Information Use Management and Policy Institute
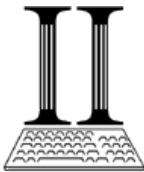## School of Information Studies, Florida State University

Evaluation Activities for the Florida Electronic Library: Data Collection Strategies and Statistics - Interim Report
August 11, 2005

Charles R. McClure, Ph. D. <cmcclure@lis.fsu.edu>
Director and Francis Eppes Professor

John T. Snead, <jsnead@fsu.edu>
Research Associate

John Carlo Bertot, Ph. D. <bertot@lis.fsu.edu>
Associate Director and Professor

School of Information Studies
Information Use Management and Policy Institute
Florida State University
Tallahassee, Florida 32306

http://www.ii.fsu.edu

EXECUTIVE SUMMARY

The Information Use Management and Policy Institute of Florida State University has been engaged in the ongoing development of the Florida Electronic Library (FEL) since 2002. During that time, the Information Institute has:

- Assisted in the clarification of the status and goals of each component of the FEL;
- Developed criteria and measures for ongoing assessment of each of these components during the component's development stages; and
- Conducted usability, functionality, and accessibility testing of selected components of the FEL as they became operational.

The purpose of the current initiative is to assess and develop long-term sustainable evaluation methods for data collection efforts of various components of the FEL. The goals of the current effort are: 1) improve the success by which the FEL meets information needs of residents of the state of Florida, the library community, and other users of the FEL; and 2) develop ongoing data collection and reporting strategies that evaluate the status and success of the FEL. Guided by these goals, the results of the current initiative will provide data that can help assure the establishment of strategies for the ongoing assessment, fine-tuning, and improvement of the FEL; and provide data that will inform the FDLIS and FLNC decision-making process for continued development of the FEL.

The current initiative of this FEL project includes the development of: 1) coordinated data collection strategies for a number of FEL components; 2) identification strategies used to select data elements for each component; 3) development of a database for the elements; and 4) sample data presentation methods for the overall use and impact of the FEL in annual and other reports. Included in these strategies is the identification and collection of an *initial* core set of data elements for each FEL component. Data development and presentation methods include guidance in the development of an initial centralized database (Evaluation Database) for both component assessment purposes and for the creation of an annual FEL report by FDLIS. The initial data elements collected within the centralized database will form a baseline for future data collection efforts and annual FEL reports. The primary product of this current initiative is the development of an *Evaluation Database*.

The intent of this initial report is to present initial data elements for each operational FEL component. These data elements will be used to inform FDLIS, FLNC, and others of the ongoing progress of the FEL and of benefits the FEL provides to the citizens of the State of Florida. The data elements selected are the results of meetings with individuals representing each FEL component and the state library liaison. This initial report is an interim presentation of collected data for available components of FEL. In fall 2005, a full final report will be issued with complete data, detailed strategies, and methodologies. Additionally, this report presents a brief list of next steps needed to complete the final report.

INTRODUCTION

   [to be added by Mark Flynn]

FEL DESCRIPTIVE DATA

This section of the interim draft report provides a summary of available data for various FEL components. Selection of the data elements collected for each component of the FEL within this report is the result of meetings with representatives for each FEL component and FDLIS. The agreed upon data collection period for inclusion within this report is January 1, 2004 through June 30, 2005. Note that data collection is an ongoing process and complete descriptive data were not yet available at the time of this interim report for all components. Any missing data collected after this report will be included within the final report.

The data presented within this report is in the form of tables and graphs. Typically, data presented in table form represents monthly statistic collection, although some annual aggregates of data have been included. Data presented within the graphs typically represents a quarterly aggregation of monthly data from the tables. Due to the time frame for data collection, quarterly aggregates do not reflect specific fiscal years.

A brief explanation is included when necessary along with some examples of potential presentations of data. Presentations of data, explanations, and examples of analysis represent the potential for use of the data within a statewide annual report. Visual elements within graphs reflect general data presentation typical of annual reports. Alternative visual data presentation graphs and analysis may be included in the final report.

The research team will continue to work closely with representatives from each component of the FEL. Continued relationships with FEL representatives throughout this process will help to ensure accuracy and reliability of: data collection methodologies, data collection efforts, and other aspects present in the development of a baseline FEL evaluation database.

1. Statewide Licensed Databases

1.1 WebTrends

WebTrends is a commercial software program that offers log analysis metrics of website usage. Collection methods of data provide limited insights into the actual usage of web pages; however, the data can provide information on trends based on web page use over time. The data elements presented below in table and graph format are just a few of the hundreds available from *WebTrends* and represent the potential for trend analysis using readily available data collection methods.

