

Status of the NDLTD
Networked Digital Library of Theses
and Dissertations



Gail McMillan

Digital Library and Archives

University Libraries, Virginia Tech

CNI Spring Task Force Meeting
April 10, 2001

The Networked Digital Library of Theses and Dissertations

www.NDLTD.org

COMMON GOALS

Expanding Access

Preserving Knowledge

Improving Graduate Education

Enhancing Scholarly Communication

Empowering Students & Universities

Leader of the Worldwide ETD

(Electronic Thesis and Dissertation) Initiative

Why ETDs?

- Help graduate students learn about e-publishing and digital libraries
- Apply that knowledge to research, build, submit ETDs
- Empower students to convey richer message through use of multimedia
- Improve graduate education through more effective sharing of information
- Timely availability of university's unlock intellectual property
- Libraries improve services and resources, save money and space

A Little ETD History

- When did it begin?
 - 1987 Michigan/UMI meeting
 - 1994 Blacksburg, VA meeting (who, formats)
 - 1996: SURA, FIPSE
 - Funding outside VT
 - Supports VT Team travel to US universities
- 1997
 - VT requires ETDs (no paper)
 - NDLTD: National/Networked Digital Library of Theses and Dissertations
- 1999-2001+:
National conferences

US University Members: 48

- ✦ Air University (Alabama)
- ✦ Baylor University
- ✦ Brigham Young University (part, whole)
- ✦ Caltech
- ✦ Clemson University
- ✦ College of William & Mary
- ✦ Concordia University (Illinois)
- ✦ East Carolina University
- ✦ **East Tenn. State U. – required fall 2000**
- ✦ Florida Institute of Technology
- ✦ Florida International University
- ✦ George Washington University
- ✦ Louisiana State University
- ✦ Marshall University (W. Va.)
- ✦ Miami University of Ohio
- ✦ Michigan Tech
- ✦ Mississippi State University
- ✦ MIT
- ✦ Naval Postgraduate School (CA)
- ✦ New Mexico Tech
- ✦ North Carolina State University
- 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 fall 1999**
- U. of Oklahoma
- U. of South Florida
- U. of Tennessee, Knoxville
- U. of Tennessee, Memphis
- **U. of Texas at Austin – required in 2001**
- U. of Virginia
- U. Wisconsin - Madison
- Vanderbilt U.
- Virginia Commonwealth U.
- **Virginia Tech - required Jan. 1, 1997**
- **West Virginia U. - required fall 1998**
- Western Michigan U.
- Worcester Polytechnic Inst.

NDLTD International Members (48)

- Australia
- Belgium
- Brazil
- Canada
- China
- Germany
- Hong Kong
- India
- Italy
- Korea
- Mexico
- Netherland
- Norway
- Russia
- Singapore
- **S. Africa**
- S. Korea
- Spain
- Taiwan
- UK

NDLTD Institutional Members (14)

- CNI: Coalition for Networked Information
- CIC: Committee on Institutional Cooperation
- Consorci de Biblioteques Universitaries de Catalunya
- Diplomica.com
- Dissertatione Online (German national project)
- Dissertation.com
- ETDweb, Division of Answers4.com
- Ibero-American Science & Technology Education Consortium (ISTEC)
- National Library of Portugal
- OhioLINK
- OCLC: Online Computer Library Center
- Organization of American States
- Solinet: Southeastern (U.S.) Library Network
- UNESCO

What are we doing?



- Helping improve the availability and content of theses and dissertations
- Educating future scholars so they can publish electronically and effectively use digital libraries (i.e., are information literate and can be more expressive)
- Aiding universities to enhance graduate education, publishing and IPR efforts

NDLTD Conferences

- 1999: Virginia Tech/Blacksburg
 - 60 attendees, 33 institutions
- 2000: Uni. of South Florida/St. Petersburg
 - 200 attendees, 14% international
- 2001: Cal Tech/Pasadena
 - 165 attendees, 24% international
 - 60 non-NDLTD universities
- *2002: U of West Virginia/Morgantown?*
- *2003: U of Montreal?*

4th International ETD Conference: CalTech, March 21-24, 2001

- 4 plenary sessions
- 15 breakout sessions
 - Implementation: graduate schools, libraries, scholarly societies
 - Technical; measures; archiving; publishers
- 12 networking lunch tables
 - Getting started; formats; vetting, archiving, collaborations, intellectual property; future
- NDLTD Steering Committee
- UNESCO Guide

Attendees not yet NDLTD members

- Arizona State University
- Cal State, Fresno
- Cal State, L.A.
- Cal State, Stanislaus
- Carnegie Mellon U.
- Cinemedia
- Eastern Michigan U.
- Ex Libris
- Florida Center for Lib. Auto.
- George Mason University
- Georgetown University
- Georgia Southern University
- Helsinki U. of Technology
- Indiana U. of Pa.
- INTIF-Agence de la Francophonie
- Jet Propulsion Lab
- Johns Hopkins University
- Los Alamos National Laboratory
- Korea Ed. & Res. Info. Service (KERIS)
- Library of Congress

Attendees not yet NDLTD members

- Michigan State University
- MindCrossing
- Montana State University
- National Hellenic Res. Foundation Northeastern University
- Northwestern University
- Ohio State University
- PUC Rio de Janeiro
- San Francisco State U.
- ScienceServer LLC
- Simon Fraser University
- State U. NY at Buffalo
- Texas A&M University
- Texas Tech University
- UMI
- University of Arkansas
- U. Ca. Irvine
- U. Ca. Santa Barbara
- University of Central Florida
- University of Chicago
- Universidad de Chile
- University of Duisburg
- University of Geneva
- University Lyon 2

