Access to Academic Theses and Dissertations

Amy Wood
Director of Technical Services, CRL
&
Marie Waltz
Special Projects Librarian, CRL

Feb. 24, 2010
CRL – the beginning
Dissertation Collection

- 800,000 dissertations
- Core from original members’ deposits
- Core expanded through exchange and deposit programs
- Demand purchase program
Dissertation dates
Oldest Dissertation at CRL
CRL Dissertations - Languages

German, 536,042, 67%
French, 136,238, 17%
English, 73,626, 9%
Dutch, 15,745, 2%
Latin, 9,333, 1%
Swedish, 8,041, 1%
Russian, 4,788, 1%
Spanish, 3,671, 1%
Other, 10,063, 1%
Access to the Collection

CRL Catalog Dissertation Scope

Features:
• Special browse by Country and Awarding Institution
• Traditional keyword, author, title and other searching and browsing capabilities
• Advanced searching and limits

http://www-apps.crl.edu/catalog/dissertationSearch.asp
Digital Access to the Collection

- Digital access begun in 2007
- Digitized over 2,700 dissertations
- 120 dissertations by nobel laurates and numerous titles by notable scholars
Improved Access through Digitization

Most used after digitization – 69 times

Most circulated – 16 times
EThOS

For the User

Features:
• Single point of access to UK dissertations
• Allows individuals to download or obtain copies unmediated
• Cost of digitization paid by awarding institution or first requestor
• Downloading is free
• Value added services at a cost

http://ethos.bl.uk/Home.do
CRL workflow for EThOS dissertations

1. Diss request received via ILL
   - Ethos database checked for title
     - Title available for immediate download
     - Title not yet digitized but EThOS will digitize free of charge
     - Title not yet digitized but EThOS will digitize for a fee
     - Title in EThOS database but EThOS can not digitize
       - CRL orders and pays
       - Ethos informs CRL when title is ready for download
       - Email EThOS to ask if we can pay for the digitization
       - Contact awarding institution to see if we can obtain a copy
       - Yes
       - No
         - CRL informs ILL dept that awarding institution will not supply
         - CRL sends request back to ILL informing of availability and how it will be delivered

2. Member ILL dept informs patron
Ensuring Author Rights

2. Copyright declaration - to be signed by all applicants requesting copies of material still in copyright.

(a) I hereby declare and undertake that:

I understand that copyright in the thesis remains with the author unless otherwise identified in the body of the thesis itself, and

(b) I will not plagiarise from the thesis in any of my own research outputs or publications or otherwise; and

(c) I will ensure that the use of the thesis is duly acknowledged in any of my research outputs or publications and quotations from the thesis are correctly referenced; and

(d) The research project identified in Section 3 is not commercial research and I am not receiving a fee for the carrying out the same; and

(e) I will only use this thesis for the purpose of carrying out my research project identified in Section 3 unless I have the prior written approval of the University of Southampton and the Author; and

(f) I will not print more than one copy of the thesis; and

(g) I will not make any further copies of the thesis, whether electronically or in print form; and

(h) I will not supply a part of or the complete thesis to any person in print or via email or in any other format, and I will not post the thesis (or a part of it) on any website; and

(i) I will destroy the CD containing the thesis and any electronic copy of the same on completion of my research.
The rest of the world

- Check institution’s library catalog to verify dissertation
- Check to see if dissertation is online through catalog or institutional repository
- Check to see if dissertation is online in regional or national repository or through national library
- Check institution’s library web site for information on obtaining a copy of dissertation (usually through document supply or department that handles reproductions)
- Do web search to see if dissertation is online or to find author’s email to request a copy directly from the author
Institutional Repositories

- University of Helsinki
- University of Pretoria
  - [http://upetd.up.ac.za/UPeTD.htm](http://upetd.up.ac.za/UPeTD.htm)
National Efforts

- Australasian Digital Theses Program
- DissOnline
  - http://www.dissonline.de/index.htm
- EThOS
  - http://ethos.bl.uk
Welcome to National Libraries Website!

Welcome to the first interactive website of the National Libraries (NL) and the revised directory which has been made possible by a partnership between Texas A&M University Libraries (TAMUL) and the National Library of Hungary (NLH).

