



Digital Library Technologies (DLT) Annual Report, July 2003 - June 2004

CONTENTS

Information Access

Research and Development

Maintenance and Support

Infrastructure

Future Initiatives

Reorganization of DLT

Staff Development

Appendices

A. LIAS Statistics at a Glance

B. LIAS E-Resource Usage

This report describes the activities of Digital Library Technologies (DLT) during the past year. Significant accomplishments include a large increase in the number and quality of information resources available through LIAS, significant additions to the emerging Digital Library Infrastructure (Shibboleth, ContentDM, PURL) and the reorganization of DLT, including the integration of the Libraries' workstation technical staff. The past year also saw continued effort on DLT's part to enhance and customize the delivered Sirsi Unicorn library management software to meet the Libraries needs.

Information Access

Addition of Large Bibliographic Files to The CAT. The last year continued a major shift in the way cataloging data is added to *The CAT*, the online catalog that represents PSUL's library holdings. Several collections of bibliographic cataloging data were purchased from external sources and added en masse to *The CAT*. Some of these collections provide for the first time full access to materials in the University Libraries for which access was previously minimal or non-existent. Many of these files received special tuning such as provision of unique call numbers to make them even more accessible. Although bibliographic collections such as these have been added to The CAT for years, the number of records added increased greatly this fiscal year.

This means of enriching *The CAT* with pre-cataloged collections will continue. In FY 2003/2004, more than 160,000 titles were added, which for the first time outpaced the number of records cataloged individually.

Did You Know?

162,669 bibliographic records were loaded into The CAT in 2003/2004:

Cornell Digital Math Books Collection	511
ACLS History E-Book Project	737
Congressional Information Service: CIS Congressional Committee Hearings, 1833-1969	42,019
Congressional Information Service: CIS Executive Branch Documents, 1789-1909	14,708
Congressional Information Service: CIS Microfiche Library, 1970-1985	43,961
Early English Books Online, 1641-1700	44,360

New Databases/Resources. Fifty new databases and other electronic resources in a wide variety of subject areas were made available via LIAS during 2003/2004 bringing the total number of e-resources accessible via LIAS to 430+.

ContentDM. A software package, ContentDM, was purchased by the Library to provide access to Digital Images as part of the VIUS project. DLT worked with the Library to customize the software to meet the Libraries' needs. Along with customizing the pages and creating CGIs to support the change, a configuration program was designed and implemented to add in the creation of new collections. Two collections, Art, Architecture and Landscape Pictures and Pennsylvania History Pictorial Collection are accessible via this software. Three new releases were installed and successfully tested with the customized interface.

Research and Development

Shibboleth. Last fall, DLT started working with Emerging Technologies to investigate using Shibboleth to access Electronic Resources. A pilot project was set up to test JSTOR. ScienceDirect and OCLC will be added to the pilot project shortly.

PURL. As part of its Digital Library infrastructure, the Libraries have expressed a need for a Persistent URL resolver. DLT investigated the service and various implementation options and decided to implement the PURL server from OCLC. The software was installed and tested and is now in production.

Software Investigations. DLT is investigating several open source products to support the Library's digital collections needs.

- Zebra, a Z39.50 server, will be used to move local library bibliographic collections from the VMS environment to a UNIX environment.
- DSpace: a digital institutional repository
- Fedora: a digital institutional repository
- DLXS: foundation/framework for digital libraries
- Greenstone: foundation/framework for digital libraries

SSN Conversion Project. Tun Chin serves on the university-wide committee charged with eliminating Social Security numbers as the main University identifier. Since the database of registered library users tracks users by SSN, DLT has designed and implemented all the programs needed to convert the SSN to the Penn State ID. The first phase of testing has been completed. The Library's lending staff is working on cleaning up all of the problem SSNs in the Unicorn database. DLT is working closely with the Library's lending staff providing monthly reports of potential problems. It is anticipated that the conversion will occur over the weekend of December 18, 2004.

PDA access to WebCAT. DLT is working with a few students in an IST class to provide access to the CAT from a PDA.

