Future Funding and the NDLTD Vision

St. Petersburg
March 18, 2000

Third International Symposium on Electronic Theses and Dissertations

Edward A. Fox
fox@vt.edu


Virginia Tech, Blacksburg, VA, USA
The Networked Digital Library of Theses and Dissertations

NDLTD

Training Authors
Expanding Access
Preserving Knowledge
Improving Graduate Education
Enhancing Scholarly Communication
Empowering Students & Universities

Leader of the
Worldwide ETD
(Electronic Thesis and Dissertation)
Initiative
Acknowledgements (Selected, VT)

**Sponsors:** ACM, Adobe, IBM, Microsoft, NSF, OCLC, US Dept. of Education, ...

**Co-PIs:** John Eaton, Gail McMillan

**Faculty/Staff:** Anthony Atkins, Thomas Dunbar, Gwen Ewing, Peter Haggerty, Gary Hooper, Len Peters, James Powell

**Students:** Emilio Arce, Fernando Das Neves, Brian DeVane, Robert France, Marcos Goncalves, Scott Guyer, Robert Hall, Neill Kipp, Paul Mather, Tim McGonigle, Todd Miller, Constantinos Phanouriou, William Schweiker, Ohm Sornil, Hussein Suleman, Patrick Van Metre, Laura Weiss
Acknowledgements (Selected)

Conference Sponsors: Adobe, CGS, Dell, Macromedia, Microsoft, USF, Xerox


Conference Chair: Joe Moxley

USF Team: Terry Beavers, Bruce Cochrane, Rosann Collins, Ilene Frank, Monica Metz-Wiseman, Anita Callahan, …
Outline

 открывание (technology base on campuses)

 Growth

 NUDL, OAI – international cooperative and new era in scholarly communication

 NDLTD, Inc. – sustainable services

 GSDI – preparing and empowering
A Digital Library Case Study

- **Domain:** graduate education, research
- **Genre:** ETDs = electronic theses & dissertations
- **Submission:** http://etd.vt.edu
- **Collection:** http://www.theses.org

**Project:**
Networked Digital Library of Theses & Dissertations (NDLTD) http://www.ndltd.org
What are we doing?

- Aiding universities to enhance graduate education, publishing and IPR efforts
- Helping 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)
Digital Libraries

- Improving Education (1995)
- Internet (1984)
- NSF DLI (1994)
- SGML (1985)
- Multimedia (1986)
- PDF (1992)
- Library Cancellations (1988)
- WWW (1994)

Info. Literacy (1995)
Key Ideas:

- Networked infrastructure
- University collaboration
- Workflow, automation
- Education is the rationale
  - 8th graders vs. grads
  - Authors must submit
- Standards
  - PDF, SGML, MM
  - MARC, DC, URNs
- Federated search
User Search Support
(multilingual, XML)

NDLTD World Federated Search

User Interface

Virginia Tech ...
(univ)

OhioLink
(lib / univ group)

Australia
(regional)

UMI ...
(corporate)

Portuguese NL ...
(national lib)

Note: All groups shown are connected with NDLTD.
www.theses.org

- XML description of each site
  - type of search engine / service
  - language
  - coverage (for resource discovery)
- Adding Z39.50 gateway capability and integrating with MARIAN, along with Harvest and Open Archives protocols (according to Santa Fe Convention) – see www.openarchives.org
# Access Possibilities

<table>
<thead>
<tr>
<th>Web search engines</th>
<th><a href="http://www.theses.org">www.theses.org</a></th>
<th><a href="http://www.openarchives.org">www.openarchives.org</a></th>
<th>library catalog clients</th>
<th>3rd Party Services (e.g., Bell &amp; Howell)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Virginia Tech</td>
<td>MIT</td>
<td>National Library of Portugal</td>
<td>CBUC (Spain)</td>
<td>Ohio Link</td>
</tr>
</tbody>
</table>
PetaPlex

Digital Library Machine ("super" object store)

Parallel computer / storage utility for terabytes to petabytes

Knowledge Systems Incorporated is supplying VT-PetaPlex-1 with

- high speed backbone connection
- 2.5 terabytes through 100 nodes:
  - Net connection + 25GB disk + 233 MHz Pentium + Linux
How does this relate to UMI?

