

## Networked Digital Library of Theses and Dissertations

Steering Committee  
meeting: September 19, 1997  
ARL Conference Room, Washington, DC

### *Participants*

David Balatti, National Library of Canada  
Joseph Boykin, Solinet Board of Directors  
George Brett, SURA  
Kenneth Crews, Indiana School of Law, Copyright Management Center  
Gail Dykstra, IHS Micromedia  
Ed Fox, VT  
Fred Friend, University College, London  
Neill Kipp, VT  
Kevin McGrath, Adobe (for Bill Lindsey)  
Gail McMillan, VT  
Bill Savage, UMI  
Craig Summerhill, CNI  
Peter Syverson, CGS

Participants introduced themselves at the video camera.

Fox: Welcome. Recap of Virginia Tech's history with ETDs, including meeting with UMI 10 years ago. Recent publicity generated in part because of graduate students' conference at Virginia Tech re electronic publishing, but really meant to focus on complaints about ETDs and controversy over prepublication vs. online access. Brought to the fore anti-computer sentiment, dislike of new requirements, and other emotional concerns. Deepest concerns are with publishing--an area where the Steering Committee (SC) can provide good guidance. On the positive side, brought publicity to the ETD project.

Summerhill: VT is holding ETDs to a higher standard than paper versions.

Brett: VT sold ETDs and their advantages too well.

Boykin: Faculty on other campuses were hesitant earlier than the VT faculty were leery.

Summerhill: Copyright, access issues; generations of scholars need to become accustomed to the new technology and availability issues that the ETD SC is aware of now.

Brett: Supportive materials -- works for teaching. More success with tailored groups

AA-level: "I'm George and I do ETDs."

OR, go for the department level where success is more likely.

Balatti: Were students aware of UMI's publishing function?

McMillan: Yes, but they don't see that that is the same as international Internet access. Copyright knowledge--students don't understand that they are creators and owners.

Fox: Discipline specific issues. Humanities and social scientists are more likely to want to publish books out of their ETDs, while scientists expect to publish articles and conference papers.

### **Participant Reports:**

#### **Adobe: Kevin McGrath**

Availability of Capture to convert legacy documents to ETDs

Brett: Is Adobe looking into how to deal with complex and massive documents?

McGrath: Adobe is working on this and has other software (e.g., Premiere)

Brett: Adobe should see the NDLTD as a great testbed

Summerhill: WYSIWYG as well as more robust SGML. Volume is not a problem. VT needs better support from Adobe.

McGrath: Have a campus liaison--one person to funnel all questions through.

#### **UMI: Bill Savage** ([www.lib.umi.com/dissertations](http://www.lib.umi.com/dissertations))

- 1) Most current 3 months (17-18,000 records) always available
- 2) Subscriptions to entire database of abstracts  
30 day trial with 70 schools (OhioLink using it on trial)

- With citation will be an icon that tells how to get a copy.
- 3) Full-text coming in mid-Nov.: Full PDF if submitted electronically, otherwise paper will be converted to TIFF wrapped in PDF; problems w/Capture--errors in chemical and physics formula (nothing is working perfectly). No longer OCRing. 600 dpi. About 30,000 documents as images. Institutions will have access to their own ETDs without charge. UMI will limit by IP address. Researchers can browse first 20 pages to determine if they want the full document.
- McMillan: Extended campus students deserve access also. VT would use its proxy server to make them identifiable to UMI.
- Fox: Improvements needed in Acrobat (2-passes like Distiller).
- Savage: UMI provides access to TIFF'd dissertations also free to home institutions.
- Fox: Reminded this Steering Committee that it agreed to encourage access to ETDs
- Savage: Current Research at \_\_\_\_: Will create a Web page listing ETDs along with subject categories, for each submitting institution. Will be available in December.
- 1998 plans: Graduate students can email UMI on dissertation defenses--which ones are in process. Graduate Schools would fill out this form. Useful to researchers and employers. Dissertation Express: Web access through ILL at the library. Receive paper unbound within 2 days for \$24.95. BS says ILL transaction is \$30.
- Summerhill: Libraries will find in 10-15 years that its cheaper and easier to have UMI do the archiving and provide the access to their institution's ETDs (than for the library to provide these functions and services).
- Savage: Saving digital library in 2 formats: CD-ROM and tape backup or hard drive backup. Sample tapes periodically and refresh as needed. As new media become available--standardized and stable, UMI will convert to that. Will microfilm anything that can be microfilmed--this is permanent and UMI's archival role.
- Summerhill: Physical storage devices will be less of a problem than the format of the ETDs. PDF is more of a concern than SGML. The future of PDF is less secure because of its single ownership.