Members of the Institute study team met with Debra Flemming (FDLIS) and Mark Flynn (FDLIS) to review WebTrend generated reports. In addition to identification of initial data elements, other potential issues were discussed that involved the use of WebTrends as part of the data collection process, such as the status of WebTrends as a reliable and ongoing data collection method.

Debra Fleming provided copies of WebTrend generated reports for analysis for the time period January 2004 – June 2005. The research team evaluated the reports and recommended the following data elements: (See Table 1.1 below)

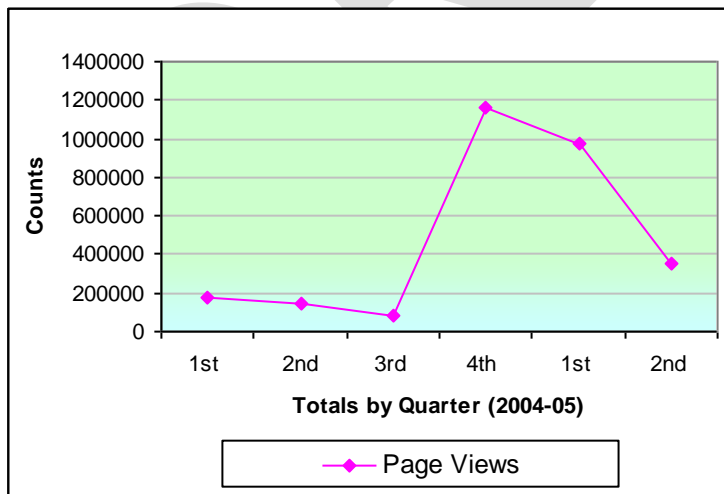| WebTrend Database Elements					
		Visits by Referring Domain	[Unique] Visitors	Total Visits	Page Views
2004	January	9013	10629	17058	60301
	February	7765	8524	14412	48490
	March	7831	9096	15409	71699
	April	6138	7322	12445	59304
	May	5116	6506	11052	41837
	June	5498	6292	11106	41260
	July	5374	6370	11583	34882
	August	4569	5584	9626	19602
	September	4097	5172	8666	27471
	October	5793	7373	13118	55526
	November	5837	7186	13178	358347
	December	4695	6100	11333	742534
2005	January	6301	7681	14309	562265
	February	9048	8491	16161	285442
	March	8765	8594	16489	126158
	April	6802	8762	16774	122280
	May	5955	8020	18086	117373
	June	7979	8216	15075	115537

Table 1.1: WebTrend Monthly Statistics per Data Element

Data collected from WebTrends is intended to supply insights into trends based on website usage. These trends can be represented in graph format. Examples are provided below: (See Graphs 1.1.1 – 1.1.3 below)
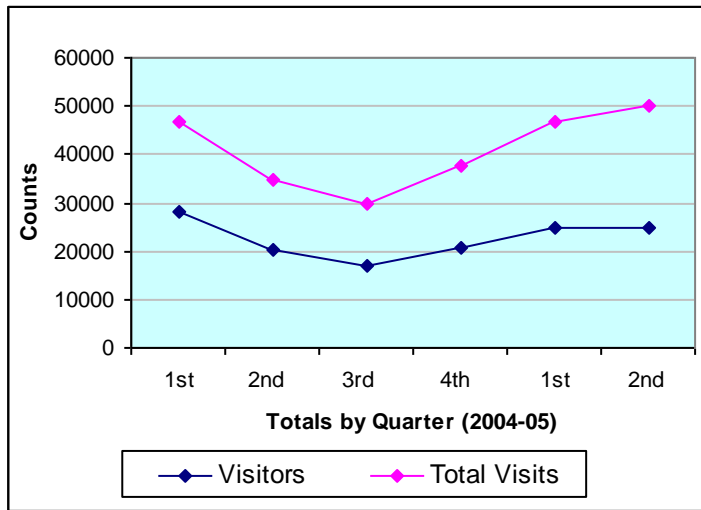
Graph 1.1.1: WebTrends Visitors and Total Visits Database Elements

Graph 1.1.1 presents a line graph comparison between the total identified unique visitors to the FEL website and total visits. Data available from WebTrends for unique visitors is also available as *Visitor, 1 time only* and *Visitor, 2 or more times* to distinguish between single use and multi-use of the FEL pages by visitors. An example of analysis of the above chart can be seen in the 2nd quarter of each year. Total visitors show a 24% increase in the 2nd quarter of 2005 compared to 2nd quarter 2004. Looking within the same time frame, total visits increased 44%. The 2nd quarter of 2005 shows a positive increase in both visitors and visits, and the higher increase of visits indicates more visits per visitor or an increase in multiple visits per visitor.