Note that the website was designed for a resolution of 1280x1024 and you might have to adjust your resolution accordingly!
Portals

- DART - Europe
  - http://www.dart-europe.eu/basic-search.php
- DiVA
Going Beyond

- OpenDOAR – Directory of Open Access Repositories
  http://www.opendoar.org/index.html

- ROAR – Registry of Open Access Repositories
  http://roar.eprints.org/

- NDLTD – Networked Digital Library of Theses and Dissertations
  http://www.ndltd.org/
Going Beyond – OpenDOAR
The Directory of Open Access Repositories

• Bridge between repository and service provider
• identify the contents of each repository
• create a list of repositories of interest
• maintain the list for currency and accuracy
• harvest the metadata and full-text re-use policies individually
• analyse the policies to check for data-mining/commercial re-use etc
• exclude those repositories without the correct rights policies
• Enable searches of content with custom Google search or with ROAR API
• http://www.opendoar.org/tools/index.html

http://www.opendoar.org/documents/beyond_the_list.html
Going Beyond – ROAR Registry of Open Access Repositories

- Offers ROARmap - *(Registry of Open Access Repository Material Archiving Policies)*
- Universities and research institutions who have adopted a mandate to provide open access to their own peer-reviewed research output are invited to register and describe their policy in ROARMAP.
- Typical entry user sees
  - Running DiVA, based in Sweden and is registered as Research Institutional or Departmental
  Registered on 2005-12-08
  Cumulative deposits: 1471 total [table] [graph]
  Daily deposits in last year: 51 days of 1-9, 3 days of 10-99, 0 days of 100+ [table] [graph](PNG format)] [interactive graph](requires SVG format support)]
**OAI Interface:** Identify List Metadata Formats List Sets [harvest status]
100% freely accessible fulltext (*estimate*)
Going Beyond – NDLTD
Networked Digital Library of Theses and Dissertations

- “international organization dedicated to promoting the adoption, creation, use, dissemination and preservation of electronic analogues to the traditional paper-based theses and dissertations”
- Guide for electronic theses and dissertations
- Search content of union archive with VTLS
  http://thumper.vtls.com:6080/visualizer/
- Information about research and experimental services and community activities

http://www.ndltd.org
Questions?
ProQuest

A Case Study of A Digital Repository
What is a Digital Repository?

Services

Administration

Digital Content
Trusted Repository Audit Checklist (TRAC)

TRAC criteria looks at:
1. Repository Management (Administration)
2. Data Management (Content, services)
3. Technology & Security

...For the purpose of certification and auditing digital repositories.
We hope TRAC may become an ISO standard soon, like OAIS
For more information on TRAC as an ISO standard

- [www.digitalrepositoryauditandcertification.org](http://www.digitalrepositoryauditandcertification.org) – A wiki for the work being done.
- [http://mailman.ccsds.org/cgi-bin/mailman/listinfo/moims-rac](http://mailman.ccsds.org/cgi-bin/mailman/listinfo/moims-rac) – email list
National Science Foundation Grant

NSF awarded CRL a grant to study eight long-lived data repositories.
Eight Repositories

- The Arabidopsis Information Resource (TAIR)
- Chemical Abstracts Service (CAS)
- National Center for Atmospheric Research: Earth Observing Laboratories (NCAR/EOL)
- Sloan Digital Sky Survey
- U.S. Geological Survey (EROS)
- National Opinion Research Center (NORC) General Social Survey (GSS)
- UMI Dissertations Publishing
- Associated Press (AP)
What is a Case Study?

An in-depth exploration of one particular case (situation or subject) for the purpose of gaining depth of understanding into the issues being investigated.

- ReCAPP Research Glossary
Our Case Studies look at:

- Key decisions
- Practices developed
- Legal obligations
- Community expectations & needs
- Interoperability/Data integration
- Incentives for acquiring content
Why did we look at ProQuest dissertations?

- A unique long-lived digital collection
- Primary sources are important to CRL’s members
- An official repository for dissertations as designated by the Library of Congress and the National Library of Canada
History of UMI/ProQuest

1938 - Eugene Power starts UMI “Edition of one”
1951 - ARL endorsement
1962 - Xerox buys UMI
1985 - Bell & Howell buys UMI
2001 - PQE stock sold on NYSE
2005 - NAPC formed with microfilm assets
2006 - Cambridge Info. Grp (CIG) buys ProQuest
2009 - NAPC sells dissertation microfilm back to ProQuest
Who is UMI’s Designated Community?
What are their needs?