Attendees not yet NDLTD members

- University of Madras
- University of Maryland
- University of Michigan
- University of Minnesota
- Université de Montréal
- University of Nebraska at Kearney
- University of Nevada, Las Vegas
- University of New Orleans
- University of Pittsburgh
- University of Puerto Rico, R.U.M.
- University of Rochester
- University of Southern California
- University St. Gallen
- University of Texas
- Southwestern Medical Center, Dallas
- University of Toronto
- University of Washington
- Washington State U.
- Washington University
- Wright State University
- York University

UMI and the NDLTD

- **Generally, independent decisions**
- 1987 UMI workshop was first to explore ETDs
- UMI wrote support letter for US Dept. of Ed. proposal
- UMI is on Steering Committee.
- We are collaborating on:
 - accepting electronic author submissions
 - standards (e.g., representation)

NDTLD encourages Standards

- Help students learn relevant skills
- Help reduce life cycle costs of ETDs
 - For students, universities, readers
- Support preservation
- Support access
 - Across all NDLTD members
 - Across all languages
 - At least some support for everyone
 - Good support for those following standards

NDLTD encourages Standards

- Documents, PDF, XML, SGML, HTML, LaTeX, multimedia...
- Metadata: MARC, DC
 - ETDMS (TheMs)
- Databases: workflow, citations
- Protocols: federated search, Z39.50, OAI

NDLTD members support NDLTD

- Share metadata
 - ETDMS
 - MARC
- Run OAI to harvest metadata
 - locally or consortially

VT Supports NDLTD

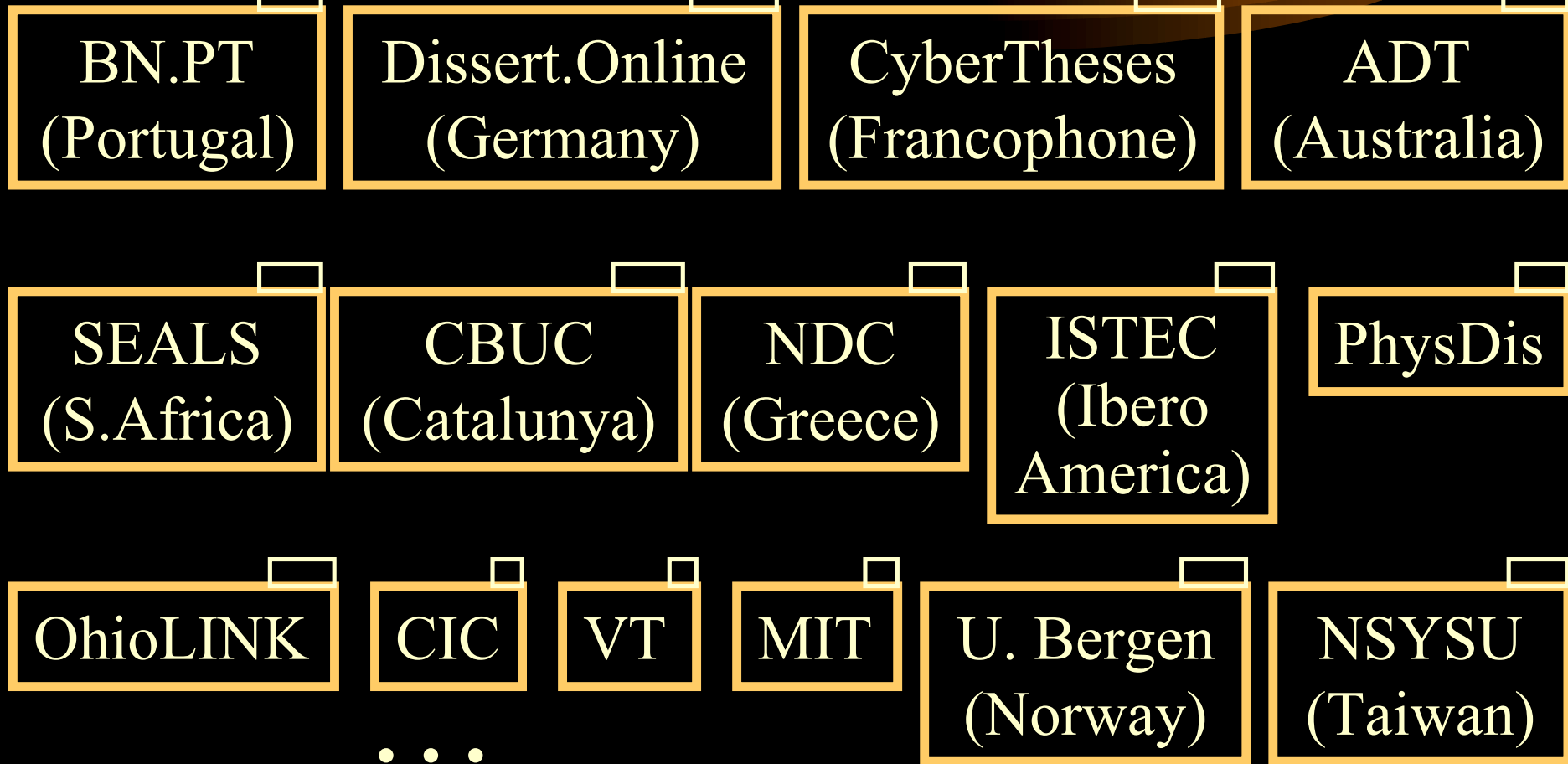
- Produce ETDMS
- OAI data provider
- Interface with authority services

OAI Philosophy

- Self-archiving = **submission** mechanism
- Long-term storage system = **archive**
- Open interface = **harvesting** mechanism
- **Data provider** + service provider
- Start with “**gray literature**”
 - e-prints/pre-prints, reports, dissertations,
...