Graph 1.1.2: WebTrends Page Views Database Element

Line graph 1.1.2 shows page views per quarter from the WebTrend database elements. Information provided by Mark Flynn of FDLIS on marketing efforts of the FEL

in the fall of 2004 suggests a possible correlation between a spike within the page views from the 3rd to the 4th quarter of 2004 and the FEL marketing efforts. Subsequent quarters in 2005 present a downward turn following the 2004 4th quarter spike; however, both quarters show a considerable increase in page views from the 1st and 2nd quarters of 2004, i.e. 539% increase and 249% increase, respectively for the 1st and 2nd quarters of 2004 and 2005.
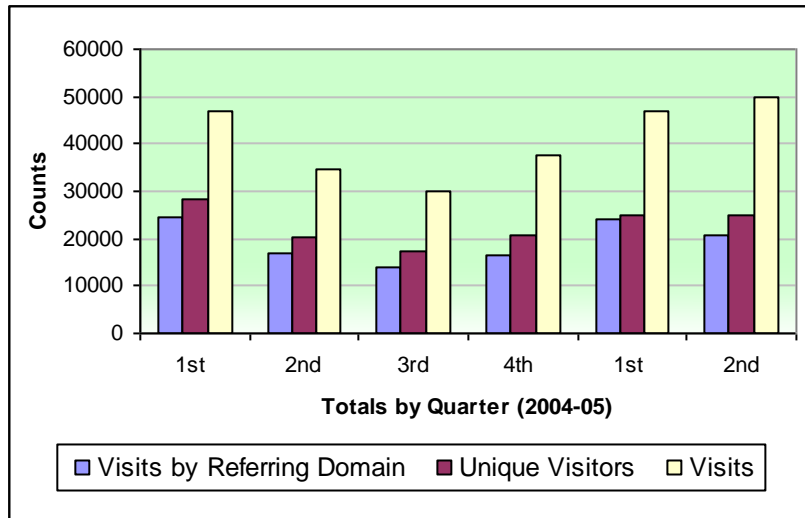
Graph 1.1.3: WebTrends Comparison of Selected Database Elements

Graph 1.1.3 shows the relationships between three WebTrend database elements: visits by referring domain, visitors (unique), and visits. This graph presents a different view of graph 1.1.1 (above) for visitors and visits. In the bar graph above, it can be seen that visits by referring domain follows a trend similar to visitors and visits from 1st quarter 2004 through 2nd quarter 2005. The relationship between visits by domain and visitors remains similar throughout the time frame with more visitors than visits from domains, as would be expected. The histogram provides a different visual, however, of the relationship between total visitors and total visits with the same results. The 1st and 2nd quarters of 2005 show a near doubling of visits to visitors compared to other quarters represented. (See Graph 1.1.1 above)

Additional Web Trend Issues:

In the past, WebTrends has undergone occasional internal changes leaving it unavailable for use for short periods of time. Based upon inquiries made by Debra Fleming, FDLIS determined these short breaks in available service did not affect the current status of WebTrends, the collection of statistics from this program, or the reliability of collected statistics. Also, WebTrends is licensed by the state. At this time, the future use of WebTrends by FEL is dependent upon the state continuing with this software for log analysis purposes.

1.2 Gale Group

        Members of the research team received prepared Gale Group reports from Mark
Flynn for the time period January 2004 – June 2005. Gale Group also provided
definitions for use in interpreting the statistics within the reports. The research team
evaluated the reports and recommended the following data elements.

Gale Group Database Elements				
	Total Sessions	Total Full-Text Downloads	Total Searches	Total Retrievals
2004 January	123972	417297	301584	636327
February	213410	1033346	490653	1560485
March	189765	808514	470235	1210476
April	73041	83116	273471	147719
May	87098	236440	247545	357134
June	131201	353358	462757	565738
July	123038	327080	428420	523518
August	64961	145609	184114	220364
September	126422	362267	536234	335257
October	175843	543021	477264	790668
November	189598	648797	502050	929923
December	81657	207717	222447	306942
2005 January	104736	226725	286185	339500
February	48351	56727	124420	80877
March	49451	64127	132295	90375
April	50872	66097	136360	91938
May	36324	39955	91345	57373
June	33098	35671	82823	51510

Table 1.2: FEL Gale Monthly Statistics per Data Element

        Gale Group provides the above elements and statistics in a monthly report to
FDLIS. Additionally, Gale Group can provide the statistics in the form of graphs for
inclusion within reports such as the proposed FEL annual report. Examples of Gale
Group graphics available to FDLIS are included within this report (see Appendix A).
The above statistics are vendor supplied, and the accuracy of the numbers is based on the
Gale Group's statistics collection methodology, policy, and procedures.

