

Digital Libraries and Virtual Universities

**Invited presentation for
Information Research for Designing and Planning
Virtual Universities**

**Universidad Nacional Autnoma de Mexico
August 11-15, 1997**

Edward A. Fox

Dept. of Computer Science

Virginia Tech, Blacksburg, VA 24061-0106 USA

<http://fox.cs.vt.edu> fox@vt.edu

Acknowledgements

- ◆ Interactive Learning with a Digital Library in Computer Science:
 - **N. Dwight Barnette**
 - **H. Rex Hartson**
 - **JAN Lee**
 - **Clifford Shaffer**
- ◆ Improving Graduate Education with a National Digital Library of Theses and Dissertations
 - **John Eaton**
 - **Gail McMillan**

Outline

- ◆ Part 1: Digital Library for Computer Science (briefly)
 - <http://ei.cs.vt.edu>
- ◆ Part 2: Digital Library of Theses and Dissertations
 - <http://www.ndltd.org>
 - <http://www.theses.org>
 - <http://www.dissertations.org>

Part 1:

Interactive Learning with a Digital Library in Computer Science

History

- ✦ 1991-93 A User-Centered Database from the Computer Science Literature - Envision Project
- ✦ 1993-97 Interactive Learning with a Digital Library in Computer Science - “EI”
- ✦ 1995- Networked Computer Science Technical Report Library - <http://www.ncstrl.org>
- ✦ June 1997 Workshop to disseminate EI results
- ✦ Fall 1997 Digital Libraries by ACM, IEEE CS, Elsevier and others
- ✦ DLI 2 (DL for undergrad SME&T education?)

Statistics

- ◆ Department has > 1000 undergraduate majors, 100 graduate students, and < 25 faculty
- ◆ 45 courses with WWW content
- ◆ 10 or more with significant WWW content
- ◆ 5 “paperless”
- ◆ 5 million accesses since 1995
- ◆ Majority of accesses from off campus
- ◆ Continual increase in access leading to upgrading and mirroring plus vBNS connection

WWW Courses

- ✦ CS1604 Introduction to the Internet
- ✦ (Honors 3004 Digital Libraries)
- ✦ CS3604 Computer Professionalism
- ✦ CS4624 Multimedia, Hypertext and Information Access
- ✦ CS5604 Information Storage and Retrieval
- ✦ (CS6604 Digital Libraries)

CRIM

- ✦ **crim-l@listserv.vt.edu**
- ✦ Curriculum Resources in Interactive Multimedia
- ✦ Focused effort out of CS4624 MHIA, in parallel with curriculum work sponsored by ACM SIGIR
- ✦ May need a Journal of Courseware to motivate and reward contributors
- ✦ May need an official repository or digital library to collect and organize and make available contributed courseware

Part 2:

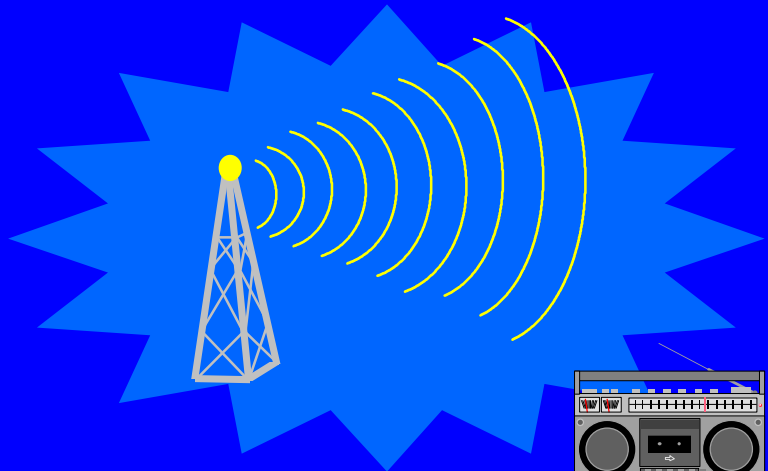
Improving Graduate Education
with a
National Digital Library of
Theses and Dissertations

ETDs Got Your Interest?

