

Understanding User OPAC Searching Habits Through WebVoyagé Log Analysis: A Comparative Study of Two Cases

Weiling Liu
University of Louisville
w.liu@louisville.edu

Research Question and Objectives

In an OPAC system, if the default options are not "Keyword" and "all of these," are they still the most frequently used options?

Objectives:

- To find out whether "Simple Search", "Keyword" and "All of these" are still the most frequently used options if the default settings are different
- To see what other information can be useful in helping us understand the user's Voyager OPAC search trends

Cases

- University of Louisville (UofL)
 - Front Page Default:
 - Simple Search
 - Search Defaults:
 - Simple Search - (Keyword – all fields)
 - Advanced Search - (“All of these” and “Keyword – all fields”)
- Western Kentucky University (WKU)
 - Front Page Default:
 - Basic Search
 - Search Defaults:
 - Basic Search - (Title – alphabetical)
 - Guided Keyword Search - (“As a phrase” and “Keyword – anywhere”)

Background

- University of Louisville (UofL)
 - Fall 2008 student enrollment: 21,761
 - Six branch libraries (about 2 million volumes)
 - Voyager (version 7.03)
 - Embedded search box on two of the Libraries' websites
 - SFX (with linking connection to Voyager)
 - MetaLib (with search connection to Voyager)
- Western Kentucky University (WKU)
 - Fall 2008 student enrollment: 19,761
 - Five branch libraries
 - Voyager (version 7.03)
 - TDNet (with linking connection to Voyager)

What is a Web Transaction Log?

“an electronic record of interactions that have occurred during a searching episode between a Web search engine and users searching for information on that Web search engine. [1]”

- Fields in a standard search log: IP, Date, Time and Search URL

Example: a Web transaction log record of a search on WebVoyagé:

```
xxx.xxx.101.96 - - [07/Jan/2008:16:24:43 -0500] "GET /cgi-  
bin/Pwebrecon.cgi?  
Search_Arg=economics&SL=None&Search_Code=CMD&PID=gP  
0n1__uryEp0-X3YO7qndbyROW08I-  
Ka&SEQ=20080107162433&CNT=25&HIST=1 HTTP/1.1" 200  
6566
```

What is Transaction Log Analysis (TLA)?

- For Web-searching research, TLA is defined as the use of transaction logs “to investigate particular research questions concerning interactions among Web users, the Web search engine, or the Web content during searching episodes. [1]”
- The three common levels of TLA [1]:
 - Term – focusing on measures term occurrence
 - Query – using the query as the base metric
 - Session – examining the interactions within a searching episode

Why Use TLA?

- “The use of transaction log analysis is a **behaviorist research method**, with a natural reliance on the expressions of interactions as behaviors. [2]”
- TLA allows us to record and study users’ searching behavior in a “**non-intrusive fashion**” [3].
- Since 1960’s, TLA has been **widely used** by researchers and practitioners to evaluate traditional information retrieval (IR) systems, library systems, Web sites and Web search systems ([4],[5], [6],[8]). As many studies ([3],[7],[9],[10],[11]) have indicated, TLA is a **useful method for OPAC usage studies** on user behavior trends.

TLA Shortcomings

- As Kurth (1993) [12] reports:
 - Transaction Logs
 - Do not record the users' Perceptions of the search
 - Emotions
 - Background skills
 - TLA's methodological issues
 - Execution –difficulty due to the hefty volume and complexity
 - Conception – difficulty due to the complexity
 - Communication - problems occur when terms and metrics are not defined in sufficient detail

TLA on OPAC Usage Studies - 1

- 1990s
 - Blecic, D. D., N. S. Bangalore, et al. (1998): **on success rate** and tested to see if it could be improved by making changes in the introductory screens
 - Wyly, B.J. (1996): on uses of specific **access points** (subject, author, title, and other fields)
 - Atlas, M. C., K. R. Little, et al. (1997): **on types of searches, hit rates** and errors to judge the effectiveness of using flip charts at the OPAC terminals
 - Ferl, T. E. and L. Millsap (1996): **on subject searching** - the results of an online survey of users accessing OPAC in the libraries and from remote sites
 - Millsap, L. and T. E. Ferl (1993): **on search behavior** (types of searches, choice of search mode and database, number of retrievals, etc.) of remote users