OAI-in-a-Box: ETD Perspective

www.theses.org



A Union Catalog for NDLTD

VTLS Service to NDLTD and Virtua

- Help students and researchers find ETDs
- Combine ETD metadata from multiple sources in multiple languages
- Provide high-level centralized user services, e.g., searching

Union Catalog for NDLTD: Goals

- Create a global union catalog of all theses and dissertations for NDLTD members and others
- Provide a single location for searching ETDs
- Single searchable database for ETDs in all languages (possible with Unicode)
- Participating institutions will provide metadata for the union catalog with a link to their ETDs
- Institutions host their own ETDs

A Union Catalog for NDLTD

- Union Catalog Agency will
 - host the union catalog
 - store the original ETD metadata (MARC, ETDMS)
 - convert metadata
 - to Unicode
 - to standardized format for loading
 - create database
 - create search indexes for database
 - provide web-based client to access database
 - provide Z39.50 server support for database

Union Catalog for NDLTD Implementation Plan - Pilot Project

- Time Frame: January - March 2001
- Five Institutions to participate
 - Three Latin based sites
 - Two in foreign languages using non-roman scripts
- Test the following aspects
 - Ease of data submission
 - Completeness of mapping
 - User interface for searching and ETD retrieval
- Requirement - submit metadata for 10% of ETD collection; test results

Union Catalog for NDLTD Implementation Plan - Test Project

- Time Frame: April - June 2001
- Ten Institutions to participate
- Test the following aspects
 - Completeness of search system
 - Ease of updating previously submitted data
 - User feedback
- Requirement - submit metadata for 100% of ETD collection; test results
- U. of New South Wales agreed to participate at Cal Tech ETD conference – four others needed

Catalog for NDLTD

Implementation Plan - Go Live Union

- Time Frame: July 2001
- Participation in Union Catalog open to all members of NDLTD
- Free access to union catalog to others for initial period
- Develop long term support strategy

OCLC supports the NDLTD

- Identify ETDs in WorldCat
- Through OAI make available ETD WorldCat metadata in both DC and ETDMS
- Provide authority service for both personal and institutional names
- Coordinate with other authority services

Responsibilities

- Handle local education and collection
 - Contact information for implementors
 - Archive
- Utilize standards
 - Metadata: MARC / DC-based specifications
- Share metadata
 - Union services, mirrored services
- Allow access
 - www.theses.org / www.dissertations.org
 - Open Archives Initiative (www.openarchives.org)

NDLTD Project Goals

- Documents: PDF + XML + ...
- Metadata: MARC21 + ETDMS
Encoded in XML, XML+RDF, using OCLC's name authority system
- Search based on harvesting



ETDs at Virginia Tech

founder of the NDLTD

- A Little History
 - Faculty governance approved 1996
 - Shepherded by Graduate School
 - Conditions before requirement
 - Digital-only required Jan. 1, 1997
 - Library receives, stores, provides access
 - Graduate School approves
- How many: 3075
- Access options: all, some, none, mixture
- Who cares?
 - US domains accessing VT ETDs
 - International domains accessing VT ETDs