#### **Council of Graduate Schools: Peter Syverson**

- 415 US members, 12 international; brings deans together at summer workshops and annual meetings. Eaton wrote letter to board asking CGS to get MORE involved. Board reviewed it last spring, July 1997 summer workshop (Gordon Batell, chair of CGS Board). Will continue to be a good strong forum for discussion of ETDs. Each December have update sessions; 1998 in Vancouver will have a plenary session to discuss ETDs, e-publishing, copyright. Will bring in a variety of people. How e-publishing is changing and being changed by graduate education. Special edition of *CGS Communicator* (3000 subscribers). Lay out issues and talk about the future.
- McMillan: What is your sense of how these deans are viewing ETDs.
- Syverson: They run the whole gamut. Higher education is slow to change.
- Summerhill: How much is published from theses and dissertations?
- Savage: Under 10%
- Summerhill: VT is doing so much of the valuable groundwork for other institutions.
- Syverson: Dean gives the impression that the students are adopting the new technologies. Surveyed graduate students chose 7 to 10 to do it on the Web than on the paper form that came to them.

#### **SURA: George Brett**

- ETD extension runs out in December. SURA will be involved after \$\$ run out. ETD could submit another proposal. Not a strong sense of support at SURA for ETDs, perhaps in part because GB heard from NCSU (John Ulmschneider) that they were waiting for us to contact them! (Has been passed on to Eric Lease Morgan). SURA may play a role of bringing the participants together. 13 states, 41 institutions = SURA with a wide continuum of educational institutions--some institutions are not well connected to the Internet and that's why GB asked for ETD package on CD-ROM. SURA could play the role of bringing a user group together.
- Summerhill: Is concerned about the sustainability of the digital library. It may be 2-300% more costly to maintain a digital library than the paper documents in a traditional library.
- McMillan: VT library's commitment re keeping access.
- Brett: Standards are important for future considerations like those of conversion to new technology and new formats. Roles and responsibilities of institutions are very important also (Protocol).
- Friend: Higher costs to sustain a library online than to build a library and shelve books.
- Kipp: Our data don't reflect that at all.
- Fox: Get the students to do it. Standards and protocols. But also Universities must retain their ability to have their own libraries.

Summerhill: VT could provide the role of a model for other institutions and SURA could support that.

Brett: To some extent VT has done that, but it is fragmented.

Brett: Make these deliverables in addition to the list of collaborators.

Fox: Shift of resources is required and occurring but need more to take place. We have also packaged the workflow.

Summerhill: need to have measurables as well as ... Remember the value added by having ETDs. Do the full gamete and measurables so that evaluation results are clear and others can benefit.

### **Solinet : Joe Boykin**

Solinet + SURA for MEL, Solinet didn't know what was involved for them in ETDs. Now perhaps Solinet is interested. May be a source of data about ETDs -- could also determine costs so that economic commitment by an institution could be known and contracted out or make any other decision based on real information (numbers). Solinet is looking for services, products, and revenue stream for its future. S could be a one stop shop for its members if they chose not to do it themselves.