- 1987 UMI workshop to explore ETDs
- Support letter for US Dept. of Ed. proposal
- Steering committee membership
- ProQuest Direct pilot of scanning works started 1/1/97, free 2 yr access to front part

Collaborating on:
- accepting electronic author submissions
- standards (e.g., representation)
ETD Initiative (and UMI)

Education

Students
Learn about DL, EPub

Universities

UMI

TDs become more expressive

Global TDs become more accessible, archived

N. Amer. (T)Ds are accessible, archived
Outline

- DLs (technology base on campuses)
- **Growth**
  - NUDL, OAI – international cooperative and new era in scholarly communication
  - NDLTD, Inc. – sustainable services
  - GSDI – preparing and empowering
US University Members (41)

- Air University (Alabama)
- Baylor University
- Brigham Young University
- Caltech
- Clemson University
- College of William & Mary
- Concordia University (Illinois)
- East Carolina University
- East Tenn. State U. – require fall 2000
- Florida Institute of Tech.
- Florida International University
- George Washington University
- Marshall University (W. Va.)
- Miami U. of Ohio
- MIT
- Michigan Tech
- Naval Postgraduate School (CA)
- North Carolina State U.
- Penn. State University
- Rochester Institute of Tech.
- U. of Colorado Health Science Center
- U. of Florida
- U. of Georgia
- University of Hawaii, Manoa
- U. of Iowa
- U. of Kentucky
- U. of Maine
- U. of North Texas – required since 8/99
- U. of Oklahoma
- U. of South Florida
- U. of Tennessee, Knoxville
- U. of Tennessee, Memphis
- U. of Texas at Austin
- U. of Virginia
- U. Wisconsin - Madison
- Vanderbilt U.
- Virginia Commonwealth U.
- Virginia Tech - required since 1/97
- West Virginia U. - required fall 1998
- Western Michigan U.
- Worcester Polytechnic Inst.
Institutional Members

- Coalition for Networked Information (CNI)
- Committee on Institutional Cooperation (CIC)
- Diplomica.com
- Dissertation.com
- Dissertationen Online (Germany)
- National Library of Portugal (for all universities)
- Organization of American States (SEDI/OAS)
- UNESCO (www.unesco.org/webworld/etd)
Australian Project Members

- U. New South Wales (lead institution)
- U. of Melbourne
- U. of Queensland
- U. of Sydney
- Australian National University
- Curtin U. of Technology
- Griffith U.
German Project Members

- Humboldt University (lead institution)
- 3 other universities
- 5 learned societies
  - Mathematics, Physics, Chemistry, Sociology, Education
- 1 computing center
- 2 major libraries
CBUC (www.cbuc.es, Spain)

- Consorci de Biblioteques Universitàries de Catalunya, as group, with 9 members:
  - Universitat de Barcelona
  - Universitat Autònoma de Barcelona
  - Universitat Politècnica de Catalunya
  - Universitat Pompeu Fabra
  - Universitat de Girona
  - Universitat de Lleida
  - Universitat Rovira i Virgili
  - Universitat Oberta de Catalunya
  - Biblioteca de Catalunya
Other International Members

- Chinese University of Hong Kong
- Chungnam National U., Dept of CS (S. Korea)
- City University, London (UK)
- Darmstadt U. of Tech. (Germany)
- Free University of Berlin (Germany - Vet. Med.)
- Gyeongsang National U. (Korea)
- India Institute of Technology, Bombay (India)
- Nanyang Technological U. (Singapore, part)
- National U. of Singapore (Singapore, part)
- Polytechnic University of Valencia (Spain)
- Rhodes U. (South Africa)
- St. Petersburg St. Tech.U (Russia)
- Univ. de las Américas Puebla (Mexico)
- Univ. of Alicante (Spain)
- Univ. of Pisa (Italy)
- U. Laval; U. of Guelph; U. Waterloo; Wilfrid Laurier U. (Canada), …
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
Economic Considerations

- Collect modestly ($20) for archiving
- PDF is a published de facto standard
- Free access to ETDs saves student researchers hundreds of dollars as study more related TDs

Consider big picture:
- $1B/yr invested on students who produce TDs
- Per student that is over $10,000
- ETDs have 10 to 1000 times the audience

If done right, ETD initiatives will save $:

Perspective: 1 Gbyte/year, DL seed on campus
Objectives

- Increase number of ETDs per year
- Increase number of NDLTD members
- Increase number of ETDs per campus
- Improve quality of ETDs

- Enabling standards  (Berlin, May 10-12)
- Enabling tools and services
- Enabling educational resources
Continuing Growth

- http://www.ndltd.org/talks
- Speaker list and circuit
- University -> City -> State -> Region
- State networks (library, networking)
- University consortium (CIC)
- National libraries (Portugal)
- Plug and Play NDLTD Server
- Awards for best ETDs
Outline

- DLs (technology base on campuses)
- Growth
- **NUDL, OAI** – international cooperative and new era in scholarly communication
- NDLTD, Inc. – sustainable services
- GSDI – preparing and empowering
1/15/99 NUDL proposal to NSF under DLI2 international program

- **VT**: Library, Grad School, Industrial & Systems Eng.
- **Partners**: UK (2), Singapore, Russia, Korea, Greece, Germany, plus Iberoamerican group (Spain, Portugal, Argentina, Brazil, Chile, Mexico)
- **Problems**: Multilingual search, multimedia submissions, requirements/usability, ...