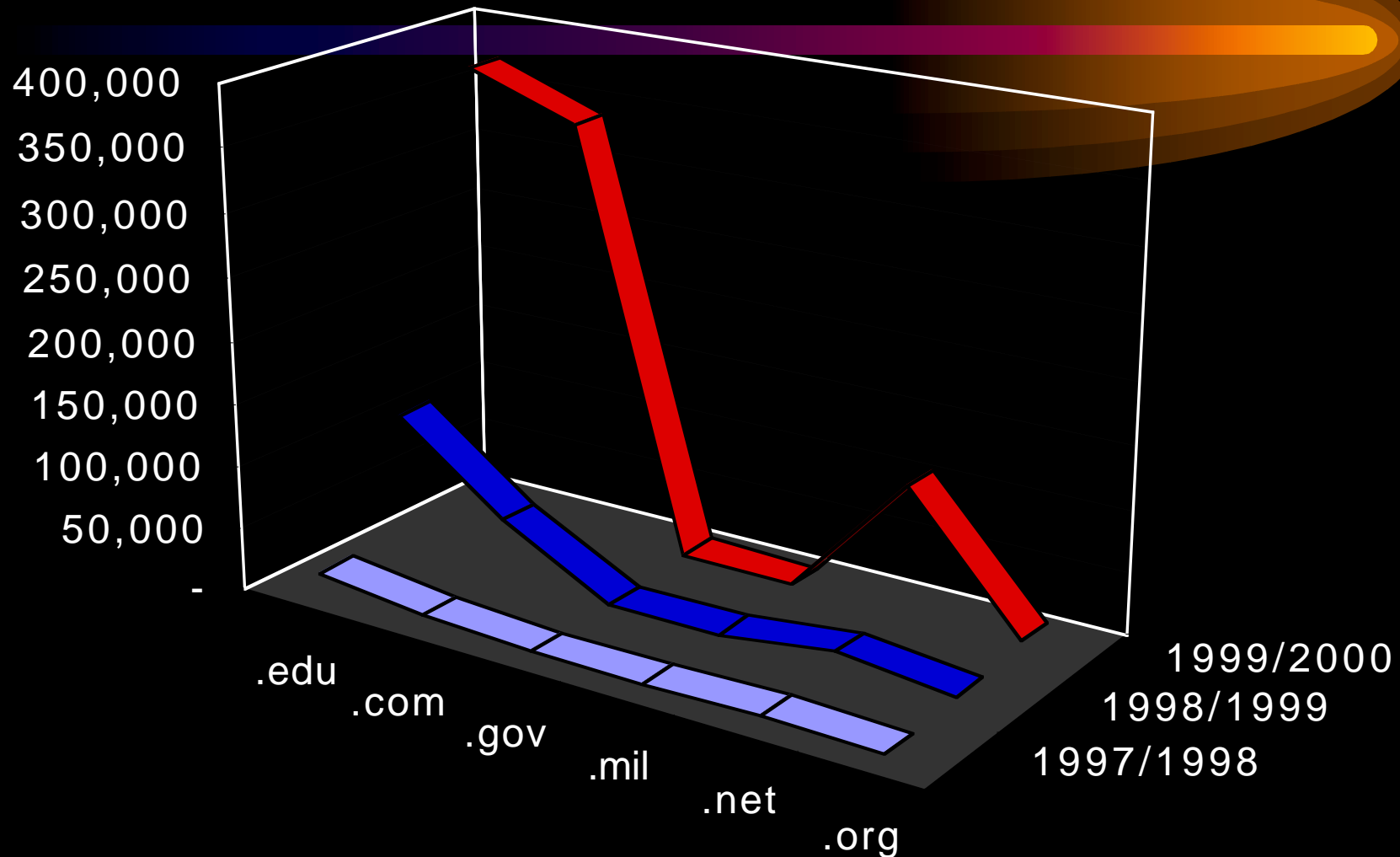
Access to VT ETDs

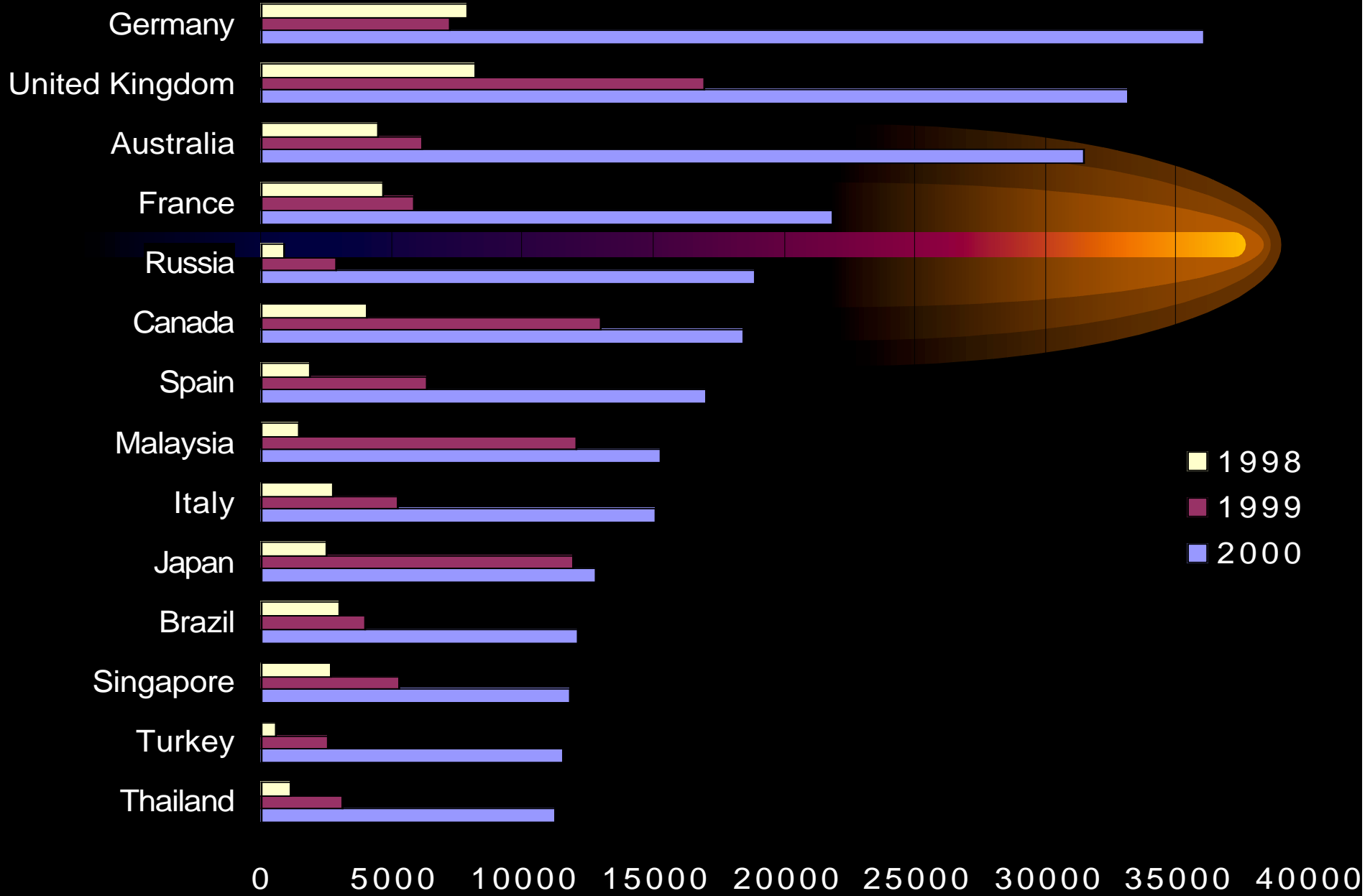
	1996	1997	1998	1999	2000
Total requests	37,171	247,537	465,974	1,190,113	1,894,510
Daily requests	102	685	1,722	3,016	5,176
ETD requests	4,600	72,854	244,987	671,981	734,807
Abstract req.	25,829	112,633	177,647	217,796	320,273
Hosts served	9,015	22,725	28,022	35,593	105,632