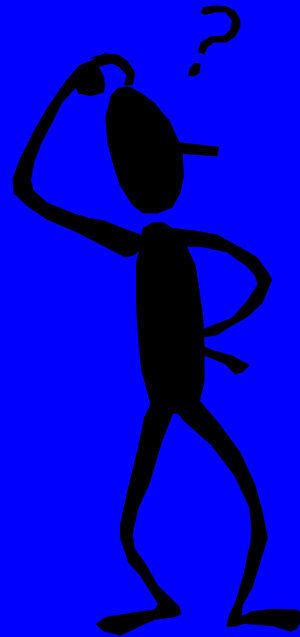


ETD Web Site

<http://www.ndltd.org/>



Media



Graduate Students

History

- ✦ 1987 mtg in Ann Arbor: UMI, VT, ...
- ✦ 1992 mtg in Washington: CNI, CGS, UMI, VT and 10 universities with 3 reps each
- ✦ 1993 mtg in Atlanta to start Monticello Electronic Library (MEL): SURA, SOLINET
- ✦ 1994 mtg in Blacksburg re ETD project: std of PDF + SGML + multimedia objects
- ✦ 1996 funding by SURA and US Dept. of Education (FIPSE) for regional, national projects (NDLTD)

What are we doing?

- ✦ Aiding universities to revitalize grad educ., publishing and IPR efforts: to help improve the availability and content of theses and dissertations
- ✦ Educating ALL future scholars so they can publish electronically and effectively use digital libraries (i.e., are Information Literate and can be more expressive)

What are the benefits?

- ◆ Save students money
- ◆ Save handling, shelf space in libraries
- ◆ Build the National Digital Library of Theses and Dissertations: with faster, broader, and less expensive access
 - UMI
 - OCLC
 - **Library catalog, local server, and:
university consortia, WWW engines**

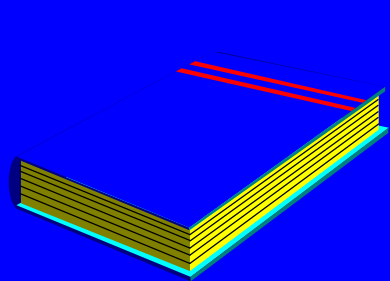
What are the long term goals?

- ◆ 400K US students / year getting grad degrees are exposed / involved
- ◆ 200K/yr rich hypermedia ETDs that may turn into electronic portfolios
- ◆ Dramatic increase in knowledge sharing: lit. reviews, bibliographies, ...
- ◆ Services providing lifelong access for students: browse, search, prior searches, citation links

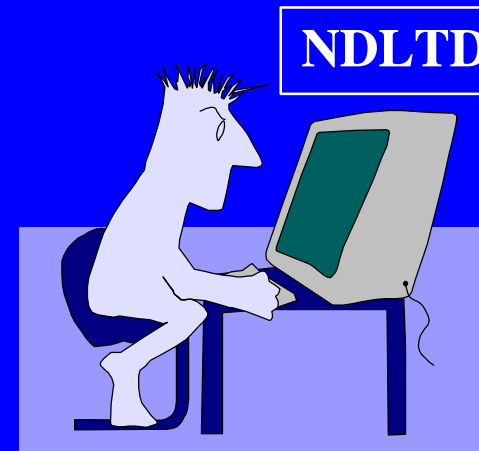
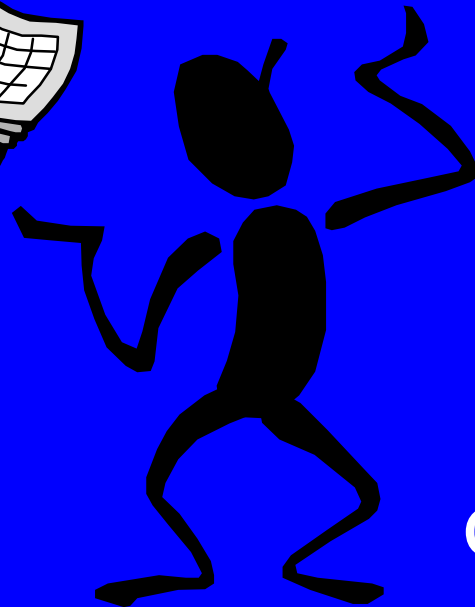
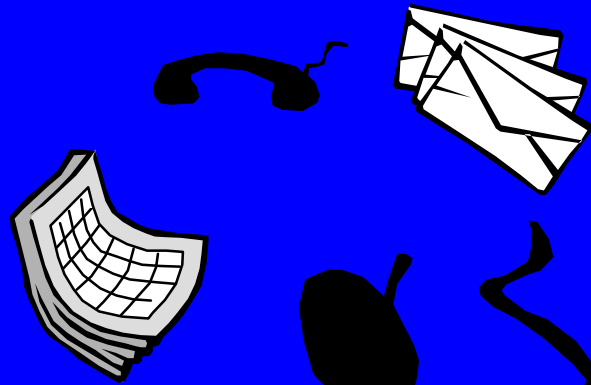
How are ETDs being done at Virginia Tech?