2. Union Catalog

Members of the research team, in coordination with Mark Flynn, met with Jennifer Pearson of OCLC to determine which initial database elements should be included within the FEL evaluation database. The research team received usage statistics from Jennifer Pearson for the time period January 2004 – June 2005. After review of the statistics, the following data elements were recommended: (See Table 2.1 below)
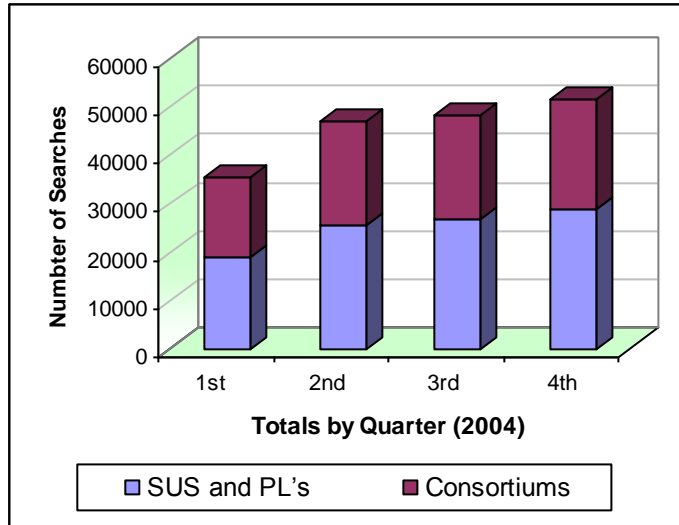
Union Catalog Database Elements			
	Searches Per Month SUS and PL's	Searches Per Month Consortiums	Total Searches Per Month
2004 January	4369	4029	8398
February	5944	5479	11423
March	8670	7075	15745
April	7360	6240	13600
May	8341	7064	15405
June	9664	8142	17806
July	9124	7097	16221
August	8748	6828	15576
September	8921	7307	16228
October	11587	9321	20908
November	9660	7617	17277
December	7551	5764	13315
2005 January	10022	7782	17804
February	10754	7766	18520
March	n/a	n/a	n/a
April	n/a	n/a	n/a
May	n/a	n/a	n/a
June	n/a	n/a	n/a

Table 2.1: Union Catalog Monthly Statistics per Data Element

Note: Data for March – June 2005 will be available for the final report. Also, data presented in the table above is per *SUS and PL's* and per *Consortiums*. Total searches per month are derived by combining statistics from data element columns 1 and 2 above.

In addition to the above database elements, OCLC Group will host and administer a survey to collect additional elements for the final database. The survey was developed by members of the research team in coordination with Mark Flynn and Jennifer Pearson. The surveys will include questions on the usability of the Union Catalog and a summary of anecdotal information collected from the surveys on overall quality and usefulness of the service. The survey form is included within this brief. (See Appendix B)

Graph 2.1: Union Catalog Database Element Statistics

The three-dimensional histogram above (graph 2.1) only shows 2004 totals by quarter. The 2005 totals were not completely available at the time of this report. The graph presents a simple, easy to read look at total searches over time for the year 2004. Looking at the bars above, an overall increase in total number of searches can be seen from the beginning of the year to the end. Additionally, a gradual year-long increase can be seen within both elements presented within the bar graph. Using table 2.1 data for a measurable increase, it can be calculated there is an overall increase in total numbers of searches from the 1st quarter to the 4th of nearly 45%.

3. Interlibrary Loan

Members of the research team coordinated with Mark Flynn and Charlie Parker to identify data elements within the evaluation database for interlibrary loan. Data were collected for the time period of January 2004 – June 2005. It was determined that the initial database elements would be limited to Courier Service data. (See Table 3.1 below)

Interlibrary Loan Database Elements		
	Total Sent Per Month	Total Received Per Month
2004 January	26197	26723
February	26689	26734
March	40707	33885
April	30714	30741
May	26919	26961
June	29169	29636
July	28732	28449
August	26643	26478
September	23393	23981
October	29900	29628
November	27907	26890
December	26422	25533
2005 January	29063	28508
February	28834	28554
March	33559	32540
April	28010	28473
May	25756	25068
June	23750	23050

