



## Digital Library Technologies (DLT) Annual Report, July 2002 - June 2003

### CONTENTS

Information Access

Maintenance and Support

Future Initiatives

Staff Development

Appendices

- A. LIAS Statistics at a Glance
- B. LIAS E-Resource Usage
- C. LIAS System Availability

*This report describes the activities of Digital Library Technologies (DLT) during the past year. Significant accomplishments include collaboration with the Libraries to make e-books available through LIAS, implementation of payroll deduction of outstanding library fees, providing access to 100 new electronic resources acquired by the Libraries available through the CAT, and implementation of the first phase of a data warehouse. 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

**E-Books.** One of the most significant changes to the Libraries e-resources during the past year was the addition of a number of full text e-book collections. In some cases, these e-books are also available in the University Libraries in paper form; in other cases, the e-book is the sole means of access to the work. E-books are available from vendors through licensing agreements, and users go through authentication to “borrow” and read the materials online. DLT makes the e-book collections accessible via the LIAS E-Resource list, ensuring compliance with license agreements. In addition, DLT loads records for the individual titles in the collections into The CAT providing a direct link to the full text of the specific title.

Among the e-book collections acquired by the Libraries is *Safari Tech Books Online* which provides searchable, full text access to over 600 technology books from the major information technology publishers such as O-Reilly, SAM, Que and others. Other e-book collections include *NetLibrary*, a collection of 4,500+ books covering a large variety of subjects, *Wright American Fiction* which contains 2,839 full texts of American novels first

published between 1851 and 1875, *Early English Books Online*, and *Evans Digital Edition*. The individual title records for the *NetLibrary* and *Wright American Fiction* collections were loaded into The CAT during FY 2002/2003; the Safari collection will be loaded in early Fall 2003.

**Accessibility Enhancements.** All public LIAS workstations in the University Libraries were upgraded with many of the standard Microsoft Windows accessibility features such as the ability to use large print, reverse negative images, etc.

**Authentication at Public LIAS Workstations.** In order to conform to Penn State security policies, DLT collaborated with the University Libraries in moving to an authenticated environment for its public workstations. In January, 2003, DLT converted the majority (670) of the public workstations in the University Libraries to authenticated stations. After authenticating with their PSU Access Account IDs, Penn State students, faculty and staff continue to have access to unrestricted Internet resources at these workstations. For other Libraries users, 80 specialized LIAS Express workstations permit access to The CAT, LIAS databases (as permitted by license agreements), and selected web sites designated as .gov, .mil, .state.pa.us and .psu.edu. Under special circumstances, one-day “Research Accounts for the Day” which grant access to the restricted resources can be issued to non-Penn Staters via a web-based management tool created by DLT.



Figure 1. Page from Safari e-book

**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*; this represents a departure from the standard means of adding titles one at a time by cataloging staff. 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.

Did You Know?	
54,582 bibliographic records were loaded into The CAT in 2002/2003:	
Cornell Digital Library Federation	288
ACLS History E-Book Project	474
History of Photography	2,198
Wright American Fiction	2,839
German Baroque Literature	3,348
Three Centuries of English and American Plays	4,106
NetLibrary	4,728
Early American Imprints, Series 2	36,601

This means of enriching *The CAT* with pre-cataloged collections will continue. Several other such projects initiated in FY 2002/2003 will go into production shortly, adding more than 100,000 titles which will for the first time outpace the number of records cataloged individually each year.

**Access to Course Reserve Materials via ANGEL.** DLT modified and adapted OCLC's WebScript program to automatically log the user on to The CAT and then issue a search for the reserve associated with the class the student is actively working on in ANGEL course management system. The search is executed using the class number and the instructor's id. On successful completion of the search, The CAT "Course Reserve" page is displayed.

**New Databases/Resources.** One hundred new databases and other electronic resources in a wide variety of subject areas were made available via LIAS during 2002/2003 bringing the total number of e-resources accessible via LIAS to 380+.

