Selection Strategies for Digital Institutional Repositories

Kent Woynowski
30 September 2004
1. Basic Selection Guidelines

2. Born-digital Materials

3. Traditional Archival Materials
Basic Selection Guidelines

- Should be basically the same as selection guidelines for print resources
Basic Guidelines

- **Scope**
  - Document the history of university and the activities of its faculty, staff, and students
  - Support teaching and scholarship at the university
Basic Guidelines

Research Value and Authority

- High priority should be given to items with long-term research value
- Innovative materials that may have future research value
- Non-authoritative materials should not be selected, but a sampling may be kept
Basic Guidelines

- Parameters
  - Break large sites into components that meet collection guidelines
  - If smaller components cannot stand on their own, but together form a valuable source of information, they should be archived in their entirety
Basic Guidelines

Priorities and Criteria
- Represents a complete version
- Represents a thematic aggregate
- Creates a comprehensive collection
- Enhances access to a collection
- Increases the use of the collection
- Helps preserve and protect materials
- Enables new kinds of research
- Does not duplicate resources already available
Born-Digital Materials

- Adapting the current collection policy
- Self-selection
Adapting the Current Policy

- Institutional Archives/Personal Papers/Organizational Records
- Visual Materials
- Theses and Dissertations
- Faculty Publications
Self-Selection

- DSpace allows content creators to select their own materials that will make up their “community”
- Removes a burden from the archives, but communities must still be checked for quality
Encouraging Participation

- Faculty (research and publishing)
  - Enhanced professional visibility
  - Higher citation rates with indexing and search capabilities
  - Register priority of ideas
  - Grassroots practice of posting research online already exists
Encouraging Participation

- Faculty (teaching)
  - Repository can include non-ephemeral teaching materials
    - Illustrations
    - Videos
    - Models
Encouraging Participation

- Academic Departments
  - Bulletins and journals online can increase visibility and status

- Students
  - Already being done, with submission of theses
Selection of Websites

- Selective archiving of static web resources
- Selective archiving of static and dynamic web resources
- Whole domain harvesting
- Combination of selective and whole domain approaches
Traditional Archival Materials

- Adapting the current collection policy
- Developing selection guidelines
Adapting the Current Policy

- Personal Papers
- Visual Materials
- Rare Books
- Science and Community
- Southeastern Textile Industry Records
- Faculty Publications
Selection Criteria & Guidelines

- From *Profiles in Science*:
  - Formal
  - Substantive
Formal Criteria

For a typical scientific collection:

1. Documents of all types of which the individual is the sole author
   a. Include drafts of documents if they differ significantly from the published version
   b. Correspondence
      1. Initiated by the subject
      2. Involves more than one exchange
      3. Longer than one page
   c. Other documents
      1. Full sentences, except for notes
Formal Criteria

2. Documents of all types of which the individual is the principal or co-author.
3. Documents written by others under the individual’s supervision.
4. Documents authored by others that recount oral statements made by the individual.
Formal Criteria

5. Documents authored by others devoted to individual and his work
6. Documents authored by others and addressed to others, with copies supplied to the individual.
7. Documents written by others collected by the individual for his/her own information.
Substantive Criteria

- Within each formal category

1. Scientific interests, projects, and production
2. Public Life
3. Professional Career
4. Private Life
Substantive Criteria

- Within each substantive category, unpublished rather than published documents receive priority.
Other Considerations

- Copyright
- Quality control
- Tracking systems
- Sensitivity
- Providing context
Strategies

- Identify collecting areas for new born-digital materials
- Encourage participation
- Establish workflow and training programs for self-selection
- Develop criteria for selection from traditional collections