Table 3.1: Courier Service Monthly Statistics per Database Element

The statistics in the table above include total number of borrows and lends of materials from within and outside of Florida. The elements are limited to Courier Service data only.
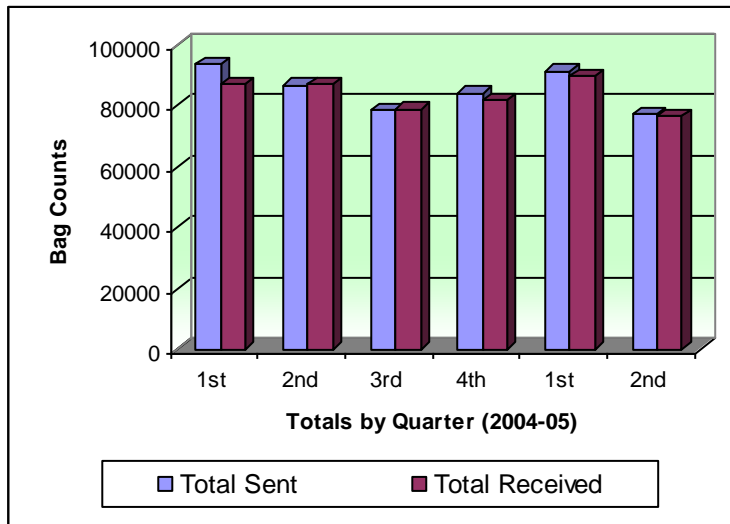
Graph 3.1: ILL Courier Service Database Element Statistics

The bar graph above shows the relationship between packets sent and packets received by quarter of the Courier Service of ILL. The number of packets sent does not give an indication of number of items actually sent as each packet could contain more than one item. The bars by quarter show a downturn in activity with a low in the 3rd quarter 2004. Additionally, the first two quarters of 2005 show a decrease in use of the Courier Service, particularly when comparing 2nd quarter counts. The information presented above could also be viewed in terms of other sources of information, i.e. data regarding an overall increase in the availability of online resources and documents from 2004 to 2005 if that data is available from other sources.
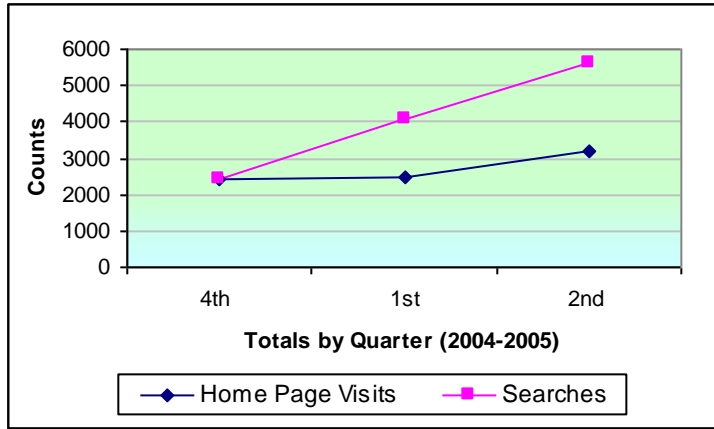
4. Florida on Florida

Members of the research team, in coordination with Mark Flynn, met with Priscilla Caplan of FCLA to determine potential database elements for the Florida on Florida FEL component to be used within the evaluation database. Priscilla Caplan provided URL's to FCLA's annual statistical reports. The research team, under the guidance of Priscilla Caplan reviewed the FCLA reports. Data collected and included within the table below came directly from the FCLA reports. Data are from the time period October 2004 – June 2005. Data are not available from January 2004 through part of October 2005; the site was in the process of development during this time frame. Additionally, the data currently collected is limited to the PALMM interface. (See Table 4.1 below)

Florida on Florida Database Elements		
	Total PALMM Home Page Visits Per Month	Number of Searches per Month (PALMM)
2004 January	0	0
February	0	0
March	0	0
April	0	0
May	0	0
June	0	0
July	0	0
August	0	0
September	0	0
*October	462	462
*November	910	910
*December	1046	1046
2005 January	847	1672
February	804	1205
March	824	1176
April	989	1888
May	1100	2720
June	1099	1011

Table 4.1: PALMM Monthly Statistics per Database Element

Note: The * shows months where the total home page visits is equivalent to the number of searches per month. A change occurred in January 2005.

In addition to the above database elements, FCLA, with the assistance of Priscilla Caplan hosted and administered a survey to collect additional data elements for the final database. The survey was developed by members of the research team in coordination with Mark Flynn and Priscilla Caplan. The surveys include questions on the usability of the Florida on Florida component of the FEL. A summary of anecdotal information collected from the surveys on overall quality and usefulness of the service, along with the usability results of the survey will be included within the final report. The survey form is included within this report. (See Appendix B)

Graph 4.1: Florida on Florida Database Element Statistics

Graph 4.1 above can be used to show a relationship between home page visits and total searches for the Florida on Florida PALMM web pages. Data collected for the 4th quarter within table 4.1 of 2004 shows equal numbers as to visits and searches in the table, unlike what is seen in the graph above. To understand what may be occurring during the 4th quarter months of 2004, a comparison of visits to searches needs to occur by looking at table 4.1 (above) along with the graph. The actual increase by percentage showing a change in the relationship between visits and searches begins in January 2005 according to table 4.1 data and not during the 4th quarter as the graph above indicates. Data within the table supports a steady increase in searches as shown above from the first to the second quarter. There is a 64% increase in searches over visits for the 1st quarter and a 76% increase in searches over visits for the 2nd quarter. These increases suggest an increase in searches per visit for each of these quarters, or in other words, each visit results in more searches per visit.