Other Activities. A variety of other software enhancements not directly related to LIAS were completed during the year. For example, the joint Libraries/DLT Helpdesk was migrated from VMS to UNIX. The Electronic Theses and Dissertations program was upgraded to reflect changes in the Graduate School thesis guide and policies.

Maintenance and Support

Sirsi Unicorn Management and Maintenance. The past year a significant amount of time was dedicated to the Unicorn 2003 upgrade. After installing and regression testing three releases, DLT and the Library determined the third release suitable for release. As part of the upgrade, DLT wrote scripts that enable the Library to use a new Lost Item Processing feature of Unicorn. DLT continued to enhance the Unicorn product by creating reports to transfer bibliographic data to RLG, OCL, ACCESSPA and provide access to Circulation Reports on the Web. Sirsi and staff members from DLT and the Libraries worked together to implement and test EDI for Serial Invoicing.

Payroll Deduction and Debt Collection. DLT worked with the Libraries to complete both the payroll deduction and debt Collection project. Both services are now in production.

Cataloging of Bibliographic Data. Late last fall, the Libraries requested that DLT investigate a new tool used for the cataloging of Bibliographic data. The tool, Connexion, is a replacement for SmartPort allowing the library staff to download multiple records to a file that would then be loaded into the Unicorn database. Custom reports were created to load Marcive, PromptCat, Blackwell, and Gobi records.

LIAS on the Web/Libraries Web Site. The merger of www.lias.psu.edu and www.libraries.psu.edu was completed. These two sites are now running as one site using the apache web server.

Outreach. DLT publicized the many resources available through LIAS at the ITS Event on March 16, 2004. Three areas in which DLT participated were: Support for Research, Bring Your Courses Alive, and Penn State from anywhere. Two DLT staff were heavily involved in the entire planning and preparation process, while 14 others participated in set-up, staffing the tables the day of the show, and in tear-down.

Infrastructure

New Hardware and Upgrades. DLT performed the following during FY2003/2004:

- Added 3 additional SUN servers to enhance library applications, running the Solaris operating system
 - 2 Production SUN V880 servers with 32GB of Memory each
 - 1 Development SUN V240 server with 8GB of Memory
 - Added connectivity hardware from SUN server to HP Enterprise Virtual Array (EVA)
 - Added 8 processor unlimited user Oracle license for one Sun Solaris server
 - Added 100 user Oracle license to be distributed between the other two Sun servers
- Added 3 additional HP Intel based servers running Windows 2000
- Upgraded HP tape library (robot)
- Upgraded 225 LIAS public PC's

Wireless. DLT has begun implementation of a wireless network in the University Park Libraries. Using Foundry Networks wireless products, the initial implementation, scheduled for the beginning of Fall 2004, will include all public areas of Pattee Library and Paterno Library as well as the Pollock Library and the ITS student lab on the 2nd floor of that Library.

Workstations for the Visually Impaired. DLT has configured and installed workstations in all Library locations at University Park as well as all campus libraries that are specifically for the use of visually impaired students. These workstations are equipped with 21" high resolution monitors as well as software that will magnify images to allow students to use Library resources that would not be possible with our standard workstations and monitors.

Space Reorganization. DLT, working in cooperation with ASET, reorganized space in the DLT computer room to allow hardware to be installed that provided ASET with redundancy and off site disaster recovery for some of their critical systems. In addition, space was also reorganized

to allow the addition of 5 temporary cubicle offices to accommodate the ITS Marketing and Communications Department while other ITS space is being renovated.

Help Desk. 4,165 trouble calls were reported to the joint DLT and University Libraries Help Desk from July 2003 through June 2004. Since the reorganization, all trouble calls are handled by DLT.

Future Initiatives

Journal Lists, Context-Sensitive Citation Linking and Federated Search. The Penn State University Libraries provides a vast array of electronic resources for their users through LIAS. The resources include the Sirsi Unicorn integrated library management system, major licensed indexes, databases and online journals, and a number of important locally produced digital collections. Many of the databases and online journals are delivered through aggregators which supply citations and full text of many electronic journals. Other electronic journals are accessed directly from the publishers' web sites. Researchers want quick access to specific electronic journals regardless of how the Libraries has contracted for that journal. Access using a variety of methods (through the CAT, via an A-Z list, keyword and subject searching) are needed to meet student and other researcher needs.