- ◆ Produced using standard word processing packages as PDF files
 - **LaTeX class, outline fonts**
 - **Word template, PDFwriter**
- ◆ Reviewed by the Graduate School
- ◆ Cataloged and archived by the library
- ◆ Downloaded by UMI from server

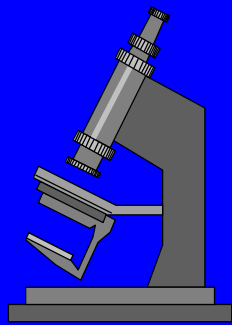
Student Prepares Thesis or Dissertation



Literature

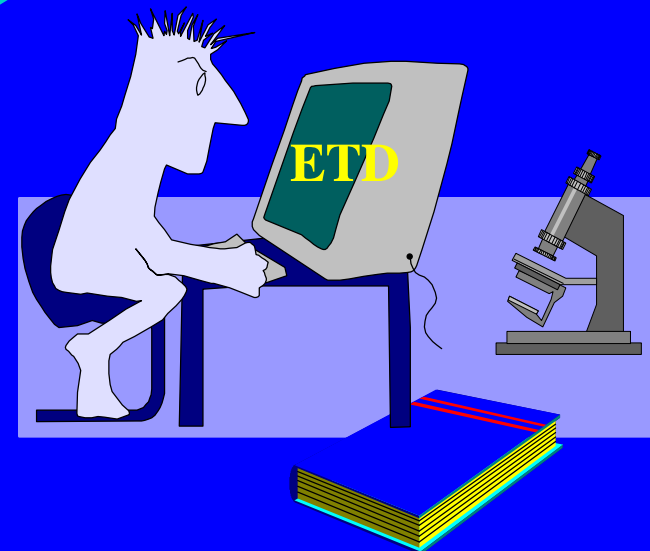
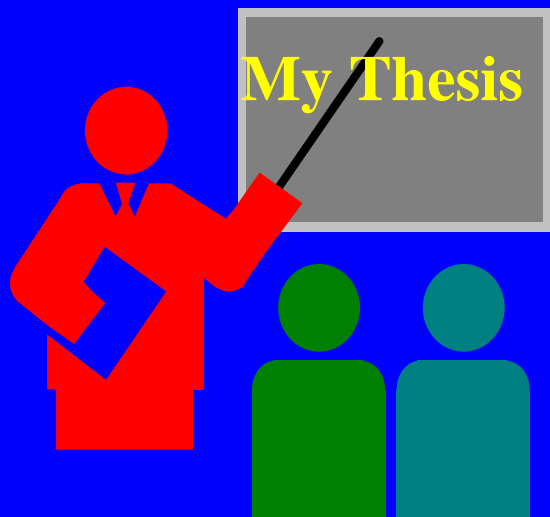


Computer Resources

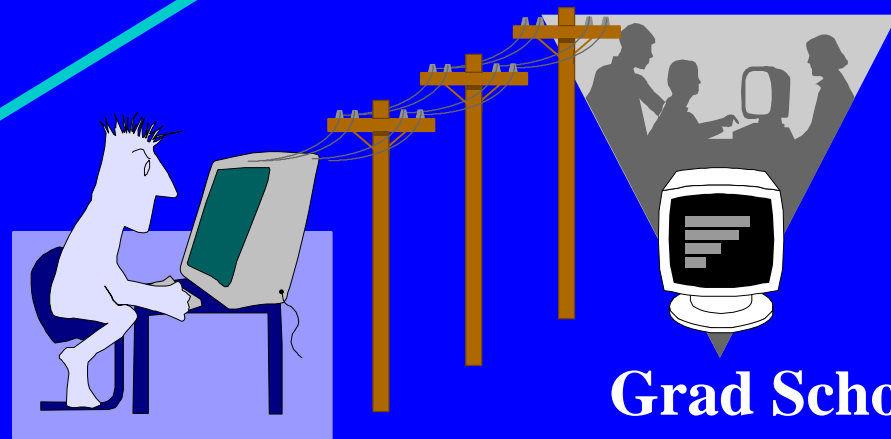
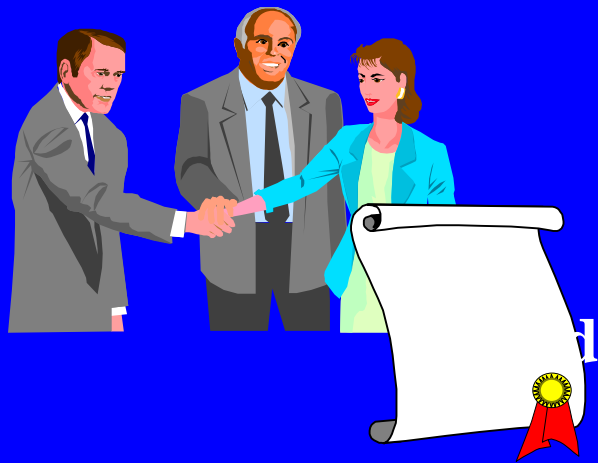


