-visible work

- Conservation students' multimedia treatment portfolios
- IS students' tutorials
- Student websites (potential)

Secure storage for learning objects

- Laboratory tutorials (learning objects now, historical documents later)
- Faculty syllabi, presentations, etc.
- Over time, many things gathered for other purposes become learning objects
- [And of course the whole repository is a learning object]

Providing IT with secure storage

- Permanent repository for locally written and legitimately downloaded code constituting dependencies in local system
- Quasi-permanent repository for major revisions of website

Providing a learning environment

- Problems in Permanent Retention of Electronic Records (INF 392K): since 2003
- Lifecycle Metadata for Digital Objects (INF 389K): beginning this fall
- Student individual study and capstone projects
- Other courses (Digital Libraries, Digitization)

Archives and historical objects: An aside on "repository"

- Digital libraries: permanence unresolved
- E-prints repositories: permanence not the main focus
- Digital archives: permanence a commitment and an ongoing negotiation
- "Departmental Institutional Repository": elements of all three

Costs by budget category

Category	Description	Annual Amount	UT Split	MIT Split[3]
Staff Salary, Benefits and Overhead	Salary, benefits and overhead (Systems Administration 5% of Time or 100 hours per year)	6K	84%	85%
Hardware	Cost system $ 2K, plus $1K expansion and maintenance with 5 year life [($2K + $1K)/5 years = ~$1K]	1K	16%	15%
Software	All open source	0K	0%	0%
Yearly Total		**7K**	**100%**	**100%**

Usage as of 6/22/2006

Title Count: Collection items	439
Title Count: Documentation items	754
Author Count	308
Average File Size	30.2 Kb
Size	36000 Kb

Collection communities

Name	Description
Digital Archive Projects	This community is an umbrella for class and individual projects undertaken by students in the School of Information. Nine testbed communities including Austin History Center, City of Austin and four iSchool courses working papers.
School of Information Administration	Policy information on retention of iSchool website, webpage retention, and iSchool Technical Services archiving of tutorials. Students engineered the archival process for this community.
iSchool Faculty	Faculty scholarship. Two current faculty and three retired faculty Students engineered the archival process for this community.
Harry Ransom Humanities Research Center	*Joyce Project*, items in this collection were created in the process of archiving the files of Michael Joyce. Joyce is a prize-winning author as well as a teacher of writing. Ongoing project is handled by alumna of iSchool who took course "Problems of Retention of Electronic Records" while a working archivist to deepen her skills for processing digital collections, continues to generate individual student projects.
Kilgarlin Center	*The Cochineal*, an online journal for student work conducted at the University of Texas Kilgarlin Center for Preservation of the Cultural Record.

Advantages of local repository

◆ Community of practice makes communication (and trust) much easier, particularly in present fraught IP environment

◆ Service learning element in student assistance to faculty is a strong benefit
- "Teachable moments"
- Knowledge continuity
- Problem solving

◆ Informal deposit plus formal preservation aids in timely capture

Steps to institutionalization

- IT/faculty/student working group on digital repositories: archives, KM, HCI interests
- Faculty/staff/student management committee for repository
- Ongoing collaborations to support student projects
- Commitment to RA/TA 10-hour position for ongoing support
- Discussions with General Libraries about federating departmental repositories and university archives

Thanks and questions

- Thanks to Ingenta and the ALA Research Round Table for a grant to support the investigation of these questions
- Thanks to Larry Stewart and Maria Esteva for their research on repository cost modeling and citation tracking
- Thanks to School of Information IT services staff for active participation in these discussions
- Thanks to Bonnie Nardi at UC/Irvine and Nancy Foster at Rochester for sharing their faculty work-practice studies