A strategic goal for the University Libraries is to "Provide quality services that ensure timely and effective access to local and remote information resources in all formats". One important objective toward this goal is to "Encourage innovation in the delivery of library services to meet the needs of onsite and remote users in an increasingly technological environment". An important strategy in meeting that object is to "Strive for virtually integrated services that allow users to search across and use geographically disparate digital collections and Web resources and services as if they made up a single online suite of collections and services".

In addition, it is very important to have the capability to allow users to extend to additional relevant resources through URL links (context-sensitive linking). Both broadcast searching and citation linking are critical applications that will make users and librarians more aware of the wide range of resources that are available to them to assist with their research and information needs.

During 2003/2004, the University Libraries and DLT developed an RFP soliciting proposals for solutions to these problems (multi-access journal lists, context-sensitive linking, and federated search). In August 2004, a contract was awarded to Ex Libris for acquisition of SFX for journal management and linking and MetaLib for broadcast searching. Implementation of SFX will begin in the Fall of 2004, with MetaLib implementation to follow in early 2005.

OAI. Penn State, along with several other CIC Libraries, is creating and implementing an OAI-PMH metadata harvesting service to aggregate metadata describing information resources held by participating CIC Libraries. The intent of this initiative is to make this metadata aggregation available to end-users (students, faculty, and the general public), both within and outside of the CIC, using appropriate, state-of-the-art search and discovery tools and browsing and navigation interfaces. This past year has been spent in identifying resources at the participating schools and beginning to build the aggregated database of metadata pointing to those resources. DLT has been enabling resources to be harvested by this project, specifically enabling the harvesting

of Electronic Thesis and Dissertations metadata and ContentDM metadata. Over the next couple of years, we will identify additional resources enable those resources for harvesting. DLT and the Libraries will also work to make the aggregated metadata available through LIAS.

Network Firewall. DLT has purchased a network firewall that will be used to protect systems and applications that do not require open public access and internal systems that are deemed mission critical where restrictions will not affect function.

UNIX Server Hardware. DLT is investigating the addition of Enterprise level servers that will likely be purchased from SUN Microsystems due to their wide acceptance in the library application arena.

Diversified Operating Systems. DLT will continue to investigate other operating system platforms in an attempt to lessen the dependency on the Windows Server and Workstation platform. Platform examples are: Novell Netware, SUN Rays, Macintosh and Solaris X86.

Reorganization of Digital Library Technologies

Digital Library Technologies split into Two Divisions. Two years ago, Computer and Information Systems reorganized and restructured to become Information Technology Services. As part of that restructuring, Library Computer Services became Digital Library Technologies, but because of the recent implementation of Sirsi Unicorn and other complications, reorganization of Digital Library Technologies was deferred.

On September 1, Digital Library Technologies was reorganized into two divisions: Digital Library Research and Development, and Digital Library Infrastructure, recognizing the changing needs of the University Libraries and the faculty and students who use digital resources. Having completed the implementation of Unicorn, the focus has evolved to include new platforms and infrastructure, recognizing both the research into and development of digital library solutions along with the need to support an increasingly larger, more critical and complex infrastructure for automated library services.

Digital Library Research and Development is responsible for research into and analysis, development and integration of digital library software components and software systems. This unit works with the University Libraries and other Information Technology Services groups to design, develop, and customize solutions to library information problems and services.

Digital Library Infrastructure is responsible for the servers, network, workstations, and systems software that underlie the digital library systems and services supported by Digital Library Technologies and the operations of those systems, with reliability and availability an ever more important element.

Each division is headed by a Director who will be responsible for all of the planning, staffing, and day to day operations of their division. Lynn Garrison assumed the position of Director of Digital Library Research and Development on September 1. All of the development staff and the system analysts reports to Lynn. Mark Saussure accepted the position Director of Digital Library Infrastructure on October 1. The systems engineers, network coordinators, and workstation

technical staff all report to Mark.