Research

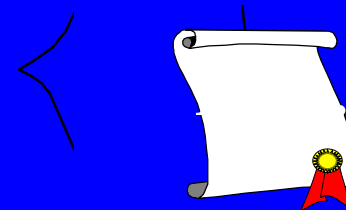
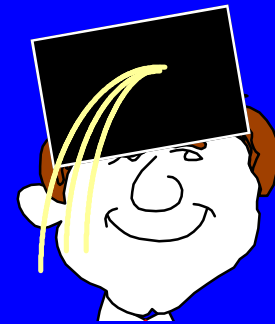
Student Defends and Finalizes ETD



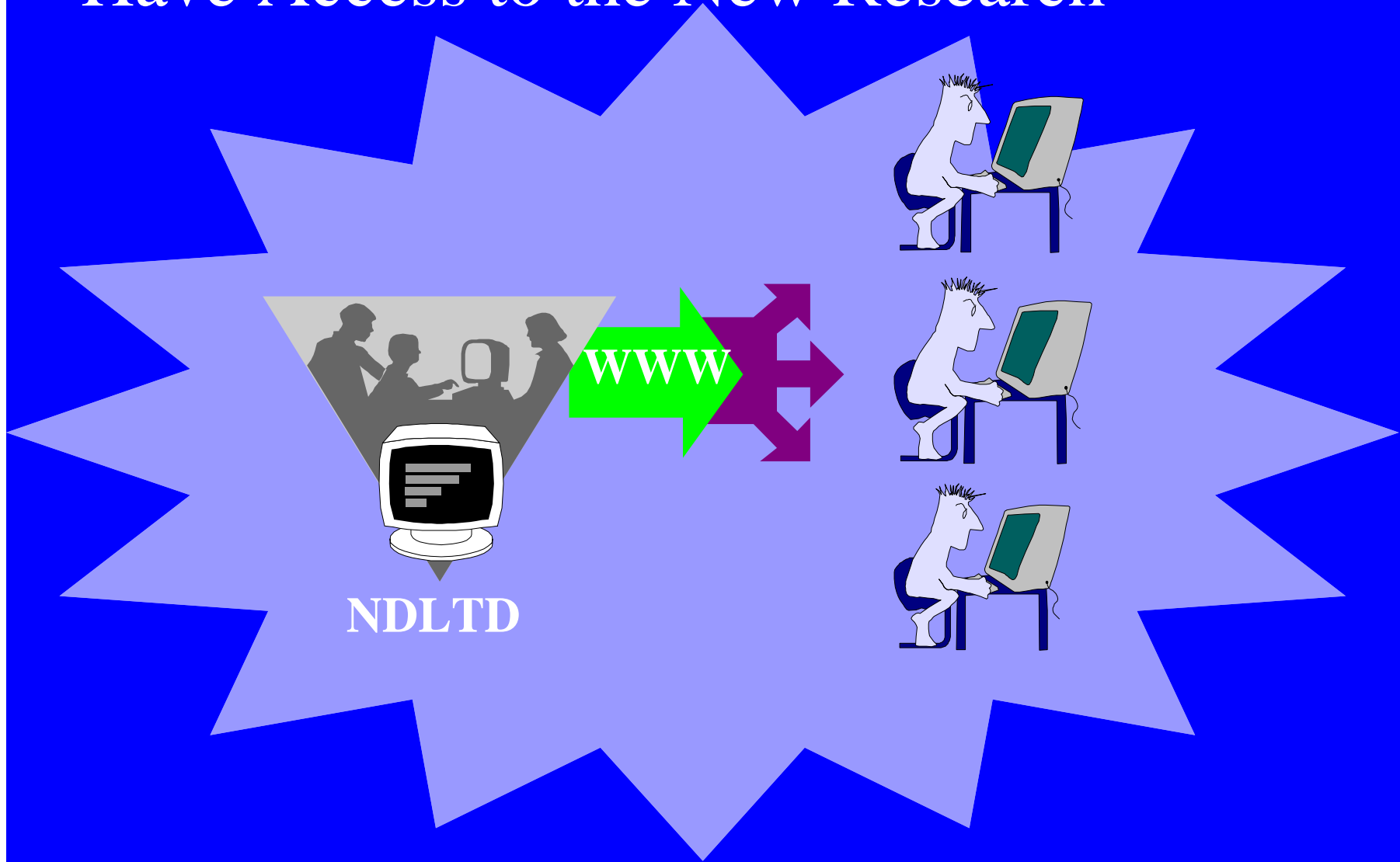
Student Gets Committee Signatures and Submits ETD