Start with ETDs, then expand to other student works, portfolios, data sets, (CS) courseware, ...
National Coverage (red/white)
NUDL Partners

- Ricardo A. Baeza-Yates, Universidad de Chile, Chile
- José Luis Brinquete Borbinha, Biblioteca Nacional, Portugal
- José Hilario Canós Cerdá, Universidad Politécnica de Valencia, Spain
- Stavros Christodoulakis, Technical University of Crete, Greece
- Lautaro Guerra Genskowsky, Universidad Técnica Federico Santa Maria, Chile
- Juan José Goldschtein, Universidade de Belgrano, Argentina
- Peter Diepold, Humboldt University, Germany
- Francisco Javier Jaén Martinez, Spain
- Sung Hyon Myaeng, Chungnam National University, Korea
- Ana Maria Beltran Pavani, Prédio Cardeal Leme, Brazil
- Lim Ee Peng, Nanyang Technological University, Singapore
- Alexander I. Plemnek, St.-Petersburg State Technical University, Russia
- J. Alfredo Sánchez, Universidad de las Américas-Puebla, Mexico
For professional societies

- Like “writing across the curriculum”
- Besides writing: computing/communications, information literacy, personal digital library management, tool use, research methods, collaboration, archiving/preservation
- Data sets, communities of users of them
- Classification systems / browsing / searching
- NRC’s “On becoming a researcher”
Open Archives Initiative (OAI)
www.openarchives.org

- Santa Fe meeting, Oct. 21-22, 1999, protopproto
- Next mtg June 3, San Antonio, between HT’00 & DL’00
- LANL, CNI, DLF, Mellon, …
- Convention (see Feb. D-Lib Magazine article)
- Archives -> Open Archives
  - Support unique archive identifiers
  - Implement Open Archives Metadata Set (DC-based, using XML)
  - Implement Dienst harvesting interface
  - Register the archive
- Build tools, layer other services: linking, searching, …
Approaches to Open Archives

Build By Institution

Build By Discipline

Access by
Author
Category
Interdisciplinary
Year
Language
Query …
Outline

- DLs (technology base on campuses)
- Growth
- NUDL, OAI – international cooperative and new era in scholarly communication
- NDLTD, Inc. – sustainable services
- GSDI – preparing and empowering
Accessibility Activities / Plans

- Interface design (simple, 3D, VR)
- Usability studies
- Generic multi-lingual support
- Support for those with disabilities
- Hybrid collection (paper, MARC, abstracts, full-text, multimedia)
- Disciplinary classifications, tools
- Visualization of results, collection
Extending Services - 1 of 2

- Working with publishers to increase level of access as much as possible
- **Interoperability** tests among universities and other interested partners to provide integrated services
- Study with testbed that emerges, to improve information retrieval, browsing, interface, and other types of user support -> NSF, DFG, …
- Evaluation, **improving learning** experience, spread to worldwide initiative, sustainable support and coordination -> FIPSE, foundations, …
Extending Services - 2 of 2

Adding services currently prototyped
- annotation and SDI (routing) capabilities
- Dublic Core metadata, crosswalk to MARC
- support for XML, *ML, preservation
- harvesting, federated search

Adding other services planned
- building/using citation DB (CiteSeer, SFX)
- implementing plagiarism check (like “SCAM”)
Sustainability

- Continuing growth, with critical mass targeted within 2-3 years
- Developing NDLTD as a sustainable self-governing global institution
- Evolving into a federated member-funded organization
  - mirroring/replication/access
  - developing/sharing standards/practices
  - clearinghouse with tools and translation
Outline

- DLs (technology base on campuses)
- Growth
- NUDL, OAI – international cooperative and new era in scholarly communication
- NDLTD, Inc. – sustainable services
- GSDI – preparing and empowering
GSDI

- Graduate Student Development Initiative
- Recall Faculty Development Initiative
- Extending efforts of CGS, CSGS
- Cast according to the campus, regional and national approaches to graduate education
GSDI Themes

- Ongoing training/education/evaluation
- Preparing scholars
- Preparing innovators/leaders
- Improving communication/writing
- IT literacy -> knowledge & skills
  - Digital libraries
  - (Electronic) publishing
ETD as Interface to Research

- Supporting
  - Communication/Collaboration
  - Organization/Understanding/Education

- Artifacts
  - Document, Database, Hyperbase
  - Simulation, VR, Active object
  - Literature review, Citation DB
  - Knowledge structure
Moving ETDs into the Future

- Preservation
  - Importance based (core / transient)
  - Aiding discovery, reuse

- Integrating with scholarly communications enterprise

- Integrating with students’ ongoing research, teaching, evolution
Remember!

- DLs (technology base on campuses)
- ETDs (growing initiative – act now!)
- NUDL, OAI (help establish international cooperative and new era in scholarly communication)
- NDLTD - improved graduate education!
  - federation to join and support!
  - www.ndltd.org/join