1990-1994: paper TD circulated 2-3/yr

- VT theses submitted 1990-94: combined average circulation was 2.24/yr per copy
- VT dissertations submitted 1990-94: combined average circulation was 3.2/yr per copy

Libraries Serve more ETD Users





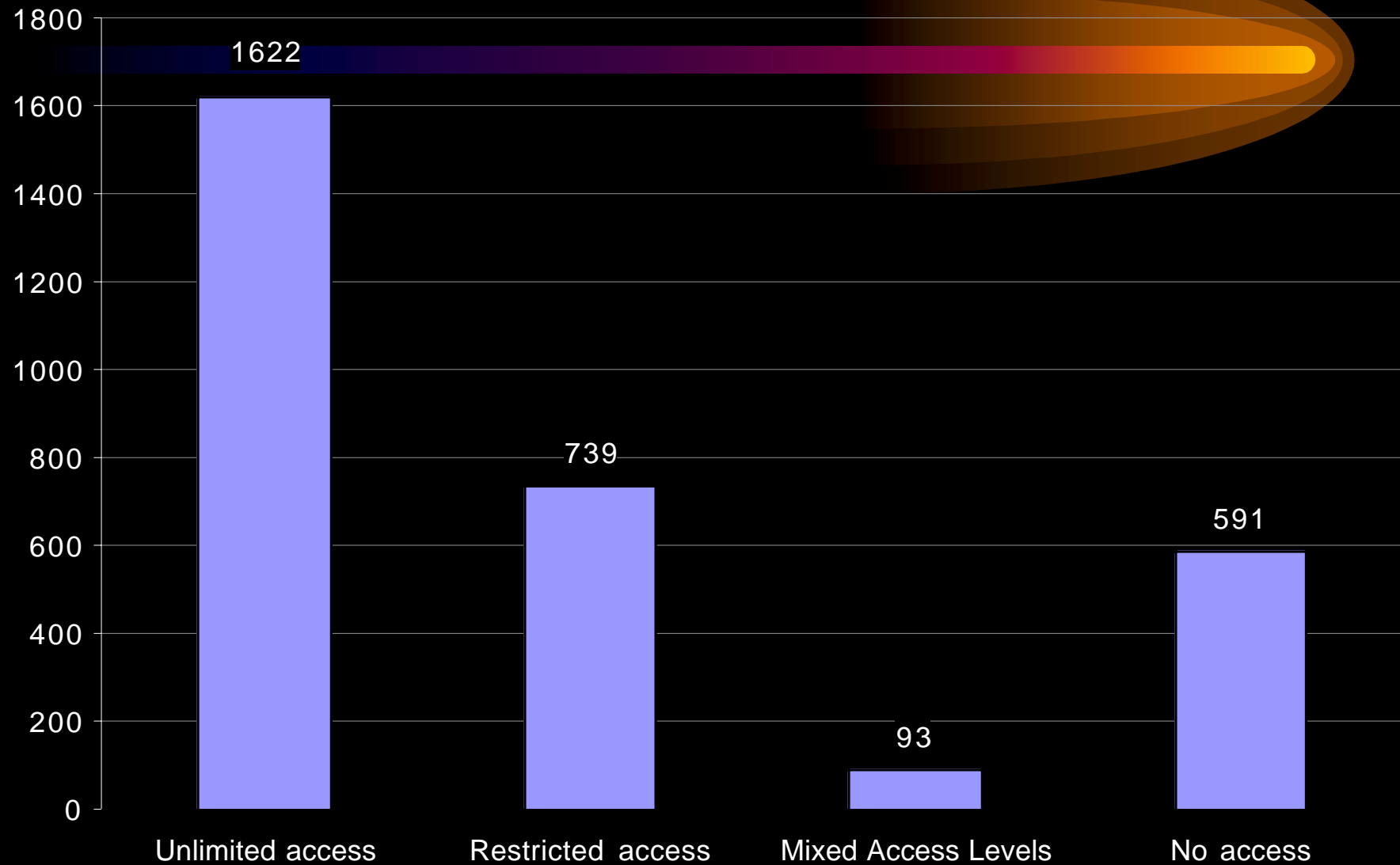
1998
 1999
 2000

Ensuring Access to VT ETDs

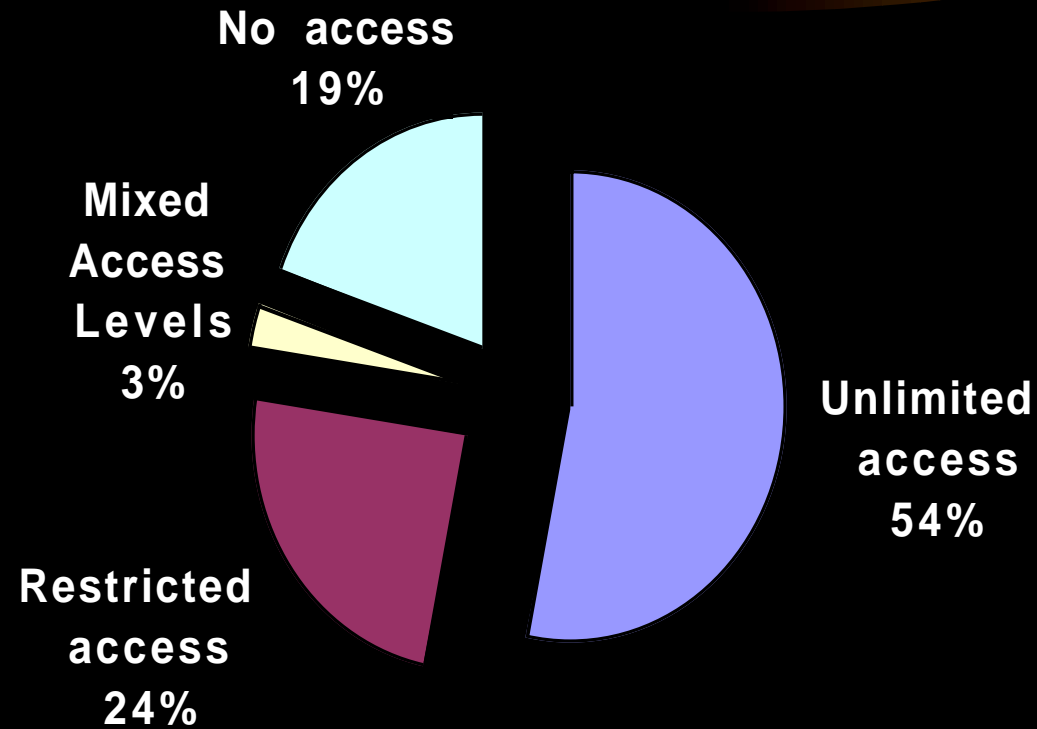


- Every 15 minutes back-ups made of newest, not-yet-approved submissions
- Hourly back-ups of newly approved ETDs
- Weekly back-ups of entire ETD collection
- Multiple copies stored on-site and off-site