Realignment of Responsibilities between DLT and the University Libraries. On July 1, 2004, four workstation technicians and a help desk coordinator were moved from the Information Technology division within the University Libraries into DLT. Prior to this move, Digital Library Technologies supported and maintained the public LIAS workstations located in the University Libraries, but the Information Technology division supported the workstations used by the librarians in the University Libraries. With this move, all of the workstations in the Libraries are now supported and maintained by DLT. The helpdesks have been integrated into a single point of contact located in DLT. This will result in better customer service, better use of technical personnel, and a more consistent approach to workstation management.

As part of the move, two librarians who had served as systems analysts and had supported much of the day to day operations of Sirsi have been moved to the Information Technology division in the University Libraries. This division will handle most of the day to day maintenance of Sirsi and other library applications, while DLT will focus on maintaining the servers, network, workstations, middleware, and daily operations in addition to research and continuing development of digital library systems.

In late June, the Dean of Libraries and the Senior Director of Digital Library Technologies agreed to a Memorandum of Understanding outlining the responsibilities of each unit. This will serve as the basis for defining our individual roles as the number of library applications and systems continues to increase along with the demands for more digital services from librarians, faculty, and students.

Staff Development

During the past year, DLT staff participated in a total of 2,819 hours of professional development activities. The activities ranged from self-paced reading, completion of web tutorials, and hands-on exploration of new software to more formal activities such as ITS' Teaching and Learning with Technology seminars, HRDC and Mastering Supervision classes, and even some week long workshops at locations other than PSU. Topics that DLT staff studied during the year included: XML, SQL, Perl, SSH, Zebra, CIW Security, Oracle, Dreamweaver and more. DLT staff also attended and participated in a number of University Libraries Technology Forums and presentations on new software systems and services.

DLT Professional Development Activity

- 2,819 total hours
- 26 Conferences/Meetings
- HRDC Training - 10 classes
- Web-based Training - 16 courses
- Courses: XML, SQL, Perl, JDBC, awk, Unix, VRML, Windows 2002 Security, Dreamweaver, Photoshop, Adobe Acrobat, LIFT from Usablenet, and others

Conferences, Meetings, Exhibits

Kurt Baker

Perl on the Web and use of Perl Modules
Oracle training manuals
Wireless Seminar
Apache Personal Web Server
ADA Web Class
XML Class
WEB 2004 Conference, University Park, 6/15/04

Joni Barnoff

JavaScript Class
CONTENTdm (self-taught)
Borland C++ Builder (self-taught)
JavaScript (self-taught)
Java loaded and created a small application (self-taught)
SSH (self-taught)
Zebra (self-taught)
XML Class
WEB 2004 Conference, University Park, 6/15/04

Paul Baughman

WBT CIW Security Professional, Part 5
Implementing and Supporting Windows XP Pro, South Hills Business School, 4/12-21/04
XML Class

Steven Baylis

Various computer courses at Penn State to earn a degree in IST

Mike Bender

I-Tech Trade Show, Valley Forge, PA 11/13-14/03
Mid Atlantic, Educause, Baltimore, MD 1/13-14/04
SUN Microsystems Training, Denver, Co 5/9-16/04

Tun Chin

ETD 2004 Conference, Lexington, KY 6/3-5/04
Oracle 9i SQL course
Oracle Database Administration Fundamentals I class
Oracle Database Administration Fundamentals II class, Pittsburgh, 3/28-4/2/04
SIRSI SuperConference 2004, St. Louis, MO 4/17-21/04
Oracle class: XML fundamentals for developers, Reston, VA 6/27-30/04

Todd Douty

XML Class

Debra Ellenberger

Searching the Web Seminar
Excel 1 class
Excel 2 class
E-mail@PSU: Fighting Spam with Eudora Filters (Windows)

Business Etiquette: Communicating in Today's Workplace
Business Etiquette: Office Protocol
Business Writing: The Fundamentals
Call Center: Inbound Customer Service
MS Excel 2002 Fundamentals
MS Excel 2002 Level 1 Competencies
MS Excel 2002 Level 2 Competencies
MS Excel 2002 Level 3 Competencies
MS Excel 2002 OPC Task 1 Skill IDs 1.1-1.4
Budgeting: Creating and Analyzing a Budget