**Usage Statistics.** In order to make compilation and retrieval of e-resource usage statistics more efficient, data were migrated to an Oracle database. The statistics are now available on the web in html or as a csv file, and can be broken down by e-resource and library. E-resource usage statistics for 2002 are summarized in *Appendix B. LIAS E-Resource Usage*; complete database statistics are available at <http://www.dlt.its.psu.edu/liasstats.html>.

**Proxy Services.** DLT enhanced its proxy services in order to ensure compliance with license agreements from the database and other e-resource vendors. DLT now provides a proxy server for full access to resources including those licensed for Hershey and another proxy server for those resources licensed for all but Hershey. As part of the enhancement, specific Hershey

e-resource lists were created, ensuring that users are presented only with those resources that they are allowed to access.

**Subject Access.** Due to the huge number of available e-resources, it has become increasingly difficult for users to identify those which may be most appropriate for them. In order to alleviate some of the problems with this, DLT and the Libraries collaborated to produce the ability to search the e-resources by broad subject areas. A list of most useful resources, referred to as “try these first”, was also generated.

**Single LIAS / Libraries Web Site.** For the past ten years, DLT and University Libraries staff maintained two web sites, one for LIAS, the other for other library services and resources. A joint DLT/Libraries SWAT team redesigned and integrated the two sites to provide one entrance point to all library resources for users. The new site was implemented in August 2002. Since the Libraries are one of the 14 “major” links on the University home page, the new design conforms to the University template for upper level homepages. As part of the project, DLT updated all of the e-resource pages to incorporate the Libraries templates for secondary pages. The SWAT team continued to review the pages, and make necessary adjustments during the course of the past year.

**Authentication.** The LIAS authentication process was updated to use MIT Kerberos 5 instead of DCE to authenticate PSU Access IDs.

**Digital Initiatives.** DLT has made a commitment to provide the digital library infrastructure as well as technical support for digitization projects undertaken by the University Libraries projects, and to work with Libraries faculty on adding services and content to the Penn State Digital Library. Included in the infrastructure is allocation of server space for the digital projects and the creation of a Persistent URL (PURL) resolver. Maintenance of URLs, which have a habit of changing frequently, is a very intensive, and time-consuming effort. To make this more efficient, DLT installed a PURL resolver which can be used for maintaining a URL database. Any URL can be stored into the database under a persistent “reference name”, and then accessed by this name. The name never changes even though the URL may change, thus guaranteeing access to the data by the “reference name”. One application of the PURL resolver during the past year was in the University Libraries/PSU Press collaborative “Time of Sorrow and Hope” digitization project.

DLT also continued to participate in the Visual Image User Study (VIUS) project by maintaining and updating the Content DM software. VIUS is a collaboration of the University Libraries, DLT, the Center for Education Technology Services, the Center for Quality and Planning, and the School of Information Sciences and Technology.

**Outreach.** DLT publicized the many resources available through LIAS at ITS’ first IT Day on October 16, 2002. Two topics were presented: *Gateway to the Information Universe* and *Today’s Library is Virtual*. Two DLT staff were heavily involved in the entire planning and preparation process, while six others participated in set-up, staffing the tables the day of the show, and in tear-down.

## Maintenance and Support

---

**Sirsi Unicorn Management and Maintenance.** By July 2002, the University Libraries and DLT had been using Sirsi's library management system Unicorn for a year. Although the initial shakedown of the new system took place the previous year, the past year saw continued effort on DLT's part to enhance and customize the delivered product to meet the Libraries needs. To that end, DLT created a number of customized reports, converted delivered reports to web-based reports, moved web reports to a secure server, and in general, made tweaks to make the system more efficient. For example, DLT staff wrote a serials control rollover script, modified and adapted a distribution and claim segments rollover script prepared at the University of Virginia, modified patron master load code, etc.

To ensure better management of Sirsi initiatives at the University Libraries, a joint Libraries/DLT Sirsi Steering Committee (SSC) was appointed in late 2002. The committee makes decisions regarding Sirsi policies, directs the activities of five Sirsi expert teams representing functional areas of the Libraries, and is responsible for the coordination of scheduling, testing and implementation of new releases. One DLT staff member serves on each expert team, and two DLT staff members serve on the steering committee.

