DSpace the “K” Way

Heidi Butler, Kathryn Lightcap, and Stacy Nowicki
NITLE DSpace User Community Meeting
University of Puget Sound, 11 June 2008
K College Process

1. Establish administrators
2. Define what the IR is, what it’s for
3. Write policies/“ideal” workflow plan
4. Write directions (“how to” materials)
5. Start practicing!
Eventually…

Marketing + Roll-out = Party time!
Starting Point

• Policies
  – Clarifies and justifies
  – Establishes guidelines
  – Plans procedures
  – Efficiency
  – Time management
  – Diversify workload
  – Get input and create “buzz”
People Involved

- College Archivist
- Information Technology staff
- Cataloger
- Dean of Students
- Administrative Assistants
- Faculty
More Documents

• Metadata charts
  – Defines required or optional elements
  – Communities, collections, items

• Instructions
  – Instructions supplied by DSpace too technical
  – “How to” for submitting, cataloging, managing

• Non-Exclusive Distribution License
  – Authors retain copyright, but K College can upload and distribute
Copyright?

• Students used to sign paper License Agreement

• Registrar and Dean of Students say:
  – Don’t need student permission to post theses
  – Not against FERPA
  – Clause in Academic Catalog
DSpace Content [539]

- **Student Work**
  - Theses (SIPs)
  - Presentation posters
  - Limited to current students/faculty/staff

- **Archival Materials**
  - College histories – searchable!
  - Photos

- **Presentations (Faculty/Staff)**
Workflow

- Student (or IS staff!) scans material
- Students (or IS staff!) uploads material
  - Adds some metadata (title, author, abstract)
- Material available immediately
- Cataloger adds detailed metadata (LCSH, dates, etc.)
Plan v. Reality

Plan: Use DSpace for access and preservation

Reality: Preservation not likely
- Pilot project’s 100 GB down to 50 GB in June 2008
- Too many big TIFF files for our storage capacity
Plan v. Reality

Plan: Uploading done by those who take responsibility for a “Community”

Reality: Most uploading done by IS staff and students
Plan v. Reality

Plan: Administrative Assistants responsible for student theses in each department

Reality: We’re still learning the system ourselves!
Heidi’s Turn

Heidi Butler,
College Archivist
About the “K” Archives

- Staffed part-time + students only
- College anniversary = special funds for digitization
- Collection includes at least 25,000 images
- Just beginning digital initiatives
Archives Materials in DSpace

- Photographs
- Historical documents
- New materials (most “born digital”) that would be archived in traditional format anyway
  - Talks, programs, etc. from anniversary events
  - Electronic version of new college history book
Choosing Materials

- "Index" information—researchers love it
- Use knowledge of what's popular
- Keep track of interesting lesser-used images when reorganizing collection
- Choose based on existing descriptive info
- Ease of digitizing non-photo materials
  - If it's a duplicate, we can chop off the spine!
  - Odd sizes, fragile items "for later"
Preservation Standards

- Photographs: 600dpi, unretouched/uncompressed TIF
  - We make JPEG derivatives for DSpace (and let DSpace make the thumbnails)
  - TIF files will go on CD/DVD backup and network tape drive
- Documents: PDF/A?
- Audio/video files: no idea!
Good Things

- We’re in OAISter
- Dublin Core & image capabilities = no need to purchase ContentDM, other Archival tools for special collections
- Thumbnail generation
- RSS and email subscriptions
- Local/regional peers don’t have IRs yet so we are cool
Collection Examples
Kathryn’s Turn

Kathryn Lightcap,
Graphics and Multimedia Designer
Types of Scanners

• HP Scanjet 7650 (ADF scanner)
  – Items up to 8.5”x14” through ADF
  – Items up to 8.5”x11” on flatbed

• Epson Xpression 100000XL (flatbed scanner)
  – Items up to 12”x17”

• Epson Xpression 1680 (flatbed scanner)
  – Items up to 8.5”x12”
Items Scanned

• Old Senior Thesis
  – Files are encrypted
  – Not PDF/A
• Photographs
• Yearbooks
• Misc documents
File Formats

- PDFs
  - Documents/books
  - Born digital files
- TIFF
- JPEG
  - Use for photographic images/images with smooth tone/color variations (lossy compression)
- GIF
  - Use for line art (lossless compression)
Student Worker Training

• File type education
  – When to use / for what
  – What is a PDF, JPEG, TIFF, GIF?

• Students do not understand the benefits/drawbacks for each of these files
Training – Scanning Images

• TIFF images converted to JPEGs in Photoshop
  – Batch conversion or use droplet
• Limited DSpace disc space / don’t upload TIFFs
• Images – 300 dpi
Training - Scanning Documents

• Senior thesis
  – Password protect thesis with Adobe Acrobat Pro 8
  – Renumber PDF pages
  – Upload DSpace
Problems / Issues

• OCR scanners jammed
  – Heavyweight paper and rice paper = bad
• Older thesis with pics/graphs/tables glued to pages
  – Graphics are barely attached with old and brittle adhesive
  – Scan these pages separately on flatbed and insert into PDF document
• Password protecting the thesis files
  – Give student workers the password that encrypts these files?
More Problems / Issues

• Student workflow
  – Permission to upload to DSpace but they do not have rights to modify their submissions

• Computer and hardware issues:
  – OCR scanner at Circulation Desk
    • Fast OCR scanner
    • Renders searchable PDF slowly
    • All-in-one computer with reasonable memory
  – OCR scanner in Center for New Media Design
    • Not fast OCR scanner
    • Creates searchable PDF reasonably fast
    • Desktop computer, more powerful/more memory
PDF/A Compatibility

- Not allowed in document
  - Audio and video content
  - Encryption
  - Javascript and .exe file launches
- Any fonts used must be legally embedded
- Colorspace must be specified
- Standards-based metadata is mandated
Born Digital Thesis Posters

- 7 years of 36”x48” PowerPoint posters
- Binders filled with 8.5”x11” proofs of thesis posters
- Convert to PDF files
- File Naming Convention
  - Name_SymposiumYear
- At thesis poster training session, I showed students that they could look at previous years
  - Very successful! 😊