### **University Theses Online Group: Fred Friend** (UTOG = 12 universities in the UK)

UTOG Web site is at [www.cranfield.ac.uk/library/UTOG](http://www.cranfield.ac.uk/library/UTOG)

Follett Report was behind starting this group. Applied for library funding to scan theses and dissertations but was turned down because the funding authority didn't think that universities were ready to get into the scanning business. Then looked to future and longer range planning. Surveyed authors and supervisors of theses--general questionnaire: re sources used to write theses, plagiarism (are you worried that yours will be stolen); how do you feel about it being electronic? Results of survey were encouraging. Plagiarism and copyright were not a big concern.

Seminar held in June, invited reps from all UK universities. Over 100 attended--twice expected turnout.

Showed the VT ETD video. Had frank discussions at the seminar. Worries were about: why are you bothering to do this? The value is in the work that goes into the research--the academic work. The content may be good and then it will get published. If it isn't good it won't get published and therefore doesn't need to be digital at all.

Trying to convince a number of universities to take part in a pilot. Would ask students to submit text in dual formats--paper and electronic. Cranfield is scanning theses and dissertations as part of a European project. UTOG wants to receive it electronically.

UTOG was attracted by Virginia Tech's vision and plan. Friend would like to see UK universities join the initiative, but today this won't happen. In the long term it will happen.

Balatti: What about interlibrary loan.

Friend: Electronic index of theses and dissertations now available. British Library handles this and it is a large operation.

Fox: Need for statistics in this area. Friend has this in mind for the UK.

Uses of ETD databases--find a graduate school, see what a particular professor has done, see if the topic of a future ETD has already been done, ETDs can be models for future graduate works.

### **National Library of Canada: David Balatti.**

In Canada, 6-12 months ago some regional discussion led to talk more broadly. November there will be a discussion. Differences of approach: some want the activity to be at the university level based on enthusiasm of a department. Libraries want a national coordination--a tradition in Canada. C Association of Research Libraries, C Association for Graduate Studies. Cautious. New contract with UMI to continue microfilming but a shorter term contract (based on VT initiative and also UMI plans). Wants data about costs--what will it cost to get involved. [Note: Preliminary costs estimates are now available at <http://scholar.lib.vt.edu/theses/data/setup.html> per McMillan 4/6/98]

**VT Library Report: Gail McMillan:** Reviewed her handouts, some of which follow.

### **What do the Library's log's files for 1997 say about VT ETDs?**

*Jan. 4, 1997 - Aug. 28, 1997 statistics from [scholar.lib.vt.edu/theses](http://scholar.lib.vt.edu/theses)*

Total successful requests: 99,388 (6,559)

Average successful requests per day: 421 (937)

Number of distinct files requested: 2,391 (770)

Number of distinct hosts served: 13,315 (632)

Number of new hosts served in last 7 days: 434

30% of requests were for .pdf files (i.e., full ETDs)  
 50.4% of requests were for .html files

65% of requests from

.edu	42.5%	42,234
.com	19.4%	19,247
.org	1.6%	1,599
.mil	1.0%	962
.gov	.6%	637

Top 5 international (academic?) sites (of 79)  
 (accessed 5.9% of files)

1639	from United Kingdom
1320	from Germany
1154	from Canada
883	from South Korea
836	from Australia

As of September 15, 1997:

409 VT ETDs available for online access (about 2/3 have unrestricted access)  
 137 ETDs have VT-only access  
 272 VT ETDs have unrestricted access  
 89 approved VT ETDs are inaccessible except by the Graduate School

per Neill Kipp: mean = 2776.48 Kb                      s.d. = 4987.93 Kb (standard deviation)  
 smallest ETD = 87 Kb    largest ETD = 40 Mb

A better measure due to this large standard deviation: median = 1003 Kb

**Multimedia content of the VT ETD collection: 327 files**

23 movies	225 still image files
8 Quicktime	154 GIFs
4 AVI	35 TIFFs
11 MPEG	36 JPEGs

85 VRML .wrl files (1 ETD)                      1 378M Macromind Director file

People come to McMillan with concerns about library issues:  
 Should we be making the compromises with publishers that we are?  
 Education needed because they don't understand what limiting access means.  
 No access to other students, faculty through interlibrary loan.