**Payroll Deduction of Outstanding Library Fees.** On June 1, 2002, the University Libraries adopted a new plan for automatic payroll deduction of outstanding library fees for all Penn State faculty and staff (excluding those at Hershey or Dickinson). By July 31, 2002, 50% of faculty and staff had accepted the new policy and agreed to payroll deductions for payment of outstanding library fees. This policy was suspended pending further discussion in the faculty Senate, however, and DLT put its work on this project on hold. Upon re-approval of the new policy at the February 25, 2003 Faculty Senate meeting, DLT resumed working on code to forward outstanding bills to the Penn State Payroll Office. By the end of FY 2002/2003, 81% of faculty and staff had accepted the new policy. It is anticipated that the first batch of outstanding fees will be forwarded to the payroll office on September 15, 2003.

**Data Warehouse.** The data warehouse is in the initial stages of development. DLT is responsible for populating the database, while Libraries staff will be responsible for creating reports and other mechanisms to retrieve the data. A base load of items, limited bibliographic data, call numbers, and command transactions was completed by the end of the reporting period.

**Migration to POPStore.** For years we had been using a POP server that would serve out normal files to our Eudora users. But scanning a VMS mail file, which was never designed as anything other than a local mainframe mail store, and reformatting new messages for download to a POP client was a slow, inefficient procedure. By formatting incoming mail as it arrives for immediate download to a POP client the PMDF POPStore product offered an opportunity to speed up the download process and improve service to our users. Other efficiency aspects include the ability of the POPStore to store one copy of a message directed to literally hundreds of recipients, saving disk storage. The migration process also gave DLT the opportunity to synchronize our local email addresses in the new POPStore with Access IDs, carrying us closer to the day when single sign-on will be a rule rather than a goal.

**Debt Collection.** In mid-2003, the University Libraries contracted with Unique Management to collect outstanding library fees from registered library users who are not affiliated with the University, i.e., are not students, faculty or staff. DLT worked with the Libraries to ensure that

debt data would be reported to the collection via secure ftp.

**Automatic Creation of Order Records.** Since 1999, DLT and the University Libraries have streamlined the cataloging process by developing the means to load cataloging copy directly into The CAT for many materials, including titles on the library's approval plan eliminating the need for staff to manually enter the data. DLT staff wrote code to take this process one step further. The new process generates not only bibliographic records, but also the corresponding order record for each item in the file. During Fall 2002 the new process was implemented for the generation of orders from Academic's Bookbag product.

**Pinyin Conversion.** Chinese language records in The CAT were converted from the Wade-Giles to the Pinyin transliteration scheme in Fall 2002. The Pinyin romanization of Chinese makes it easier for users to locate Chinese language materials. To make this conversion possible, DLT extracted and sent copies of LIAS records to OCLC Online Computer Library Center, Inc. for processing. After conversion, the changed records were reloaded back into The CAT. The entire process was done without any interruption to cataloging or patron access.

**Hardware Upgrades.** DLT performed the following during FY20032/2003:

- Upgraded the network hardware to equipment from Foundry Networks
- Upgraded the virtual array storage with 7 TB of storage
- Upgraded hardware on all 15 Windows 2000 servers

**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 upgraded, and DLT staff participated in an evaluation of commercial helpdesk software. The Electronic Theses and Dissertations program was upgraded to reflect changes in the Graduate School thesis guide and policies. DLT staff also modified the LDAP server for Access ID and password authentication, and added it to the front end of the Libraries new Interlibrary Loan ILLiad software. DLT upgraded the Libraries TechSmart training registration software code, and also moved it from Linux to Unix.

**Help Desk.** 3,990 trouble calls were reported to the joint DLT and University Libraries Help Desk from July 2002 through June 2003. Of these, 1,746 involved DLT hardware and software, and were resolved by DLT staff.