Eric Ferrin

Seminar on Academic Computing, Snowmass, CO 8/4-6/03
HP Enterprise Technical Symposium, Atlanta, GA 8/9-15/03
CIC LITD Meeting, Chicago, IL 10/20-21/03
CIC Tech Forum, Madison, WI 10/29-11/1/03
Educause 2003, Anaheim, CA 11/4-7/03
Fall 2003 DLF Forum, Albuquerque, NM 11/17-19/03
RLG Digital Repositories Forum, Chicago, IL 12/11-12/03
CNI Fall 2003 Task Force, Portland, OR 12/8-10/03
Net@EDU, Tempe, AZ 2/1-3/04
CNI Spring 2004 Task Force, Alexandria, VA 4/13-16/04
DLF Spring Forum 2004, New Orleans, LA 4/18-21/04
Benchmarking with University of Washington, Seattle, WA 5/10-13/04
CIC Library Information technology Directors Meeting, Chicago, IL 5/6-7/04
JCDL 2004, Tucson, AZ 6/6-11/04

Koren Foster

Introduction to Excel
Illustrator Basics
FIT Class
Managing Your Matrix
Dreamweaver Basics
Making the most of non-verbal communications

Dace Freivalds

SIRSI SuperConference 2004, St. Louis, MO 4/17-21/04
SIRSI, Bucknell Univ., Lewisburg, PA 6/13-15/04
ALA, Orlando, FL 6/24-29/04

Lynn Garrison

Performance Management: Building for the Future
Apache (self paced)
Educause 2003, Anaheim, CA 11/4-7/03
IOUG LIVE! 2004, Toronto, Canada 4/17-22/04
CAMP Shibboleth, Broomfield, CO 6/26-30/04
XML Class
ALA Mid Winter Meeting, San Diego, CA 1/8-13/04
WEB 2004 Conference, University Park, 6/15/04

Maryam Kutchemeshgi

XML Class

Sylvia MacKinnon

ALA Mid Winter Meeting, San Diego, CA 1/8-13/04
ALA, Orlando, FL 6/24-29/04
SIRSI SuperConference 2004, St. Louis, MO 4/17-21/04

Janis Mathewson

ETD 2004 Conference, Lexington, KY 6/3-5/04
Oracle 9i SQL course
Oracle Database Administration Fundamentals I class
XML Class
WEB 2004 Conference, University Park, 6/15/04

Jack Orlandi

LITA Forum, Norfolk, VA 10/1-5/03
2004 DSpace Federation User Group Meeting, Boston, MA 3/9-12/04
Mastering SuperVision
XML Class
WEB 2004 Conference, University Park, 6/15/04

Mark Saussure

SQL Server Administration Class
IBM Briefing, Palisades, NY 12/11-12/03
SUN Microsystems Training, San Jose, CA 5/16-22/04

Jeff Shawver

Implementing and Supporting Windows XP Pro, South Hills Business School, 4/12-21/04

Wayne Stump

Implementing and Supporting Windows XP Pro, South Hills Business School, 4/12-21/04
XML Class

J. Lance Wilkinson

Novell TTP (Technology Transfer Partners), Salt Lake City, UT 7/11-20/03
SUN Microsystems Training, San Jose, CA 6/6-12/04
XML Class

Presentations**Tun Chin**

XML, XSLT
Class for DLT

Eric Ferrin

Content Management Systems
Technology Update, Penn State, 9/10/03
LionShare
Net@EDU, Tempe, AZ 2/1-3/04
Peer to Peer Task Force, Washington, DC, 2/18/04
CM, DAM, DRM, KM, DM, DMM: What does all of this alphabet soup mean?