TLA in OPAC Usage Studies - 2

□ Recent:

- Knievel, J. E., et. al., (2009): on user **search choice** (keyword, title, author, subject, ISN, and LC call number) to investigate if interface design had a measurable influence
- Moulaison, H. L. (2008): on **the use of advanced search** and user reaction when getting zero hits
- Malliari, A . & Kyriaki-Manessi, D. (2007): on users' **searching behavior** (user profiles, patterns, errors, success and level of satisfaction)
- Villén-Rueda, L., Senso, J. A., & Moya-Angón, F. de. (2007): on **types of searches** to find out which was the most frequent type of search among different user types

Voyager's Web Transaction Log

1	xxx.xxx.101.96 - - [07/Jan/2009:16:24:33 -0500] "GET /cgi-bin/Pwebrecon.cgi?DB=local&PAGE=First HTTP/1.1" 200 4335
2	xxx.xxx.101.96 - - [07/Jan/2009:16:24:33 -0500] "GET /css/webvoyageStyles.css HTTP/1.1" 200 345
3	xxx.xxx.101.96 - - [07/Jan/2009:16:24:33 -0500] "GET /js/emich.js HTTP/1.1" 200 2115
4	xxx.xxx.101.96 - - [07/Jan/2009:16:24:33 -0500] "GET /images/DownSearch.gif HTTP/1.1" 200 964
5	xxx.xxx.101.96 - - [07/Jan/2009:16:24:33 -0500] "GET /images/DisabledHeadings.gif HTTP/1.1" 200 905
6	xxx.xxx.101.96 - - [07/Jan/2009:16:24:33 -0500] "GET /images/DisabledTitles.gif HTTP/1.1" 200 770
7	xxx.xxx.101.96 - - [07/Jan/2009:16:24:33 -0500] "GET /images/DisabledHistory.gif HTTP/1.1" 200 1131
8	xxx.xxx.101.96 - - [07/Jan/2009:16:24:33 -0500] "GET /images/UpPatron.gif HTTP/1.1" 200 1434
9	xxx.xxx.101.96 - - [07/Jan/2009:16:24:33 -0500] "GET /images/UpRequest.gif HTTP/1.1" 200 1408
10	xxx.xxx.101.96 - - [07/Jan/2009:16:24:33 -0500] "GET /images/UpHelp.gif HTTP/1.1" 200 649
11	xxx.xxx.101.96 - - [07/Jan/2009:16:24:33 -0500] "GET /images/UpExit.gif HTTP/1.1" 200 759
12	xxx.xxx.101.96 - - [07/Jan/2009:16:24:33 -0500] "GET /images//limit.gif HTTP/1.1" 200 1168
13	xxx.xxx.101.96 - - [07/Jan/2009:16:24:43 -0500] "GET /cgi-bin/Pwebrecon.cgi?Search_Arg=economics&SL=None&Search_Code=CMD&PID=gP0n1__uryEp0-X3YO7qndbyROW08I-Ka&SEQ=20090107162433&CNT=25&HIST=1 HTTP/1.1" 200 6566
14	xxx.xxx.101.96 - - [07/Jan/2009:16:24:44 -0500] "GET /images/DownTitles.gif HTTP/1.1" 200 770
15	xxx.xxx.101.96 - - [07/Jan/2009:16:24:44 -0500] "GET /images/UpSearch.gif HTTP/1.1" 200 924
16	xxx.xxx.101.96 - - [07/Jan/2009:16:24:44 -0500] "GET /images/UpHistory.gif HTTP/1.1" 200 1243
17	xxx.xxx.101.96 - - [07/Jan/2009:16:24:44 -0500] "GET /images//d-prev.gif HTTP/1.1" 200 1021
18	xxx.xxx.101.96 - - [07/Jan/2009:16:24:44 -0500] "GET /images//s-next.gif HTTP/1.1" 200 766