## Future Initiatives

---

**Web-based Sirsi Unicorn Tools.** To provide better access to and retrieval of data in Unicorn, DLT started writing a Perl program to provide a web interface to Sirsi's selection and print "tools". The web version of these tools will enable Libraries staff to generate delimited files for import into databases, spreadsheets, etc. This is not possible within Unicorn as it is delivered. Preliminary work on this was completed by the end of the reporting period with implementation planned for Fall 2003. In addition, DLT is looking at ways to update Sirsi Unicorn policy tables via the web so that policy updating can be delegated to individuals in the Libraries.

**GOBI 2.** Specifications for the loading of orders from Yankee Book Peddler GOBI 2 (Global Online Bibliographic Information Edition 2) interface were prepared during the past year. This process differs from the Bookbag process implemented in Fall 2002 in that Bookbag orders

are placed directly with the vendor while GOBI 2 orders are selected via the GOBI 2 interface, but orders are placed directly through Unicorn. Implementation of the GOBI 2 process will continue into the next year.

**SSN Conversion Project.** DLT Director 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 will need to develop mechanisms for replacing the current User IDs with the new, randomly generated nine-digit identifiers.

**Enhanced Public Workstations.** During the coming year, DLT will convert approximately 200 of the 750 public LIAS workstations in the University Libraries to “enhanced public workstations” which will include the Microsoft Office Suite (Word, Excel, etc.).

**EndNote and The CAT.** At the present time it is possible to connect to and retrieve citations from The CAT from within the EndNote software. Or, it is possible to go into The CAT, and save citations for importing into EndNote. DLT and the University Libraries are looking at a way to streamline this process so that users can simply click on a link in The CAT and automatically move marked citations to EndNote.

**Requesting MTSS Materials.** Media and Technology Support Services (MTSS) is working with DLT staff to develop a mechanism for requesting audio-visual materials directly from The CAT. Since MTSS does not use the Unicorn circulation component, and requests can come from users not registered with the Libraries, requesting these materials via “I Want It” is not an option. Instead, DLT will develop a link at the record level that will take users to the MTSS request form. MTSS and DLT hope to have this feature in place early in Fall 2003.

## Staff Development

---

During the past year, DLT staff participated in a total of 2,390 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 Penn State Leader and Mastering Supervision classes, and even some week long workshops at locations other than PSU. Topics that DLT staff studied during the year were: XML, SQL, Perl, JDBC, awk, Unix, VRML for Programmers, CIW Security, Windows 2000 Security, LIFT from Usablenet, Dreamweaver, Photoshop, Adobe Acrobat and others. DLT staff also attended and participated in a number of University Libraries Technology Forums, presentations on new software or services, etc.

### DLT Professional Development Activity

- 2,390 total hours
- 23 Conferences/Meetings
- Mastering Supervision - 3 staff
- Penn State Leader - 2 staff
- Courses: XML, SQL, Perl, JDBC, awk, Unix, VRML, Windows 2002 Security, Dreamweaver, Photoshop, Adobe Acrobat, LIFT from Usablenet, and others

## *Conferences, Meetings, Exhibits*

### **Mike Bender**

Comnet, Washington, DC, 1/28-31/03  
HP Storage Days, Scottsdale, AZ, 11/16-21/02

### **Tun Chin**

CIC OAI Metadata Harvesting Serv. Conf., Chicago, IL, 10/7/02  
Web Development with XML: Design & Applications, Tucson, AZ, 1/20-24/03  
SIRSI Super Conference, St. Louis, MO, 3/16-19/03  
DLF Forum, New York City, NY, 5/14-16/03

### **Lynn Garrison**

Educause, Atlanta, GA, 10/1-4/02  
IOUG LIVE! 2003 (International Oracle Users Group), Orlando, FL, 4/27-5/2/03  
WEB 2003 Conference, 5/17-18/03