**Copyright discussion re library issues**

Crews: Make the student/author responsible for notifying the university if the ETD cannot be released, otherwise university will release it. CS said this also.  
 Fox: We are gathering statistics from log files as well as from students.  
 Fox: Approval form  
 Friend: You wouldn't need the letter from publishers in the UK. Students can publish chapters before dissertation is approved.  
 Crews: Prepublication is how publishers feel about ETDs on the Web.  
 Syverson: Faculty loyalty is to the association before the university.  
 Fox: Need to air issues and address through a user group meeting, perhaps.  
 Summerhill: Gets students caught between university and publishers. Educate students about their copyrights as a service to the students.  
 Crews: Multitude of issues, name of the effort is digital library, but in effect you are a digital publishing house. Consequences: fewer alternatives than otherwise. Like other publishers, don't try and answer these questions. We can provide a service: identifying the questions--full disclosure is a service. Survey publishers (leaders) of scholarly journals re policies re ETDs today. Finally, the decision lies with the authors  
 Summerhill: Agrees with Crews completely. But you have an element of [the perception of] coercion with students, who have different people influencing them. Other authors don't have this opportunity for coercion.

Friend: Opportunity to train students and help them provide more knowledgeable decisions.  
Dykstra: The ETD project is trying to get something that is unattainable and if you got a lot of permissions from publishers, what about all those that you don't get. Why not focus on success? Rather than try and convince all publishers.  
Dykstra: We have ETDs and some are not accessible. That's the way it is.  
Summerhill: ETDs are not published if available from the library as the paper copies were.  
Crews: Required commitment from each NDLTD member as a minimum -- e.g. put it in the database, provide some level of access, equivalent of library shelf access, must be usable in a way consistent with current use of other library materials--i.e., copyright and fair use for libraries.  
Fox: Our goal is also to make students aware of the issues involved--part of the educational goal of the ETD project. Therefore, we can't ignore the problem. Likes Crews' idea to survey publishers to collect information. Should meet needs of students and faculties.  
Boykin: We want to make ETDs available without restrictions. How important is this goal? If you can't gather the permission from publishers, then students will restrict access.  
McMillan: Why not survey the universities and libraries to gather more favorable data re Do the ETD thing.  
Fox: UMI advice?  
Savage: Relationships now with permission to republish newspapers and books. We publish ETDs. Dissertations are required to be "published." This is an economic issue, not a copyright issue. Publishers want first rights. UMI gets permission from ACS to "publish" dissertations in paper and microfilm--not electronically. "UMI hasn't raised its prices in 5 years."

**VTDL report: Ed Fox**

SiteSearch donated to our campus. URNs and federated searchers across institutions. Cooperation with UVa to do this also.  
Dissemination of information: In person is more effective than giving a talk to a group. Teleconferencing is good method also.  
Syverson: Video needs to be shortened to 20 minutes. Liked the nature and sense of it--from us to you. [video format NTSC for Canada and US, UK needs PAL]  
Kipp: New fliers for "early majority" Look at productivity and bottom line. Does the new flier target and attract this audience?  
Syverson: Improve graduate education. Are they really going to save money? John should show this to other deans and ask them. Graduate Deans are comfortable with UMI, so say these dissertations still go to UMI? Does this upset the process. Say that this works with UMI--so why do it in addition to UMI in 8 seconds.  
Summerhill: Some of this is commercially viable. That's OK. You're making available all that is not commercially viable. We are worrying too much about the black hole--in between unlimited and totally restricted access. You have a hook on selling the piece of immortality to students. If it isn't commercially viable, they can publish and contribute to the academic enterprise.  
Savage: thesis.com is now available. Look at his site registration on Internic. This person assigns an ISBN and registers them with amazon.com.  
Summerhill: We can do better than vanity publishing. We can promote these works for scholarly purposes.  
Crews: Royalties should be available to the authors. If you have advertising, there will be revenue and the authors should get some of it.  
Summerhill: 1 to 10,000 bytes from royalties.  
Kipp: Need comments from libraries.  
Summerhill: Saving \$ should not be the only argument.  
Balatti: and Friend: Saving \$ is important to libraries. Leave this on the front.