Voyager's OPAC Search Log

Date	stat_string	session_id	search_type	Search_string	Limit	Limit_s	Ind	Rel	Hyp	Hits
2009.01.07.16.2	WebOpac	2009010716243	Keyword--boolean (and,	economics	N		K	N	N	18316
2009.01.07.16.2	WebOpac	2009010716243	Simple Search	smith	N		B	N	N	-1
2009.01.07.16.2	WebOpac	2009010716243	Author--Sorted by Title	smith	N		B	N	N	-1
2009.01.07.16.2	WebOpac	2009010716243	Call Number	qa 76.76	N		B	N	N	-1
2009.01.07.16.2	WebOpac	2009010716243	Journal Title	JALL JOURNAL OF LIBRARY TECH	N	TYPE=?S		N	N	0
2009.01.07.16.2	WebOpac	2009010716243	Journal Title	JALL LIBRARY JOURNAL?	N	TYPE=?S		N	N	5
2009.01.07.16.2	WebOpac	2009010716243	Keyword	economics	N		K	Y	N	18316
2009.01.07.16.2	WebOpac	2009010716243	Subject	economics	N		B	N	N	-1
2009.01.07.16.2	WebOpac	2009010716243	Title	TALL ECONOMICS?	N			N	N	1310
2009.01.07.16.2	WebOpac	2009010716243	Advanced	(GKEY economics)	N		K	N	N	18316
2009.01.07.16.3	WebOpac	2009010716243	Advanced	(TKEY economics)	N		K	N	N	5039

Web Log vs. OPAC Search Log

Web Log Description	Web Log Field	Search Log FIELD	Search Log DESCRIPTION
	Date	<u>Search_date</u>	Search date and time.
	n/a	<u>Stat_string</u>	The requesting client's 10-character ID number from the OPAC.ini file (see step 2 of "Setting Up OPAC Search Logging" on 15-4).
	SEQ	<u>Session_id</u>	String generated by OPAC client when started or reset (yyyymmddhhmmss)
Can tell which Boolean option is used	Search_Code	<u>Search_type</u>	Type of search performed (Author, Subject.).
	Search_Arg, SAB1, SAB2, SAB3, etc	<u>Search_string</u>	Query entered as the criteria for the search.
	SL	<u>Limit_flag</u>	Were limits in affect? Y/N
	SL	<u>Limit_string</u>	Details on limits.
	n/a	<u>Index_type</u>	A=authority, B=browse, K=keyword or L=leftanchored.
	n/a	<u>Relevance</u>	Were the results relevance ranked? Y/N
	n/a	<u>Hyperlink</u>	Was the search the result of clicking a hyperlink? Y/N
	n/a	Hits	The number of hits returned by the search. Note: Hit count will not apply to Heading Subject and Heading Call Number searches which are browse type searches of an entire index. For these searches, a "-1" will be recorded.
Determined by query string	Search_Code	<u>Search_tab</u>	0-5 search tab number.
Determined by query string Can't tell if it is from Z39.50, but it is possible to tell if a search was initiated from the embedded search box in a Web page or Linked from another page, like SFX	n/a	<u>Client_type</u>	W=Web OPAC A=ASCII OPAC Z=Z39.50
	IP	<u>Client_ip</u>	The IP (Internet Protocol) address of the requesting client, for example, 128.218.1.38.
	n/a	<u>dbkey</u>	Unique identifier of the database which the user performed the search against.
Determined by query string It is possible to tell if it follows a link to browse Author, Subject or Call Number	n/a	Redirect Flag	Y= Search performed was redirected. N= Search performed was not redirected.

Note: n/a – not applied to the current data set for this particular study

Why not Voyager's OPAC Search Log?

- No information about the use of Boolean Search options (e.g. “all of these”)
- No way to identify searches initiated from other Web systems
- The logging function needs to be enabled before data collection
- Additional maintenance on the logs

Methodology

- Scope:
 - One year's (2008) Web transaction logs for Voyager on the searches initiated from WebVoyagé classic interface:
 - Boolean Search tab (AD)
 - Simple Search tab (S)
- Analysis Level
 - Query
- Tools
 - Perl script, XML, MS Access and MS Excel
- Processes
 - Collect data and identify search query string patterns and options
 - Clean and prepare data
 - Analyze data and generate reports

Data Collection

- Collecting the data
 - Download the 2008 logs for both UofL and WKU
- Identifying search query string patterns and options
 - Enable OPAC Search Log
 - Perform test searches with each search type and option
 - Extract Web log and OPAC Search Log for the test searches
 - Compare both logs and identify patterns

UofL-Simple Search (default)