Graduate School Approves ETD Student is Graduated



Library Catalogs ETD and New Students Have Access to the New Research



Status of the Local Project

- ◆ Approved by university governance
Spring 1996; required starting 1/1/97
- ◆ Submission & access software in place
- ◆ Submission workshops for students
(and faculty) occur often
- ◆ Faculty training as part of Faculty
Development Initiative
- ◆ Over 300 ETDs in collection
- ◆ IBM h/w in place, DL s/w installed

Universities that have Expressed Interest in the Project

- ✦ Cal Tech
- ✦ College of W & M
- ✦ Duke Univ.
- ✦ George Mason U.
- ✦ Mississippi State
- ✦ Oklahoma State
- ✦ Penn State
- ✦ Portland State
- ✦ Tech. U. Munich
- ✦ Texas A&M
- ✦ Texas Womens U.
- ✦ Univ. AZ, CO, MI, Umea (Sweden)
- ✦ Washington State
- ✦ West Virginia Univ.

CIC Considering Joining

Universities (to be) Visited by Virginia Tech Staff

- ✦ Auburn
- ✦ CMU
- ✦ Columbia
- ✦ Georgia Tech
- ✦ MIT
- ✦ (Nat'l Lib. Canada)
- ✦ Nat'l Univ. Singapore
- ✦ New York U.
- ✦ Ohio State U.
- ✦ Rice Univ.
- ✦ Rutgers Univ.
- ✦ San Jose State U.
- ✦ Stanford Univ.
- ◆ U. Alabama
- ◆ U. Al. Birmingham
- ◆ U. CA Berkeley
- ◆ U. CA Santa Barbara
- ◆ U. Central Florida
- ◆ U. Delaware
- ◆ U. Ill. Urbana Ch.
- ◆ U. NC Charlotte
- ◆ U. North Florida
- ◆ U. Pennsylvania
- ◆ U. Utah
- ◆ U. Waterloo
- ◆ U. Wisc. Madison

Universities Officially Part of ND LTD

- ◆ **Clemson University**
- ◆ **Concordia University (Illinois)**
- ◆ **Darmstadt University of Technology (Germany)**
- ◆ **Florida Institute of Technology**
- ◆ **Michigan Tech**
- ◆ **Naval Postgraduate School**
- ◆ **North Carolina State University**
- ◆ **Rhodes University (South Africa)**
- ◆ **Rochester Institute of Technology**
- ◆ **University of Florida**
- ◆ **University of Georgia**
- ◆ **University of South Florida**
- ◆ **University of Tennessee: Knoxville, Memphis**
- ◆ **University of Virginia (Bachelor's theses)**
- ◆ **Vanderbilt University**
- ◆ **Virginia Tech**

Who are the sponsors and cooperators?

- ✦ ACM
- ✦ Adobe Systems
- ✦ Arbortext
- ✦ CGS, CSGS, CNI
- ✦ IBM
- ✦ Microsoft
- ✦ OCLC
- ✦ SURA, SOLINET
- ✦ UMI
- ✦ US Dept. of Education (FIPSE)

How does this relate to UMI?