5. Virtual Reference

In addition to the database elements provided in Sections 5.1 and 5.2 below, CCLA regularly collects comments on the overall quality and usefulness of the service. A summary of collected anecdotal information will be included within the final report.
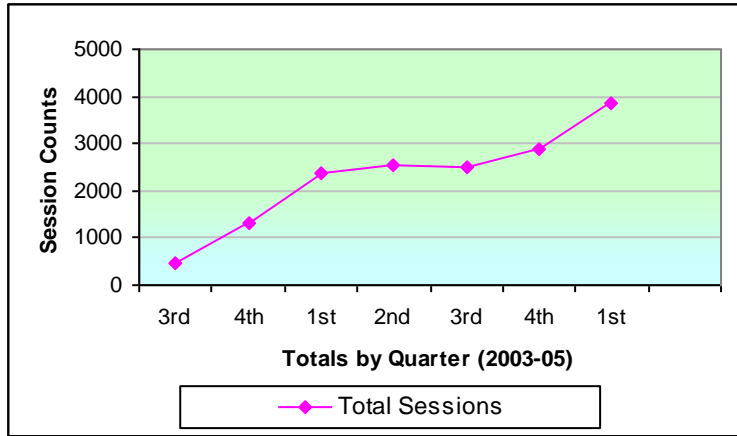
5.1 Virtual Reference Database Elements (Monthly Data)

Members of the research team met with Vince Mariner of CCLA to identify potential database elements for the Virtual Reference component of FEL. The research team along with Vince Mariner determined the initial database elements for the database. Vince Mariner provided the data for each of the elements provided in tables 6-9 below from reports generated by CCLA. The Virtual Reference service was activated July 23, 2003. Data provided is from the date of activation, July 2003 – June 2005.

Virtual Reference Database Elements (Monthly)			
		Total Sessions per Month	Total Participating Libraries per month
2003	July	40	13
	August	99	25
	September	315	38
	October	525	43
	November	305	52
	December	501	54
2004	January	721	60
	February	735	64
	March	901	64
	April	863	64
	May	761	66
	June	914	70
	July	870	73
	August	882	73
	September	753	76
	October	943	76
	November	979	76
	December	975	76
2005	January	1130	78
	February	1313	81
	March	1401	83
	April	1217	85
	May	1025	85
	June	n/a	86

Table 5.1: Virtual Reference Monthly Statistics per Database Element

Note: Data provided is from activation of the service, July 28, 2003; June 2005 session data not available.

Graph 5.1: Virtual Reference Sessions Database Element Statistics

Graph 5.1 above indicates a steady increase in the number of sessions occurring each quarter from the inception of the service during the 3rd quarter 2003 to the 1st quarter 2005. Data for the 2nd quarter 2005 was not included for the graphic above. Data has not been received for June 2005 to give an accurate reflection of 2nd quarter 2005 sessions. Another example of identifying change can be the measured difference from the 1st quarter 2004 to the 1st quarter 2005, i.e. a 63% increase in use of this service.

5.2 Additional Virtual Reference Database Elements (Annual Data)

5.2.1 Virtual Reference Sessions by Library Type

Virtual Reference Annual Database Element	
	Total Number of Sessions by Library Type - (January 1 – June 20, 2005)
Askalibrarian.org*	3016
Public Libraries	2216
Academic Libraries	1156
School Libraries	48
Special Libraries	4

Table 5.2.1: Virtual Reference Sessions by Library Type, Jan-Jun 2005.