- Authors
- Universities
- Libraries
- Researchers

- Library of Congress
- National Library of Canada
- Database producers/vendors
What services does UMI offer for its dissertations?

- A searchable dissertations interface
- Full-text dissertations
- Easy citations
- Many channels for distribution & purchase of dissertations
- Open Access
- Author copyright and embargo services
- A persistent digital archive for dissertations
- Ph.D. Verification
Sources of Income for UMI from dissertations

- Subscriptions (mainly libraries)
- Dissertation Sales (mainly researchers)
- Authors’ payments (one time payment)
Dissertation Publishing Process Flow

**ProQuest**

1. **Paper and Electronic Dissertations Received**
   - Unpack and sort dissertations, check for completeness and apply tracking barcode

2. **Scan Vendor**
   - Scan paper dissertations per requirements
   - QC for missing pages

3. **ProQuest**
   - Key metadata into production workflow system (Exodus)
   - Scan Abstract, OCR, add SGML tag and index terms
   - If Paper: Pack and ship Dissertations to scan vendor
   - If Native PDF: Scan Abstract, OCR, add SGML tag and index terms
   - Loading system (EIR) automatically merges metadata with images and loads database
   - PDF wrapper assembles TIFF scans in PDF

4. **Scan Vendor**
   - Prepare Print Index indexes for DAI, MAI, ADD and CDI
   - Upload index and abstract records to PQDD and PQDT

5. **Print Vendor**
   - FTP Print Book to print vendor
   - Books printed and bound by vendor
   - Books Shipped to Customer
   - Ship Confirm Returned to PQ

**Products:**
- ProQuest Dissertations and Theses (PQDT)
- Print Dissertation Indexes
- CD-ROM Dissertation Index
- Print on Demand Dissertations
- High Resolution Repository
- Books printed and bound by vendor
- Books Shipped to Customer
- Ship Confirm Returned to PQ
- FTP Print Book to print vendor
- PQDD & PQDT
- Print Book Vendor
Preservation Strategies

- Two copies of microform stored in climate & humidity controlled vaults
- TIFF’s and PDF’s in the online electronic vaults
- Daily backups
- Offsite digital storage
- Logs of retrieval failures
Successful Strategies

- “Edition of One” was a good idea at the time and continues to be one.
- The distributed cost model works
- Collaborations have been important: ARL, LC, NLC, etc.
- Uniformity of the product.
- Good relations with academic community
- Incentives for the designated communities
Google Dissertations and Stanford

- Electronic submission will reduce time and expense for authors
- Stanford Libraries will provide a persistent, stable URL to dissertations in perpetuity
- Makes it available in Google Scholar, versus Google's more generic index.

Stanford Continues:

- To Catalog dissertations in their online library catalog
- To archive the file's in the Stanford Digital Repository
Is this a threat to UMI?

1. Stanford will continue to allow students the option of submitting to ProQuest.
2. ProQuest is on the verge of signing a deal to put their dissertations into Google Book Search (not Google Scholar).
3. Universities want to maximize the dissemination and discovery of dissertations and putting them in Google book search helps widen the audience for dissertations.
Vulnerabilities

- Self-publishing/Open access
- More complexity content (larger size, more formats)
- Dissertations’ low market value
- UMI’s Preservation planning is not transparent, so we aren’t certain if it will work should there be a need.
Questions?
CRL Contacts

Amy Wood, Director of Technical Services, awood@crl.edu

Marie Waltz, Special Projects Librarian, mwaltz@crl.edu
Thank you for joining us today

Please take a few moments to fill out our followup survey:
http://www.surveymonkey.com/s/XPT5NQN

You can access this PowerPoint presentation here:
http://www.crl.edu/events/6519/follow-material

You will also receive a followup e-mail with this information.
Upcoming CRL Webinars

CRL Access and Collections Webinar
Wednesday, April 14, 2010, from 1:00 to 2:00 p.m. CDT.
This no-cost event provides staff and faculty at CRL libraries with a great overview on CRL collections, services, and programs.

61st Annual Meeting CRL Council of Voting Members Collections Forum
Friday, April 23, 2010, from 1:30 to 4 p.m. CDT
For the first time, CRL will open its Annual Meeting to all representatives of CRL libraries for a special afternoon collections forum, with a focus on print and digital repositories and access to humanities and social science databases.
For more information, visit www.crl.edu

To sign up for our online mailing list, visit www.crl.edu/connect