- ✦ 1987 UMI workshop to explore ETDs
- ✦ Support letter for US Dept. of Ed. proposal
- ✦ Steering and technical committee membership
- ✦ ProQuest Direct pilot of scanning works started 1/1/97
- ✦ Collaborating on:
 - **accepting author submissions**
 - **standards (e.g., representation)**

Interoperability Tests Planned

- ✦ IBM DL: donated equipment, technical support, powerful IPR (see TOIS, D-Lib)
- ✦ Z39.50 / DL protocol / OCLC SiteSearch
 - university libraries w. catalogs of freely shared MARC records pointing to archival copies
 - via URNs: handles & PURLs
- ✦ Dienst / NCSTRL - www.ncstrl.org: CS depts., DARPA, NSF, CNRI, Cornell - UVA is working on extensions for ETDs

Access Approaches

- ◆ Goal: Maximize access and services, e.g., by encouraging:
 - ◆ UMI and OCLC centralized services
 - ◆ Distributed service: Dienst, Z39.50
 - ◆ Regional services (e.g., CIC, MEL)
 - ◆ Local servers with browse, search
 - From local catalogs to local archives
 - ◆ WWW robot indexing and search services

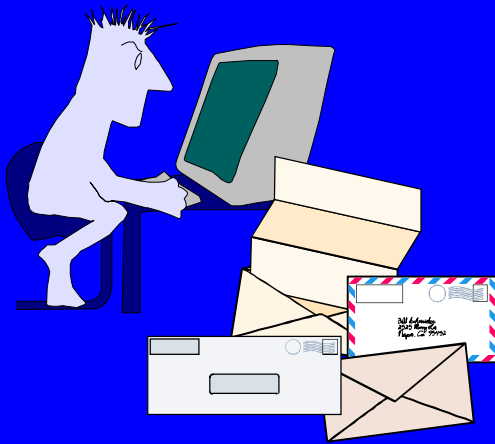
Why might your university want to be involved?

- ✦ To improve graduate education / better prepare your students
- ✦ To unlock university information
- ✦ To save money for students and for the university / improve workflow
- ✦ To build an important digital library supported by SURA, FIPSE

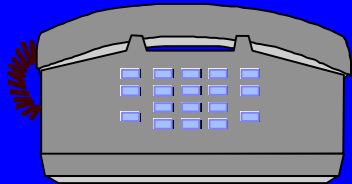
How can your university get involved?

- ✦ Select planning/implementation team
 - **Graduate School**
 - **Library**
 - **Computing / Information Technology**
 - **Institutional Research / Educ. Tech.**
- ✦ Send us letter, give us contact names
- ✦ Adapt Virginia Tech solution
 - **Build interest and consensus**
 - **Start trial / allow optional submission**

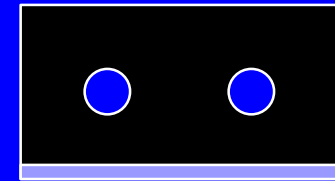
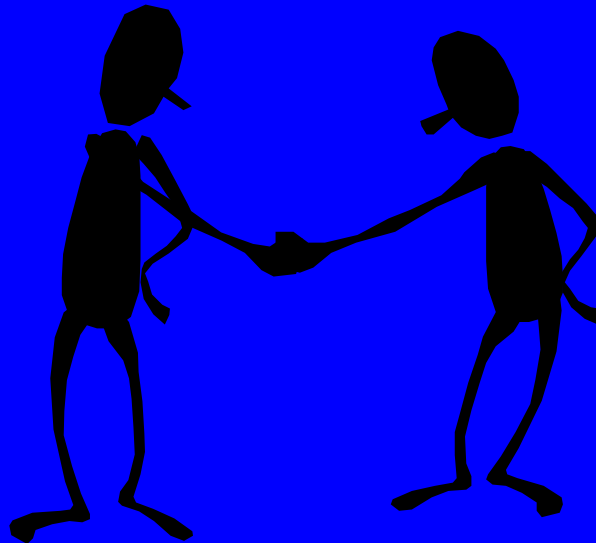
Contact Our Project Team



