Digital Project Management

Jenn Riley
Metadata Librarian
Indiana University Digital Library Program
You are not alone.

  - Collections (organized groups of objects)
  - Objects (digital materials)
  - Metadata (information about objects and collections)
  - Initiatives (programs or projects to create and manage collections)
- Principles from this framework integrated into this presentation
Achieving good collections

- Curated, cohesive group of materials
- Adherence to standards and best practices
- We must do things the way that others do them
- We must use commons systems or create metadata in ways that allow us to share

*Summary by Kristine Brancolini*
Collaboration

• Many people must actively work together to implement a digital project
• We must collaborate within and between institutions
• Collaboration requires work!
  ▫ Managing expectations and personalities
  ▫ Understanding each other’s perspectives
  ▫ Group decision-making
Communication and documentation

• Effective collaboration requires open communication
• Consensus may not always be necessary
  ▫ But allow all participants the opportunity to present their views
  ▫ Give them adequate time to formulate opinions
• Documentation is key, but don’t expect people to read it
Why is all of this necessary?

Sustainability.
Digital Project Workflow

PLAN → DIGITIZE → STORE → ASSESS → DESCRIBE → DELIVER → PLAN
Planning is the most important stage.

• Clearly articulate your goals – the entire project plan extends from this
  ▫ Envision what services you will provide
  ▫ Define your primary user groups

• Rights issues
• Timelines
• Funding
Timelines

• Should be clearly articulated, even for ongoing work
• Will need constant revision
• Build in flexibility
  ▫ Staff turnover
  ▫ Even average production rates aren’t average
• Be aware of dependencies
Funding

- Will be discussed tomorrow
- Funding model depends on the situation
- Some possible sources
  - State/Federal granting agencies – if you’re doing something innovative
  - Campus – for starting up new initiatives
  - Incorporate into operating budget – moving to a sustainable model
• You’ve heard a great deal about this already!
• Takes up a significant proportion of the project budget
• Designing a good workflow
  ▫ Connect milestones to your project timeline
  ▫ Set production benchmarks, but alter them when necessary
  ▫ Be sure to include preparation/preservation of physical materials
  ▫ Ensure effective communication with prior and later stages
Objects must meet defined uses

- Master and derivative image specifications should support services defined in the project plan
  - Masters should also be created to support future uses, at least as far as we can imagine
  - Don’t just copy others’ specifications without understanding if they will work for you
- Bottom line: there is no truly neutral view of a resource
Automation

- We have too many important things to do to spend our effort performing repetitive tasks
  - Manually distributing files to multiple places
  - Creating derivatives one by one
  - Emailing lists of images to people who need to know
- Small, nimble tools should be used for these tasks
  - Programming support when you can get it
  - Maximize use of existing tools (e.g., Office)
• If no one can find your resources, the rest isn’t of much use
• The VR community is among the most active specialist community in this area
  ▫ Lots of opinions and experience out there
  ▫ Sharing of expertise is common, but sharing of data and practices is not!
• Pick the right standards
• Design an effective workflow
• Need good tools
Selecting appropriate standards

- Elings/Waibel: “Metadata for All” article in *First Monday*, 2007
  - Argues standards should reflect materials, not holding institution
  - Slides and books, for example, should be treated differently

- Many decisions must be made
  - Descriptive metadata structure standards
  - Content standards
  - Controlled vocabularies
  - Technical, administrative, structural metadata
Workflow issues

- Optimize use of existing metadata
- Put the right information in front of the right person at the right time
- Automate when appropriate
- Design to be re-usable
- Many of the same issues apply from the digitization phase
  - Connect milestones to your project timeline
  - Set production benchmarks, but alter them when necessary
  - Ensure effective communication with prior and later stages
Metadata quality control

• Patterns (and outliers) emerge from data in the aggregate
• Reporting capabilities
  ▫ Sortable, deduplicated lists of values from a given field or set of fields
  ▫ How many of this field per record
  ▫ How many distinct values used in this field
  ▫ Data overlap between fields
Tools

- Must be re-usable
- Modularize when possible
- Types of tools that might help
  - Auto-complete
  - Spell-check
  - Integrated controlled vocabularies
  - Data integrity checking
  - Metadata transformation stylesheets
• Will be discussed tomorrow
• Not just sticking files somewhere
• Likely can’t do it effectively on your own
• Some initiatives to be aware of
  ▫ ICPSR Digital Preservation Tutorial
    http://www.icpsr.umich.edu/dpm/dpm-eng/contents.html
  ▫ Open Archival Information System (OAIS)
    http://public.ccsds.org/publications/archive/650x0b1.pdf
  ▫ RLG/OCLC Trusted Digital Repositories
• Will discuss in great detail later today
• Delivery services drive all other project requirements
• Discovery
  ▫ Allows users to find resources that meet their needs
• Use
  ▫ Allows users to do interesting things with the resources once they find them
• Promotion
  ▫ “Build it and they will come” hasn’t been effective in this area
Building and maintaining systems

- Plan for ongoing improvement
  - User expectations change
  - Technology changes
  - New usage scenarios emerge
- You *will* migrate to another system eventually
  - Plan for persistent URLs
  - Ensure conformance to standards
  - Limit the degree to which you design for a specific system
Interoperability and sharing

• Local users are often the first priority
• But we need to start thinking more widely
  ▫ We’re duplicating too much effort at our individual institutions
  ▫ Our data could be used to support services we haven’t even imagined
  ▫ Increasingly, working on our own is not sustainable
• Consider this when choosing and implementing delivery systems
• Digital projects are never done
• Ongoing development is more than simply adding more content over time
• In the current environment, need to repeatedly justify our investments
  ▫ Progress reports
  ▫ User-focused assessments
  ▫ Comparative analyses
• Don’t forget usage statistics
• Must constantly monitor new developments in the field
Bottom line

- Learn from experience (yours and others’)
- Project management is an active process
- Most difficult parts
  - Managing dependencies
  - Ensuring effective communication
  - Balancing the ideal and the practical
- Think big!