Web 2004 Pre-Conference Tutorial, Penn State, 6/14/04
A Digital Library Tool for Real People
Web 2004 Pre-Conference Tutorial, Penn State, 6/14/04

Dace Freivalds, Tun Chin

Tired of Begging Your SysAdmin? Tools for Doing It Yourself
SIRSI SuperConference 2004, St. Louis, MO 4/17-21/04

Sylvia MacKinnon

GOBI2 interface to Unicorn
SIRSI SuperConference 2004, St. Louis, MO 4/17-21/04

Jeff Shawver

Networking and Backup
Presentation to Workstation Liaisons

Lance Wilkinson

PreciseMail Anti-Spam Gateway
Technology Update

Appendix A. LIAS Statistics at a Glance

Did You Know That ...	
	<ul style="list-style-type: none"> • 50 new databases and resources were made available through LIAS during 2003/2004.
	<ul style="list-style-type: none"> • 87% of PSU faculty and staff had accepted payroll deduction of library fines and fees by the end of the fiscal year.
	<ul style="list-style-type: none"> • 476 bills were paid through payroll deduction totally more than \$6600.
	<ul style="list-style-type: none"> • 4,165 Helpdesk tickets were reported and resolved.
	<ul style="list-style-type: none"> • The CAT accounted for 42% of LIAS e-resource usage during 2003/2004.
	<ul style="list-style-type: none"> • 83% of all users receive their library notices via email.
	<ul style="list-style-type: none"> • 162,669 bibliographic records were loaded into The CAT in 2003/2004.
2003/2004 in Numbers...	
Databases and Other E-Resources	
Number Available	430+ (13% increase)
Added in 2003/2004	50
Total Number of Sessions	3,272,611 (7.1% increase)
The CAT	
Bibliographic Records in Database	2,680,190 (10.4% increase)
Bibliographic Loaded in 2003/2004	162,669
Total Number of Sessions	1,381,556 (2.3% increase)
Circulation	
Items in Database	4,900,362
Users in Database	129,406
Charge Transactions (Charges and Renewals)	651,040
Hardware	
LIAS Express Workstations in the Libraries	80
Authenticated Public Workstations in the Libraries	670
Staff Workstations in the University Libraries	800
Disk Storage	22 TB
Servers	18 Windows, 6 Unix, 1 VMS

Appendix B. LIAS E-Resource Usage Statistics

LIAS Database and Other E-Resource Usage Statistics, 2000 - 2003				
	2000	2001*	2002	2003
Totals				
Total No. Of Sessions Across All Resources	2,299,067	2,542,038	3,055,648	3,272,611
The CAT				
CAT - PSU Libraries Web Catalog	1,087,740	1,103,812	1,350,682	1,381,566
Top Twenty (based on 2003 usage)				
<i>(Statistics for all resources are available at http://www.lias.psu.edu/stats/contents.htm)</i>				
CAT - PSU Libraries Web Catalog	892,588	1,016,243	1,350,682	1,381,556
ProQuest Direct (newspapers, magazines, business)	140,307	214,786	285,101	320,069
LEXIS-NEXIS Academic Universe	39,948	45,081	90,046	86,097
PsycINFO (psychology)	65,726	52,859	70,901	78,479
Web of Science	34,325	46,328	70,246	69,968
JSTOR (full text journal articles)	31,116	42,724	56,830	61,413
WorldCat	11,728	19,582	38,401	50,243
Electronic Reserves	28,129	36,310	14,162	49,151
ProQuest Psychology Journals	--	--	--	48,650
Elsevier ScienceDirect Web Editions (fulltext journal articles)	6,761	32,836	48,234	42,825
ABI/INFORM (business)	24,947	24,473	33,220	38,277
Sociological Abstracts	8,785	18,430	25,826	36,317
MEDLINE (PubMed)	41,461	43,013	53,631	28,351
Dow Jones Interactive (now Factiva)	19,337	22,876	29,621	26,793
IEEE Electronic Library (IEEEXplore)	17,487	26,682	27,902	25,285
CQ Researcher (Contemporary Issues)	--	--	12,603	23,979
Education Abstracts Full Text [Wilson]	3,992	--	13,623	22,785
Elsevier Science Direct (full text journal articles)	--	--	16,705	21,135
Compendex (Engineering)	15,408	11,125	16,047	20,936
Blackwell Science/Blackwell Publishing Journals	--	--	18,817	20,690
Note:				
<i>*Data for October - December 2001 were not available. Annual statistics for 2001 were calculated using the data from the first nine months.</i>				