Note: askalibrarian.org is the main entry point to the service. Users who enter
        this link are not attributed to a participating library
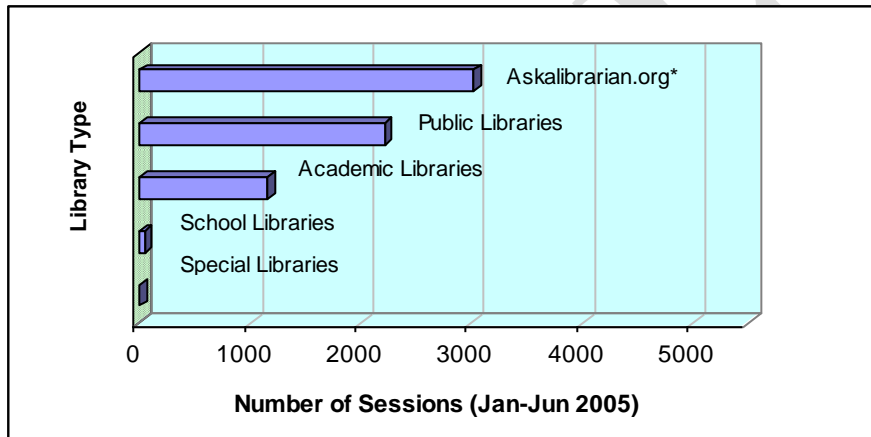
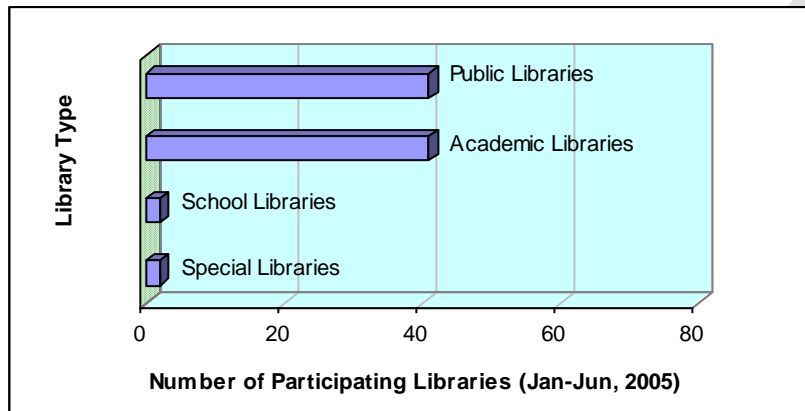Graph 5.2.1: Virtual Reference Sessions by Library Type, Jan-Jun 2005.

Graph 5.2.1 is a representation of a breakdown of total number of sessions for each library type for the six-month period from January 2005 to June 2005. This data has only been collected beginning January 2005. An example from this graph shows the askalibrarian.org element as having around 3000 sessions during this six-month period compared to about 2200 sessions from public libraries. The graph allows not only a breakdown of sessions per library type, but also a comparison between library types. The graph raises a number of potential issues related to directing marketing efforts, identifying preferred entry points, potential demographics of users, etc.

5.2.2: Virtual Reference Participating Libraries by Type

Virtual Reference Annual Database Element	
	Total Participating Libraries by Type January 2005 - June 2005
Public Libraries	41
Academic Libraries	41
School Libraries	2
Special Libraries	2

Table 5.2.2: Virtual Reference Participating Libraries by Type, Jan-Jun 2005.

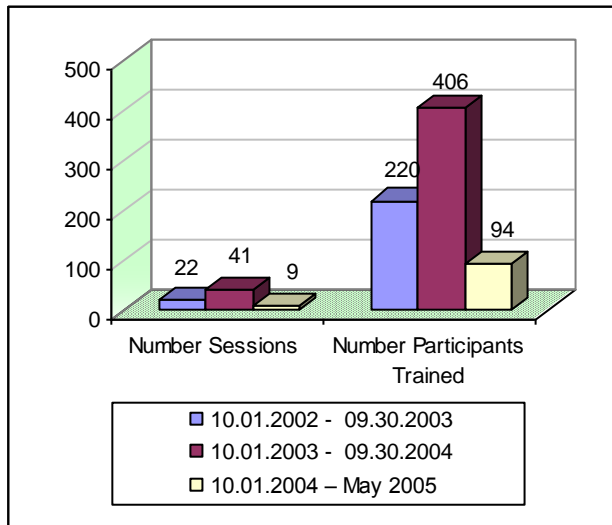Graph 5.2.2: Virtual Reference Participating Libraries by Type, Jan-Jun 2005.

Graph 5.2.2 shows relationships between the total numbers by type of library that use the Virtual Reference service. As in graph 5.2.2, these relationships can be used to identify specific types of libraries to target for marketing efforts, potential demographics of the users of the service, etc.

5.2.3 Virtual Reference Training Sessions and Participants

Virtual Reference Annual Database Elements		
Training Sessions	Number Sessions	Number Participants Trained
10.01.2002 - 09.30.2003	22	220
10.01.2003 - 09.30.2004	41	406
10.01.2004 – May 2005	9	94

Table 5.2.3: Number of Participants Trained and Number Sessions by Fiscal Year.

Note: The number of sessions is based on average of 10 participants per session.

Graph 5.2.3: Virtual Reference Training Sessions

Graph 5.2.3 above shows potential problems in presenting relationships between two elements, i.e. comparisons of increases in sessions to increases in participants. The numbers for the sessions presented above were derived form total participants based on an average of 10 participants per session (derived calculation method provided with the data received for the Virtual Reference services). There is an 86% increase in sessions from 2003 to 2004 and a nearly identical 84% increase in participants trained.