E-mail
etd@ndltd.org



Phone Call

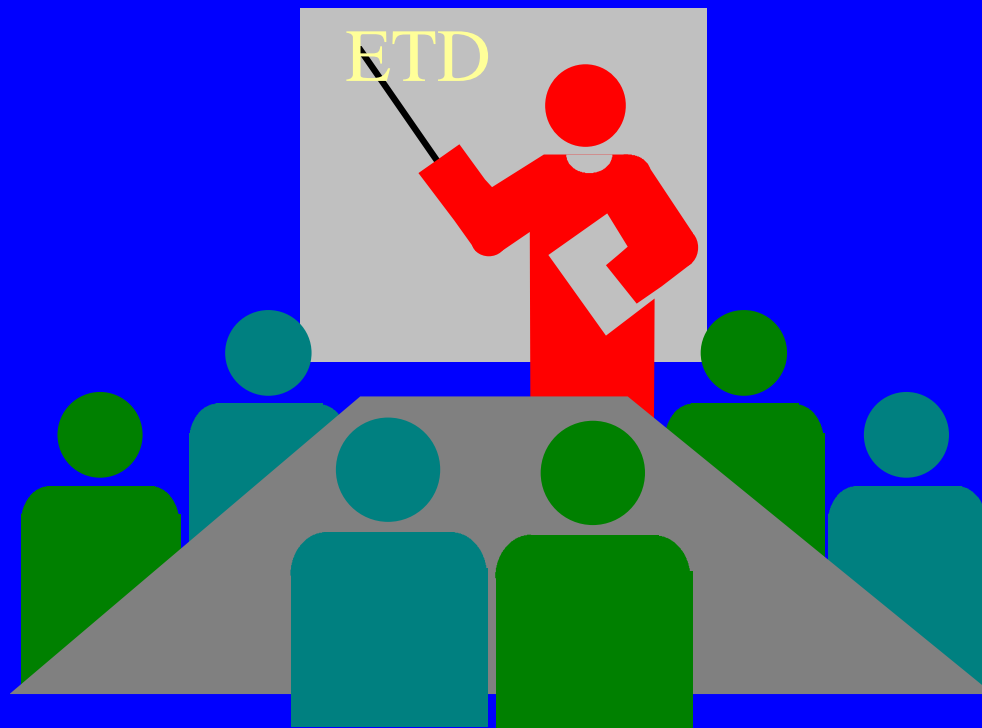


Video Tape



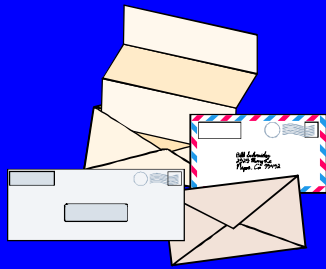
Visit

Convene Local Planning Group

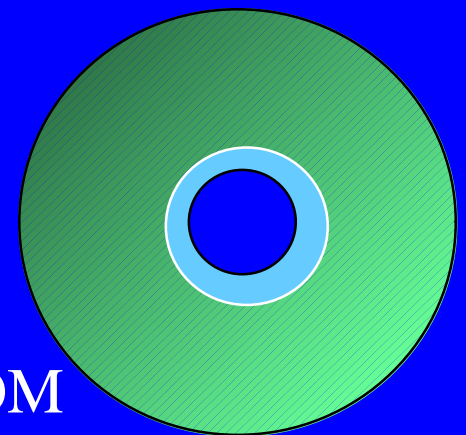
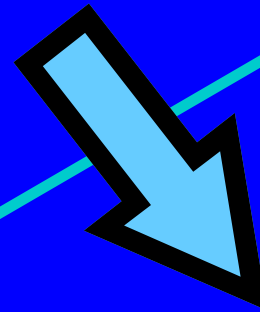


