Shaping Digital Preservation Programs in Research Libraries

March, 2003
Theme 1

In the digital realm, it’s as much about **constraints** as it is about selection.
Theme 2

What gets preserved will reflect who is does the preserving
To put it another way

* Motivation
  (who wants to play)
  (what do they want to do)

* Ability
  (what can they do)
Theme 3

Constraints can be loosened through sharing

But does that sharing require structure?
Topics

1. The issue of scale
2. The issue of control
3. The issue of format
4. The influence of relationships
5. Roles and passions
6. The influence of unique collections
7. The influence of national programs
1. The issue of scale

- If you are doing small things, you can:
  - Get by with chewing gum and baling wire, or
  - Buy services

- Large-scale programs will need expensive things:
  - Technical infrastructure
  - Expertise
Technical infrastructure

- Formal repository
- Well defined metadata for preservation and management
- Ingestion facilities
  - with sophisticated quality control
- Access management
- Persistent identifier system
- Preservation framework (migration, emulation)
Technical infrastructure

* Turnkey solutions are lacking ("digital library-in-a-box" products oblivious to preservation issues)

* In the absence of commercial solutions, can we help each other developing the tools?
Expertise

• A serious program requires a significant investment in sophisticated staff
  – Technical formats
  – Metadata
  – Rendering applications
  – Large scale storage
  – Etc., etc., etc.
Expertise

Potentially so easy to share…

But the need for continuity of program over very long time-frames may require some formality in sharing
Scaling

Unless infrastructure becomes easy and affordable, few will play at scale

The more players, the better overall program (*diversity of approaches, interests, priorities adds robustness*)
2. The issue of control

* Physical control: it’s hard to preserve what you don’t hold

* Legal control: you need the right to preserve objects
Preserving what’s already in hand

- It’s instinctive
  - Replicates the physical world: if I have it, I will preserve it
- It’s easier
  - Remote resources will take a lot more effort technically and legally
Most of what we preserve is owned by others

Many will care about what we do to their creations!

("migrations", for instance)
Thus, the natural beginning...

- Most programs begin with the institution’s own digital resources
  - Products of library digitization programs
  - University records
  - Institutional repositories
  - University press publications
  - Course web sites
3. The issue of format

- Digital preservation is a format-by-format issue
  - Cost and complexity as much related to the number of formats involved as the number of objects

- Formats determine
  - Expertise required
  - Technical facilities needed
• Programs likely constrained by format
  – If more than a bit-store, may need to restrict formats accepted
• Uncomfortable
  – Technical format is orthogonal to traditional concerns of topic and publication format (books, serials, etc.)
This could change...

*Commercial services* or *shared expertise and technology* could mitigate the constraint of format.
4. The influence of relationships

- Trust is key to a digital preservation program
  - Will the IP owner trust the preserving institution?
  - Will the affected user community trust the preserving institution?
- Note the influence of using community on IP owners
  - As authors and editors
  - Through discipline norms
Strong relationships with disciplines or IP producers can:
* create the trust needed to allow the archiving and preservation of resources
* create the expectation/demand for a preservation effort
5. Roles and passions

- Digital preservation is hard
  - Serious program will consume significant amounts of money, attention, and talent

- Digital preservation is the un-glamorous brother of digital services
  - Digital services yield pats-on-the-back, faculty praise, smiling provosts
  - Digital preservation frequently yields invisible effects
Justifying & maintaining programs will be hard

• Programs likely to develop where
  – Strong institutional commitment to preserving research resources
  – Curators & administrators have enough understanding and passion to overcome the obstacles

• Specifics of what is preserved will reflect the specifics of those institutions and individuals
7. The influence of unique collections

- Existing strong programs of collecting contemporary primary materials may lead institutions to digital preservation
- The “digital gap” in collections of contemporary arts, political, or social documentation will become increasingly obvious
- Ephemeral nature of the materials likely to generate a sense of urgency
Non-unique collections

The same pressures unlikely in areas widely “collected”:

* curatorial intensity is lower

* “the other guy will do it”
8. The influence of national programs

Observation 1

Contributing to a shared national program is a powerful motivator.

The existence of a national program can generate institutional activity.
National programs

Observation 2

A dangerous gap: remote resources, widely used, for which everybody feels some responsibility, but no one has a particular responsibility.

Natural focus for a shared national program?
In summary….

1. Digital preservation programs are hard and costly. There will be few large players.
2. What gets preserved will be constrained by rights, remote storage, and format.
3. Institutional priorities will be influenced by relationships, perception of institutional role, and the understanding and passion of key individuals.
4. Serious programs of sharing technical solutions and expertise can encourage more players.
In summary...

5. Natural priorities will be
   * things created within the institution
   * other objects held locally
   * digital resources related to existing collections of primary materials

6. National programs can motivate new activity

7. Widely used, remotely stored materials are problematic, and are a prime area for shared programs