### **Eric Ferrin**

Educause, Snowmass, Denver, CO, 8/3-8/02  
Educause, Atlanta, GA, 10/1-4/02  
HP Enterprise Tech. Symposium (ETS), St. Louis, MO, 10/7-11/02  
CIC Library Info. Tech. Directors, Chicago, IL, 10/21-22/02  
DLF Forum, Seattle, WA, 11/2-6/02  
HP Storage Days, Scottsdale, AZ, 11/16-21/02  
IT Directors/CNI, San Antonio, TX, 12/3-8/02  
CSG Meeting, Austin, TX, 1/15-16/03  
SIRSI Super Conference, St. Louis, MO, 3/16-19/03  
Brainshare, Salt Lake City, UT, 4/12-19/03  
CNI, Washington, DC, 4/28-5/1/03  
CUMREC 2003, Orlando, FL, 5/10-16/03  
JCDL'03, 3<sup>rd</sup> Annual Joint Conference on Digital Libraries, Houston, TX, 5/28-21/03  
SUN Microsystems, Menlo Park, CA, 6/4-5/03

### **Dace Freivalds**

ALA (American Library Association) Midwinter Meeting, Philadelphia, PA, 1/24-25/03  
SIRSI Super Conference, St. Louis, MO, 3/16-19/03  
WEB 2003 Conference, 5/17-18/03

### **Sylvia MacKinnon**

ALA (American Library Association) Midwinter Meeting, Philadelphia, PA, 1/24-27/03  
WEB 2003 Conference, 5/17-18/03

### **Janis Mathewson**

WEB 2003 Conference, 5/17-18/03

### **Jack Orlandi**

Educause, Atlanta, GA, 10/1-4/02  
Web Development with XML: Design & Applications, Tucson, AZ, 1/20-24/03  
WEB 2003 Conference, 5/17-18/03

### **J. Lance Wilkinson**

Novell TTP (Technology Transfer Partners), Cambridge, MA, 7/6-13/02  
Brainshare, Salt Lake City, UT, 4/12-19/03

## Appendix A. LIAS Statistics at a Glance

---

### Did You Know That ...

- 7,569 full text e-books were available in The CAT (University Libraries online catalog) by the end of June, 2003.
- 81% of PSU faculty and staff had accepted payroll deduction of library fines and fees by the end of the fiscal year.
- 100 new databases and resources were made available through LIAS during 2002/2003.
- 1746 Libraries Helpdesk tickets were resolved by DLT staff.
- The CAT accounted for 44% of LIAS e-resource usage during 2002/2003.
- 73% of all users receive their library notices via email.
- 54,582 bibliographic records were loaded into The CAT in 2002/2003.

### 2002/2003 in Numbers...

#### Databases and Other E-Resources

Number Available	380+
Added in 2002/2003	100
Total Number of Sessions	3,055,648

#### The CAT

Bibliographic Records in Database	2,426,013
Bibliographic Loaded in 2002/2003	54,582
Total Number of Sessions	1,350,682

#### Circulation

Items in Database	4,403,904
Users in Database	131,649
Charge Transactions (Charges and Renewals)	847,570

#### Hardware

LIAS Express Workstations in the Libraries	80
Authenticated Public Workstations in the Libraries	670
Staff Workstations in the University Libraries	800
Disk Storage	20 TB
Servers	15 Windows, 3 Unix, 1 VMS