Join NDLTD: Get a CD-ROM

WE JOIN

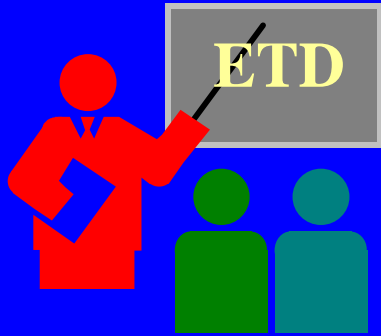


Signed Letter

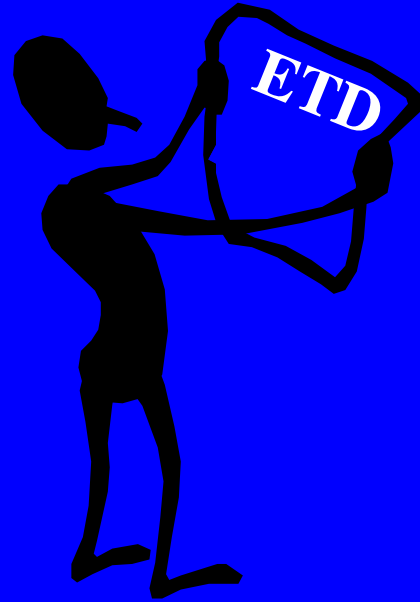


ETD CD-ROM

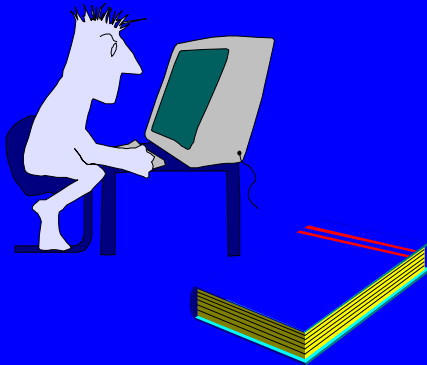
Build Your ETD Site



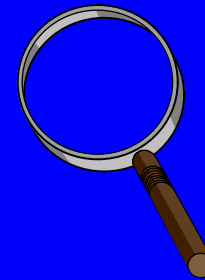
Workshop/Training



Policies



Digital Library



Inspection/Approval

Level 0 Involvement

RISK FREE - allow students

- ◆ Adobe Acrobat in bookstore
- ◆ Submission allowed
- ◆ Archive/access through Virginia Tech, OCLC, and/or UMI
- ◆ (Local) WWW site, publicity
- ◆ (Local) Assistance provided as requested: email, phone, listserv(s)

Level 1 Involvement = Level 0 +

LOW COST - help & encourage students

- ◆ Install our software, change practices in graduate school and library
- ◆ Train students
- ◆ Build grass roots support
 - **Advisory committee: representative? expert?**
 - **Champions to spread by word of mouth**
 - **Approval: Senates, Commissions, Deans, Students**
 - **Publicity to reach community**

Level 2 Involvement = Level 1 + EVENTUAL FULL INVOLVEMENT

- ◆ Require electronic submission
- ◆ Have firm arrangement with local library, OCLC, VT and/or UMI re archival services
- ◆ Share MARC records, with URNs pointing to archive copy
- ◆ (Stock laboratories)
- ◆ (Run servers: search, URN)
- ◆ (Launch evaluation program)

Support Services Developed

- ◆ WWW site with > 300 Mbytes, CD
- ◆ Automated submission system
- ◆ Student guidelines, style sheets, training materials, FAQs, press info
- ◆ SGML DTD for ETDs
- ◆ SGML to HTML (web generator)
- ◆ LaTeX, Word templates, converters
- ◆ Multimedia educ. materials, Videotape

Support Offered

- ◆ Software, documentation, tech support
- ◆ Email, listservs (etd-1@listserv.vt.edu, -eval, -grad, -library, -technical)
- ◆ Donations: Adobe, Microsoft
- ◆ Evaluation: instruments, analysis
- ◆ (Temporary storage / archiving)
- ◆ (Aid - in setting up a national service and archive)

Relationship with publishers

- ✦ Concern of faculty and students that still wish to publish books or journal articles, voiced: campus, Chronicle, NPR, Times
- ✦ Solution: Approval Form gives students, faculty choices on access, when to change access condition; use IPR controls in DL
- ✦ Solution: by case, work with publishers and publisher associations to increase access
 - AAP, AAUP
 - AAAS, ACM, ACS, Elsevier, ...

Some statements from publishers

- ✦ Elsevier: the two genre do not conflict
- ✦ AAAS: Science wants first publication
- ✦ Textbook publishers: different market, manuscript significantly reworked
- ✦ General: restricting access to local campus will not cause any problems

Scenario 1: Patent/Proprietary

- ✦ Student work cannot be released since plans to file a patent
- ✦ University holds till can be filed - work is “locked up”

Scenario 2: Unrelated

- ✦ Student has no history of, or plans for, publishing related to ETD

Scenario 3: Hold parts

- ✦ Some parts (chapters, figures) relate to published works or works to be submitted
- ✦ Those are secured for campus-only access

Scenario 4: Delay till after publication

- ✦ The work relates to a publication, so the ETD is made accessible on campus only.
- ✦ If the relationship is not strong, then the ETD is released only after the journal publication has appeared, plus say 3 months

Scenario 5: Point to electronic journal articles

- ✦ A part of an ETD relates closely to a published work.
- ✦ The “normal” version of the ETD is released on campus, with credit to the copyright owner on the first page of each part that relates.
- ✦ The “public” version of the ETD has each part, that relates to a published or to-be published work, replaced by a ptr/link .

Future Work

- ✦ Working with publishers to increase level of access as much as possible
- ✦ Interoperability tests among universities
- ✦ Study with collection that emerges, to improve information retrieval, browsing, interface, and other types of user support
- ✦ Evaluation, improving learning experience, spread to worldwide initiative, sustainable support and coordination

CONCLUSIONS

- ◆ There is a rich collection of CS courseware at <http://ei.cs.vt.edu>
- ◆ All universities should join the NDLTD (see <http://www.ndltd.org>)
- ◆ Universities can work together to improve education worldwide
- ◆ Having a DL can change the way students learn, and make it easier to have virtual universities