Fox: Funding opportunities: (but Fox won't address whether \$ will become a revenue issue for the ETD project now)

Interesting problems to be solved on the technical side. Sunday morning, Oct. 26 in Minneapolis--just prior to CNI (I could send Joanne for me). SGML, protocol, and digital library structures.

Funding opportunities: Joint EC and US D of E, 3 European countries and 3 US states could be funded. Due in May

Friend: A very good opportunity. DG13 has been funding the library projects, but there may be an international office in the EC.

Fox: Possibly next spring NSF will also have international digital library funds. Network traffic delays and currency (\$) issues can be addressed.

Technical issues of this project provides a chance to explore all of the hard problems in digital libraries, multilingual, multimedia, policies, workflow, searching methods, etc. There will be an NSF announcement soon re DL research proposal request. ETD will be one of Ed's focus. Wants to look at federated searching. Fox will be approaching members of this group for support.

Friend: GrayNet international conference about gray literature (which includes ETDs). That may be an opportunity to find more international opportunities.

Fox: NDLTD Steering Committee, therefore, you are advocates of ETDs and how to make them successful. This is a serious concern with a worldwide vision, not a VT project. UMI is a partner.

Fox: Is membership on the NDLTD Steering Committee complete?

Crews: Perspectives that should be invited, e.g., a graduate student--oops Kipp. Graduate School processor. Faculty advisor. Technical support--that's Kipp again.

Friend: ARL connection.

McMillan: Include a librarian without a technical background--reference librarian.

Crews: Invited advisors with different perspectives--rather than additional SC members. Publishers didn't want to participate so they must not think its important enough.

Summerhill: How about an Australian participant.

Fox: We'll have an article in an Indian journal so that may generate interest.

Steering Committee will meet 2/year since there's enough to talk about. Continue with March and September meetings.

User Group: Have it in conjunctions with another meeting. Graduate Schools must be in favor so this Dean or Provost, so have in conjunction with December CGS meeting in DC, next December in Vancouver.

Balatti: A short list of universities most likely to attend.

Syverson: Work through a regional group such as Council of Southern Graduate Schools. Feb. 20 is the next meeting (100 members) in a small southern city.

Kipp: gave Bill's football report.

Fox: Ivys + Berkeley, Chicago and Stanford: Small group of grad school deans: Seven Dwarfs vs. Big Ten?!?

Crews: We haven't talked much about options students have. They don't leave the local library the opportunity to work with an ETD like they can with a paper TD. We have done ourselves a disservice. We need to use different language.

Summerhill: Discovery can continue through UMI--unchanged. Institution must be able to honor requests as it would paper works.

Summerhill, Crews, Dykstra, Boykin, Friend, Syverson: Modify option 2 on the VT ETD Approval Form to allow libraries to do what they can with paper TD. [See library-friendly draft, unapproved form at <http://scholar.lib.vt.edu/theses/approval.htm>.]

Savage: LC gets 200 database inquires for theses and dissertations and 85 phone calls a day re theses and dissertations.

Dykstra: Let's keep the right perspective and appreciate that most graduate students do want their works to be available to the world without restrictions.

Fox: Thanks for coming and for your advice.

Meeting adjourned at 4 pm, September 19, 1997.

GMc 9/22/97

<http://scholar.lib.vt.edu/theses/NDLTD/SC970919.pdf>