## Appendix B. LIAS E-Resource Usage Statistics

<b>LIAS Database and Other E-Resource Usage Statistics, 1999 - 2002*</b>				
	<b>1999</b>	<b>2000</b>	<b>2001</b>	<b>2002</b>
<b>Totals</b>				
Total No. Of Sessions Across All Resources	2,080,468	2,299,067	2,542,038	3,055,648
<b>The CAT</b>				
CAT - PSU Libraries Web Catalog	1,039,199	1,087,740	1,103,812	1,350,682
<b>Top Twenty (based on 2002 usage)</b>				
<i>(Statistics for all resources are available at <a href="http://www.dlt.its.psu.edu/liasstats.html">http://www.dlt.its.psu.edu/liasstats.html</a> )</i>				
CAT - PSU Libraries Web Catalog	952,372	892,588	1,016,243	1,350,682
ProQuest Direct (newspapers, magazines, business)	87,737	140,307	214,786	285,101
LEXIS-NEXIS Academic Universe	--	39,948	45,081	90,046
PsycINFO (psychology)	63,394	65,726	52,859	70,901
Web of Science	16,716	34,325	46,328	70,246
JSTOR (full text journal articles)	23,276	31,116	42,724	56,830
MEDLINE (PubMed)	44,450	41,461	43,013	53,631
Elsevier ScienceDirect Web Editions (full text journal articles)	--	6,761	32,836	48,234
Academic IDEAL (full text journal articles)	--	17,632	34,352	39,174
WorldCat	7,217	11,728	19,582	38,401
ABI/INFORM (business)	35,142	24,947	24,473	33,220
Dow Jones Interactive	15,232	19,337	22,876	29,621
IEEE Electronic Library (IEEEExplore)	7,179	17,487	26,682	27,902
Sociological Abstracts	8,032	8,785	18,430	25,826
MUSE (full text journal articles)	13,607	12,731	14,651	20,109
CatchWord	--	1,127	8,463	18,480
Education Complete (Proquest)	--	9,551	24,932	18,851
Wiley Interscience Journals	--	--	--	17,481
Ingenta	--	--	4,506	17,254
Elsevier Science Direct (full text journal articles)	--	--	--	16,705

**Note:**

\*Data for October - December 2001 were not available. Annual statistics for 2001 were calculated using the data from the first nine months.

## Appendix C. LIAS System Availability

System Availability Based on University Libraries Hours, July 2002 - June 2003*													
	2001						2002						
	JUL	AUG	SEP	OCT	NOV	DEC	JAN	FEB	MAR	APR	MAY	JUN	Totals
<b>DOWNTIME(In Hours)</b>													
<b>Environment</b> (Air conditioning, scheduled power outages, unexpected power outages, etc.)													
	0	0	0	0	0	0	0	0	0	0	0	0	<b>0</b>
<b>Hardware</b> (CPU, MCU, HSC and disk failures, cluster interconnect, etc.)													
	9	0	0	1	0	0	0	6	0	0	9	0	<b>25</b>
<b>Scheduled Maintenance</b>													
	0	0	0	0	0	0	1	0	0	1	0	0	<b>2</b>
<b>Software</b> (Oracle, Unicorn, VMS, Unix, Netscape, Apache, E-Z Proxy, Kerberos, etc.)													
	2	7	7	1	3	1	1	4	1	2	4	2	<b>35</b>
<b>Total Downtime</b>	<b>11</b>	<b>7</b>	<b>7</b>	<b>2</b>	<b>3</b>	<b>1</b>	<b>2</b>	<b>10</b>	<b>1</b>	<b>3</b>	<b>13</b>	<b>2</b>	<b>62</b>
<b>TOTAL AVAILABLE HOURS</b>													
Maximum Available	368	280	428	434	398	344	357	404	406	436	364	359	<b>4578</b>
Total Downtime	11	7	7	2	3	1	2	10	1	3	13	2	<b>62</b>
<b>Uptime</b>	<b>357</b>	<b>273</b>	<b>421</b>	<b>432</b>	<b>395</b>	<b>343</b>	<b>355</b>	<b>394</b>	<b>405</b>	<b>433</b>	<b>351</b>	<b>357</b>	<b>4516</b>
<b>AVAILABILITY IN %</b>	<b>97.0</b>	<b>97.5</b>	<b>98.3</b>	<b>99.5</b>	<b>99.2</b>	<b>99.7</b>	<b>99.4</b>	<b>97.5</b>	<b>99.8</b>	<b>99.3</b>	<b>96.4</b>	<b>99.4</b>	<b>98.6</b>

\*Availability percentages are based on the maximum available hours at University Park Libraries and may not be a true reflection of LIAS availability at the CES campus libraries. LIAS availability at the campuses can be impacted by network downtime and other regional factors. LIAS is available for searching twenty-four hours a day, seven days a week.