ETD Technology for Implementers

Presented March 22, 2001 at the 4th International Symposium on Electronic Theses and Dissertations

Anthony Atkins
anthony.atkins@vt.edu

Digital Library and Archives
VirginiaTech

Overview

- Stages in Implementing an ETD collection
  - Identify System Goals
  - Identify Communities of users and their needs
  - Translate Needs into Workflow
  - Develop Prototype
  - Test Prototype and Develop Production System
  - Refine Production System to meet changing awareness of needs.

Identifying System Goals

- What should the technology do?
  - Based in part on needs of users
  - Based in part on policy
  - Based also on ideals (what would we like the technology to do?)

Anthony Atkins
anthony.atkins@vt.edu

Digital Library and Archives
VirginiaTech

Identifying Communities of Users

- Students
- Faculty
- Graduate School
- Library
- System Developers
- System Maintainers
- Researchers

Anthony Atkins
anthony.atkins@vt.edu

Digital Library and Archives
VirginiaTech

Student Technology Needs

- All components must be easy to use without training
- Wide variety of user locations and client software
- Creation of content must continue to use tools familiar to students
- There should be few if any additional software costs
- Technology should respect student copyright concerns

Anthony Atkins
anthony.atkins@vt.edu

Digital Library and Archives
VirginiaTech

Technologies for Meeting Student Needs

- PDF
  - Portable Document Format. A technology for preserving the content and look of a document.
  - Creation of PDF files is based on the normal process of printing.
- Document Templates
  - Students can more easily comply with formatting standards.
- Web-based submission
  - Requires an internet connection and a web browser.
  - Submission process can be completed from on-campus user labs, home computers, even computers in geographically distant areas

Faculty Technology Needs

- All components must be easy to use with minimal training
- Variety of user locations and client software, but more controllable than students
- Process of reviewing content should use tools familiar to faculty
- Technology should respect faculty copyright concerns

Technologies for Meeting Faculty Needs

- Web-based review process
- Central Password Authorities for Student and Faculty Functions
- Secured Access to Full Text, as Requested

Anthony Atkins
anthony.atkins@vt.edu

Digital Library and Archives
VirginiaTech

Graduate School Technology Needs

- Components must be easy to use with specific training
- Components must be able to handle high enough volume
- Small number of users; proprietary technologies are acceptable for this community, although familiar technologies are preferable.
- Must be able to give feedback to authors

Technologies for Meeting Graduate School Needs

- Integrated Feedback (e-mail notices, etc.)

Anthony Atkins
anthony.atkins@vt.edu

Digital Library and Archives
VirginiaTech

Library Technology Needs

- Must be able to preserve electronic content
- Must be able to access electronic content as technologies change
- Small number of users; proprietary technologies are acceptable for this community, although familiar technologies are preferable.
- Should be able to integrate ETDs into existing catalog in some way

Technologies for Meeting Library Needs

- PDF
- XML
- Redundant Technologies
  - Mirroring of content on separate equipment
  - RAID (striping and/or mirroring of technology across drives)
- Backup technologies
  - Tape Drives
  - CD Burners

System Developer Technology Needs

- Technologies must be familiar to developers
- Technologies should allow incorporation of existing technologies (password authority, etc.)

Technologies for Meeting Developer Needs

- Commercial Technologies
  - Commercial technologies shift development time towards configuring rather than developing from scratch. Tend to provide greater reliability with a shallower learning curve.
  - Examples:
    - Integrated Library Systems
    - Oracle

Technologies for Meeting Developer Needs (continued)

- "Open" Technologies
  - Require much development time but offer great customizability and closer fit to user needs.
  - Examples:
    - Perl
    - PHP
    - Apache
    - MySQL
  - Combination Systems
    - Example: Using PHP to interface with Oracle.

System Maintainer Technology Needs

- Technology must be familiar enough that the function of scripts is understood by maintainer. This allows maintainer to more easily:
  - Fix problems that arise
  - Respond to changes in user needs (example: newer browsers)

Technologies for Meeting Maintainer Needs

- Commercial Technologies
  - Tend to be easier to install and upgrade for power users rather than system administrators.
  - Example:
    - Internet Information Server
- "Open" Technologies
  - Tend to be easier to install and upgrade for system administrators.
  - Examples:
    - MySQL
    - Apache

Researcher Technology Needs

- Components must be easy to use with no training
- Greatest variety of locations and user client software
- Technology should be integrated into systems researchers are accustomed to using.

Anthony Atkins
anthony.atkins@vt.edu

Digital Library and Archives
VirginiaTech

Technologies for Meeting Researcher Needs

- Web-based searches
- Integrated Library Systems that support links

Anthony Atkins
anthony.atkins@vt.edu

Digital Library and Archives
VirginiaTech

Bringing It Together: Translating Needs into Workflow

- Workflow:
  - The sequence of user and system actions that affect an ETD throughout its life cycle.
  - Based on the needs of all communities of users.
  - Provides a road map for the development of a prototype

Developing a Prototype: Hardware

- Low Volume Prototype Server:
  - Existing hardware
  - Use available hard drives

Anthony Atkins
anthony.atkins@vt.edu

Digital Library and Archives
VirginiaTech

Developing a Prototype: Software

- Take your detailed workflow, adapt or implement software for your needs.
- Research existing solutions
- Look at existing technologies that can be integrated into the system (ILS, etc.)
- For custom solutions, consider the expertise required to support existing systems, the same technologies may make sense.

Production System: Hardware

- Production Server:
  Dedicated server equipment with a lot of:
  Processing power (dual processors can be a good buy)
  Memory (500Mb - 1 Gb)
  Storage Space (dedicated 9+ Gb drive for theses)
  Bandwidth (100BaseT or other high speed connection if available)

Testing the Prototype and Developing a Production System

- Test using a variety of equipment, and from a variety of locations
- Try out as many hypotheticals as you can think of (what if the user does this and then doesn't do this?)
- Demonstrate the system for intended users, get their feedback
- Recruit intended users of the system, watch them as they use the system

Case Study: The Virginia Tech ETD-db system

- Hardware:
  - Dual-processor Sun Enterprise 250 with 384Mb of RAM
  - 18Gb drive for theses
- Software
  - Solaris
  - Apache
  - MySQL
  - Perl
- ETD-db suite
  - Web-based workflow management tools
  - Freely available at: http://scholar.lib.vt.edu/ETD-db/

Demonstration of ETD-db

- Demonstration Server Available
  http://lumiere.lib.vt.edu/ETD-db/
- Slide Show of Production System

Anthony Atkins
anthony.atkins@vt.edu

Digital Library and Archives
VirginiaTech

Case Study: North Texas

- Presented by Julian Long