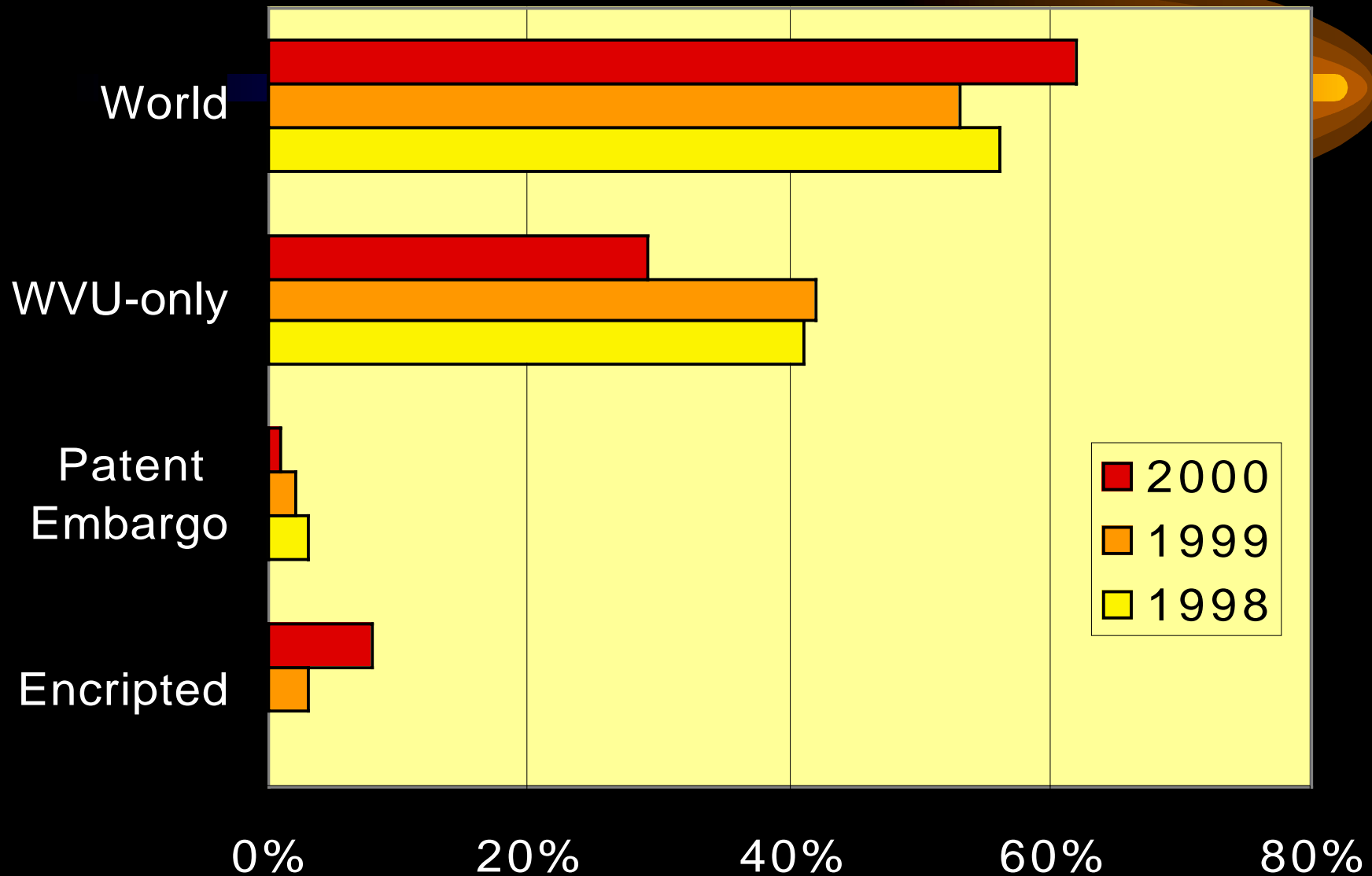
VT ETDs as of March 19, 2001



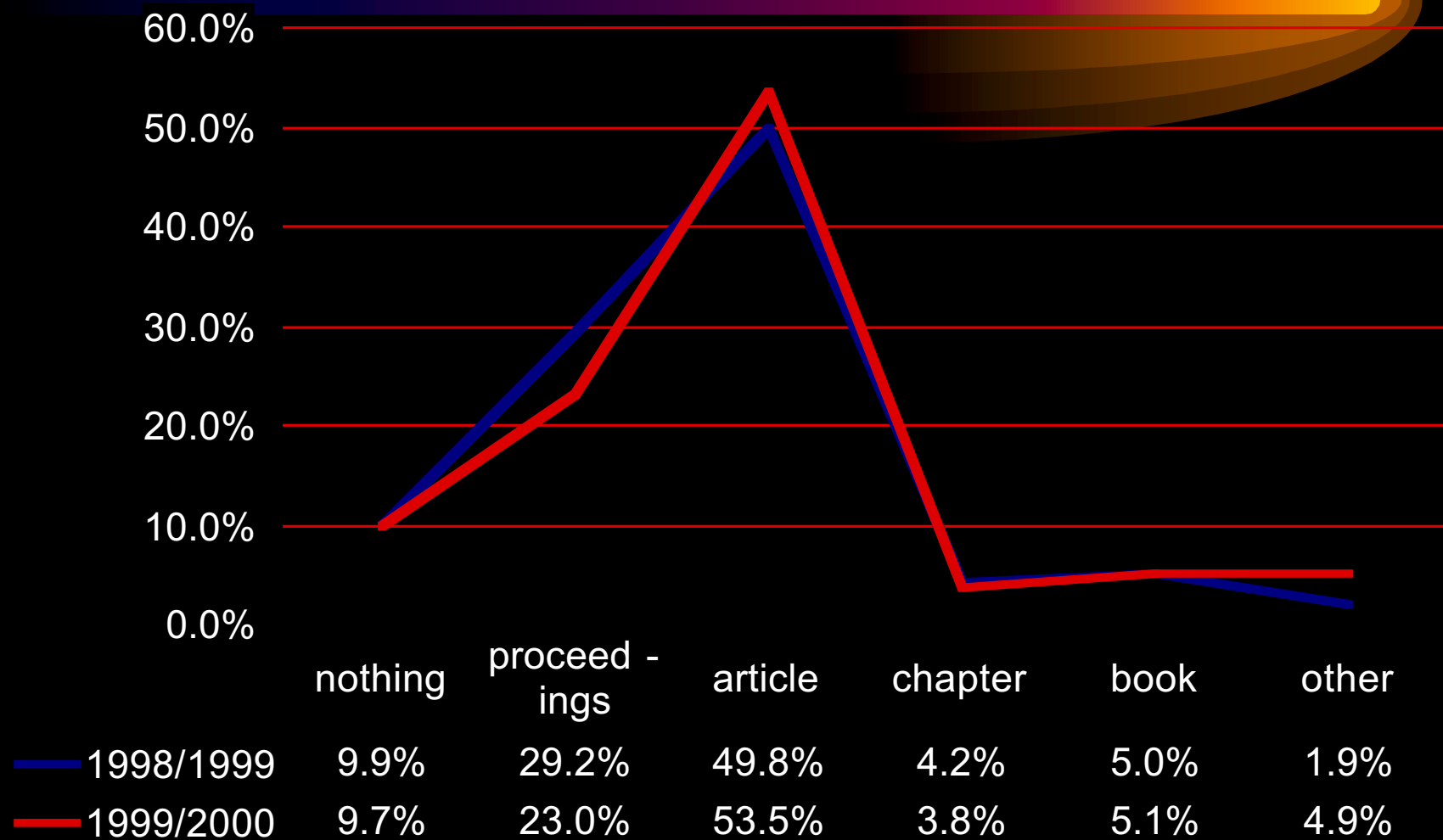
VT ETDs as of March 19, 2001



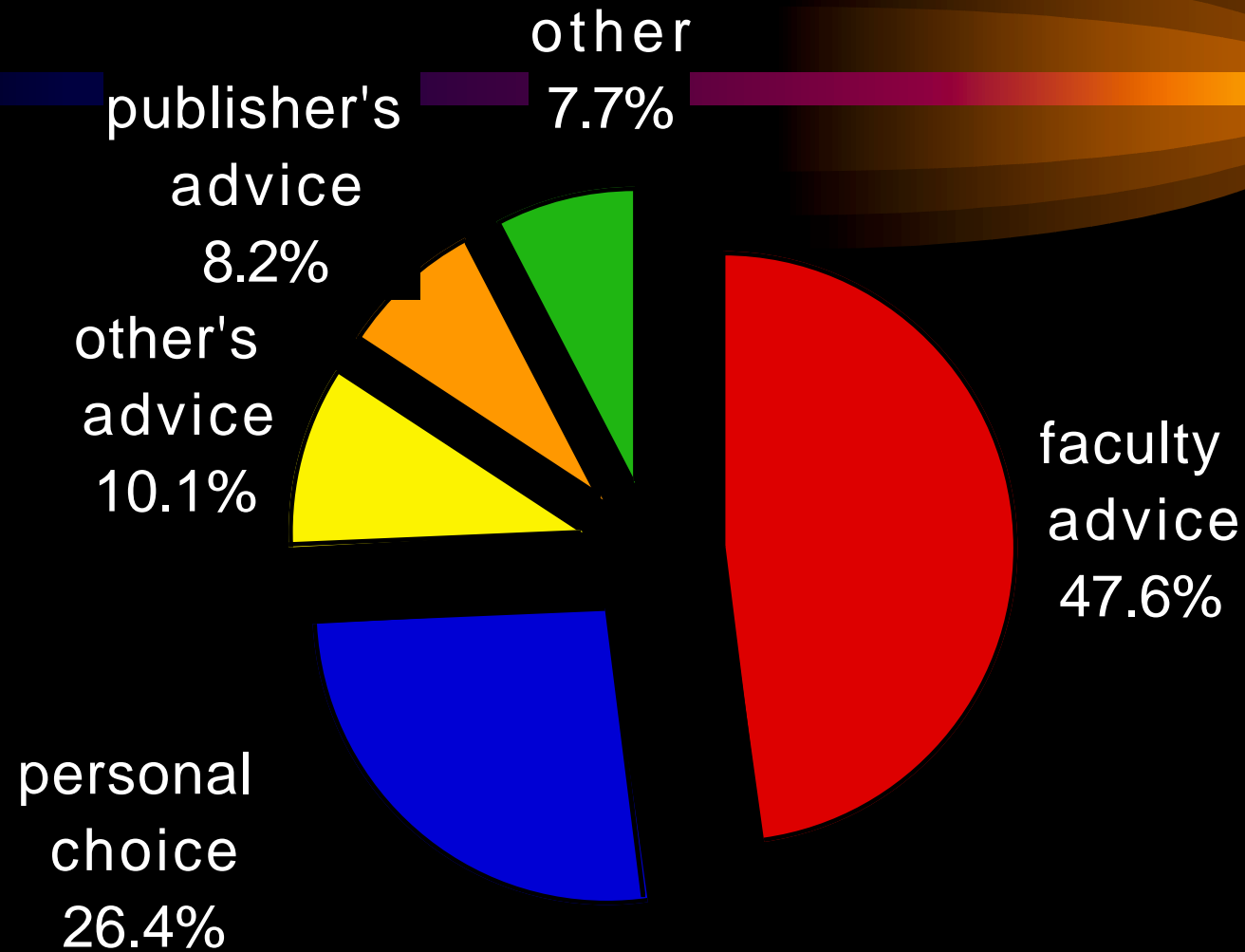
Availability of WVU ETDs



What do graduate students plan to publish from their ETDs?



Why VT Graduate Students Restrict Access



VT Graduate Student Alumni Surveys



- 21 had published by fall 1998
- 48 had published by fall 1999
- **0% encountered publisher resistance** to accepting manuscripts derived from their ETDs

1999 Publishers Survey re ETDs

Joan Dalton, U. of Windsor

- 200 scientific journals: 29% response rate
 - 94% had policy on prior publication
 - 68% did NOT have a policy specifically for online publications
- 10% -- ETDs is published if its on the Web
- 4% -- would not consider a manuscript from ETD
- More a perception of a problem than a problem

2001 Publishers Survey re ETDs

Nan Seamans, Virginia Tech

- 160 Science/Technology Studies journals:
55% response rate
- Majority have prior-publication policy