NEXT STEPS

The intent of this initial report is to present in tabular form selected initial data elements identified for each operational FEL component collected to date. Final data collection efforts for the FEL evaluation database will also be submitted electronically. Next steps include:

- Final data collection – collect missing data needed to complete the tables above along with final collection of data from surveys conducted for the Florida on Florida and the Union Catalog components of the FEL;
- Produce Final Report;
- Provide Mark Flynn of FDLIS with an electronic version of all data collected as the baseline for the creation of the FEL evaluation database;
- Recommend additional data elements to be provided in the final report; and
- Recommend procedures, strategies, and methodologies in the implementation of the FEL evaluation database and in the creation of the first FEL annual report.

Data element identification and data collected within this report are part of a joint effort between the research team, individuals associated with each of the FEL components, and individuals from FDLIS and FLNC.

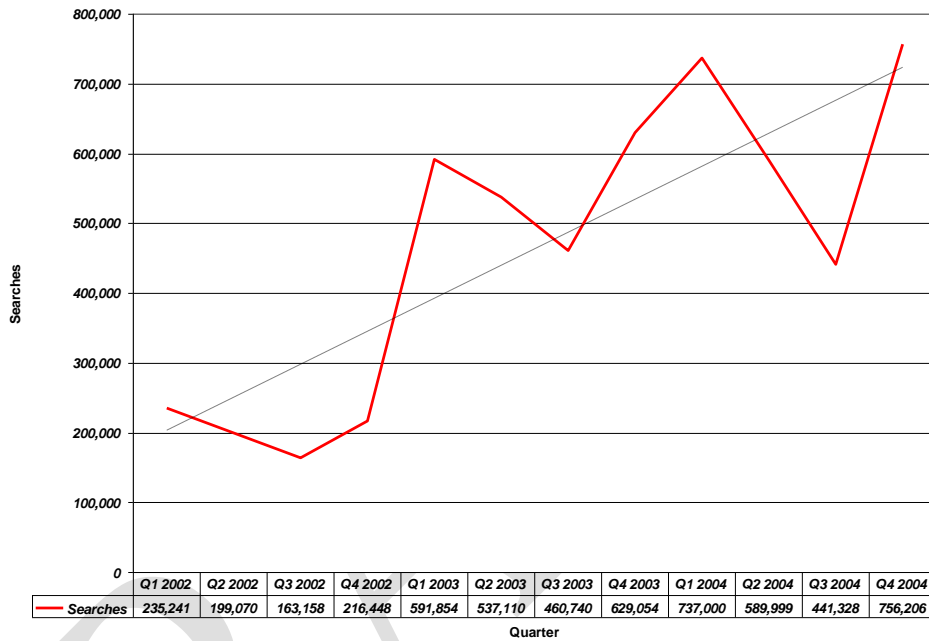Appendix A: Examples of graphs supplied by Gale Group to FDLIS.

Example 1

Through the Florida Electronic Library, over 5 million searches have been served in the past two years.

The content offered has increased by nearly 3,000 full-text periodicals and newspapers. The overall title count for the periodical products increased from 6,325 to 9,775. InfoTrac Onefile increased from a 45% full-text product, to over 60% full-text.

Usage of Gale products is rapidly increasing with the highest levels reached in the last two months.

2005 year is extremely exciting for both the Florida Electronic Library and Thomson Gale with the launch of Thomson Gale PowerSearch. This new platform allows users to cross-search all of their resources from one, easy to use interface.

Florida Electronic Library Searches on Thomson Gale Products

	Q1 2002	Q2 2002	Q3 2002	Q4 2002	Q1 2003	Q2 2003	Q3 2003	Q4 2003	Q1 2004	Q2 2004	Q3 2004	Q4 2004
Searches	235,241	199,070	163,158	216,448	591,854	537,110	460,740	629,054	737,000	589,999	441,328	756,206

Note: Sidebar with explanatory text of graph has been modified from the original provided by Gale Group by increasing font size from 5.5 to 8 for readability purposes.

Example 2

Florida Electronic Library Searches on Thomson Gale Products

	Q1 2002	Q2 2002	Q3 2002	Q4 2002	Q1 2003	Q2 2003	Q3 2003	Q4 2003	Q1 2004	Q2 2004	Q3 2004	Q4 2004
Searches	235,241	199,070	163,158	216,448	591,854	537,110	460,740	629,054	737,000	589,999	441,328	756,206

Appendix B: Online Survey for Union Catalog and Florida on Florida anecdotal data collection, website usability, and website user satisfaction.

Online User Survey				
	Strongly Agree	Agree	Disagree	Strongly Disagree
1. Overall, this website is easy to use.	O	O	O	O
2. The [Florida on Florida/Florida Group Catalog] *search techniques* are easy to use.	O	O	O	O
1. I am very satisfied with my search results or with the information I received from the website.	O	O	O	O
4. This website and service could be improved by doing the following: a. b.				