- Majority do NOT have a policy specifically for works available online/Web
- 18% -- ETDs constitute prior publication (+4% from Dalton's survey)
- Typical publishers' comments: impermanent, reworked/new

Publishers and ETDs

- **Concern** of faculty and students that plan 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 faculty, publishers, publisher associations to increase access

If you're thinking about ETDs

- Participate in NDLTD and its conference
- Engage local community broadly
- Find local leaders, advocates, and champions
- Implement a process that is easy and inexpensive for students, faculty, and staff
- Integrate the process with other campus efforts
 - evolve technology and education
- Run Open Archives, register with www.theses.org, and share metadata

Library Goals compliment ETD Goals

- Improve access to information
 - Better turn-around time
 - Always available
- Save space
- Reduce work (i.e., save money)
 - catalog from e-text
 - eliminate handling: mailing to UMI, bindery prep, check-out, check-in, reshelving, etc.

Costs/Savings at VT

- Graduate School stopped shipping 3000 copies of paper TDs/year to the library
- Library stopped handling (binding, shelving, circulating, etc.) 3000 copies of TDs/year
- 166 ft of shelf space saved/year by the library
- VT used existing library resources
 - vs. start-up costs for staff, hardware, software from from a zero-base estimate: \$65,000

Support Available from Virginia Tech

- Software, documentation, tech support
- Email, listservs (etd-1@listserv.vt.edu, -eval, -grad, -library, -technical)
- Evaluation: instruments
 - <http://lumiere.lib.vt.edu>
- Analysis: statistics from log files
 - http://scholar.lib.vt.edu/www_stats

Support for NDLTD

- OCLC
 - Hosting Jan. 9-10 DC-XML metadata std meeting
 - Financial support (student wages, travel): white paper
- VTLS
 - Union catalog server: multilingual, MARC/DC-XML
 - Machine, software, support staff
- UNESCO
 - ETD Guide in English, French, Spanish: book, Web
 - Scholarships to attend ETD Conference
 - Assistance with regional workshops: Columbia, ...