LIBRARIES : MINERVA

New Search

Index List

Title List

Search History

My Library Account

Request RRS Item

Help

Logoff

University of Louisville Libraries

Basic Search

Advanced Search

Course Reserves

New Items

Journal Finder

1. Search For:

Quick Limit (Optional)

Limits apply only to Title and Keyword searches

None

2. Find Results in: (*Can Use Limits)

Keyword--all fields (and, or, not, "phrase")*
Author (Last Name, First Name)
Author--Sorted by Title
Call Number
Journal Title*
Keyword--all fields (ranked)*
Subject
Title*

25 records per page

Search

Reset

Set More Limits

Basic Search Tips

Below are some tips for each type of search. If you have questions, please contact the Ekstrom Reference Desk at [\(502\) 852-6747](tel:502-852-6747). Current U of L students, staff, and faculty can use Interlibrary Loan for materials not available in Minerva.

There are no journal articles in Minerva. To find magazine, journal, or newspaper articles, return to the [Libraries Home](#) page and use the [Databases A-Z](#) or [Subject Guides](#) buttons.

Keyword--all fields (and, or, not)

- Combine search words with Boolean operators **and, or, not** [students and stress](#)
- Use opening and closing parentheses to group search words ([college or university](#)) [and students](#)
- Use ? to truncate [teen?](#) Finds [teenage, teens, teen, teenager, teenagers](#)

Journal Title

- Type as much of the name of a journal, magazine, newspaper, or other periodical as you know [sports illustrated](#)
- Do not type initial articles at the beginning of a title (a, an, the, la, los, die, der)

Keyword Search--all fields (ranked)

UofL - Simple Search Options

Option default: (Keyword—all fields)

Code - Web Display

CMD - Keyword-all fields (and, or, not, “phrase”) *

NAME + - Author—(Last Name, First Name)

AUTH + - Author—Sorted by Title

CALL+ - Call Number

JALL - Journal Title *

FT* - Keyword—all fields (ranked) *

SUBJ+ - Subject

TALL – Title *

(Note: “+” is recorded as “_” in the log file, e.g. NAME+ is logged as NAME_)

UofL-Advanced Search

LIBRARIES: MINERVA

[New Search](#) [Index List](#) [Title List](#) [Search History](#) [My Library Account](#) [Request RRS Item](#) [Help](#) [Logoff](#)

University of Louisville Libraries

[Basic Search](#) [Advanced Search](#) [Course Reserves](#) [New Items](#) [Journal Finder](#)

Search for: [all of these](#) Search in: [Keyword--all fields](#)

AND OR NOT

Search for: [all of these](#) Search in: [Keyword--all fields](#)

AND OR NOT

Search for: [all of these](#) Search in: [Keyword--all fields](#)

25 records per page [Search](#) [Clear](#) [Set More Limits](#)

Advanced Search

An Advanced Search finds records using keywords located anywhere in a record, or only in specific fields. To perform an Advanced Search:

1. Type in the search word(s) you want to find in the *Search For* free text field. (Punctuation, case, and word order are ignored.)
2. Default search criteria can be modified for your search:
 - o Select from the *Search In* drop-down list whether you want to search for all the search terms (all of these), any of the search terms (any of these), or the search terms as a phrase (as a phrase).
 - o Select which Boolean operator you want applied to the next search word(s): AND, OR, or NOT.
 - o Narrow your search by further limiting by entering information in the next free text fields and selecting options and Boolean operators from the drop down menu as needed.
3. Click the Limits button if you want to limit your keyword search based on other criteria.
4. Click the Search button to begin your search. Click the Reset button to clear the search page.

[New Search](#) [Index List](#) [Title List](#) [Search History](#) [My Library Account](#) [Request RRS Item](#) [Help](#) [Logoff](#)

UofL – Advanced Search Options

Default: (all of these) in (Keywords—all fields), operators: AND

Code – Web Display

GKEY - Keywords-all fields

TKEY - Title

ISBN – ISBN

ISSN – ISSN

NKEY - Author Name

SKEY - Subject

260B - Publisher: Name

SERI - Series

260C - Publisher: Date

100A - Personal Name

KEYW - New Keyword

WKU-Basic Search (default)

TOPCAT - Western Kentucky University Libraries Catalog

Search Headings Titles Patron History Remote Library Help Exit

Database Name: WESTERN KENTUCKY UNIVERSITY LIBRARIES

Basic Search

Guided Keyword Search

Course Reserve

Search for:

Quick Limit (Optional):

None

Search by (*limits can be used):

Title* (Alphabetical)
Title* (Keyword in Title)
Keyword AND*
Keyword*
Author/Creator (Last name first)
Author/Creator sorted by Title
Subject (Browsable)
Subject* (Keyword in Subject Heading)

25 records per page

Search

Reset

Limits

Off-Campus access to licensed web resources

- Off-campus users who wish to access links to licensed electronic materials **must** [Login](#) here.
- Most licensed resources are available only to Western Kentucky University students and employees from off campus.
- Username and password from your WKU issued e-mail account are required for login.
- Usernames and passwords from other e-mail accounts such as AOL, MSN Hotmail, Yahoo, etc. will **not** work.

Internet

100%

WKU – Basic Search Options

(Option default: Title (Alphabetical))

Code - Web Display

TALL – Title* (Alphabetical)

TKEY^ - Title* (Keyword in Title)

GKEY^* - Keyword AND*

FT*- Keyword*

NAME+ - Author/Creator (Last name first)

AUTH+ - Author/Creator sorted by Title

SUBJ+ - Subject (Browsable)

SKEY^ - Subject (Keyword in Subject Heading)

JALL - Journal Title (Alphabetical)

JKEY^ - Journal Title (Keyword in title)

CALL+ - Call Number (Browsable)

CALL - Call Number (Exact)

TITL+ - Series/Uniform Title

CMD* - Boolean with Relevance

008D - Date

ISBL - ISBN

ISSL – ISSN

NULL - _____

WKU-Guided Keyword Search

TOPCAT - Western Kentucky University Libraries Catalog

[Search](#) [Headings](#) [Titles](#) [Patron](#) [History](#) [Remote](#) [Library](#) [Help](#) [Exit](#)

Database Name: WESTERN KENTUCKY UNIVERSITY LIBRARIES

[Basic Search](#) [Guided Keyword Search](#) [Course Reserve](#)

Search for: as a phrase Search by:

AND OR NOT

Search for: as a phrase Search by:

AND OR NOT

Search for: as a phrase Search by:

25 records per page

[Limits](#)

Guided Keyword Search

[New Search](#) [Headings List](#) [Titles List](#) [Patron Information](#) [Search History](#) [Other Libraries' Catalogs](#) [Library Information](#) [Help](#) [Exit](#)

Need assistance? [Ask Us!](#)

Internet

100%

WKU – Guided Keyword Search Options

Default: (as a phrase) in (Keyword Anywhere), operators: AND

Code – Web Display

GKEY - Keyword Anywhere

TKEY - Title

SKEY - Subject

ISSN - ISSN

NKEY - Author Name

ISBN - ISBN

100A - Personal Name

260C - Publisher: Date

260B - Publisher: Name

SERI - Series

K505 - Table of Contents

K696 - Subject Local 696

Search Query String Patterns - Examples

Simple/Basic Search

```
XXX.XXX.XXX.XXX - -  
[01/Jul/2008:09:47:58 -0400]  
"GET /cgi-bin/Pwebrecon.cgi?  
Search_Arg=social+work&SL=Non  
&Search_Code=JALL&PID=  
pBECivvOWH62pqf9Y3IEC2o0q  
oN&SEQ=20080701094749&CNT  
T=25&HIST=1 HTTP/1.1" 200  
41703
```

Advanced/Guided Keyword Search

```
XXX.XXX.XXX.XXX - -  
[22/Feb/2008:10:33:54 -0500]  
"GET /cgi-bin/Pwebrecon.cgi?  
SAB1=economic&BOOL1=as+a+  
phrase&FLD1=Keyword+Anywher  
e+%28GKEY  
%29&GRP1=AND+with+next+set  
&SAB2=&BOOL2=as+a+phrase&  
FLD2=Keyword+Anywhere+  
%28GKEY  
%29&GRP2=AND+with+next+set  
&SAB3=&BOOL3=as+a+phrase&  
FLD3=Keyword+Anywhere+  
%28GKEY  
%29&PID=gPdJeyJCkbMthPzGyCL  
W8F4stk4rGSV0lp&SEQ=200802  
22103344&CNT=25&HIST=1  
HTTP/1.1" 200 6859
```

Search Query String Patterns - Extraction Criteria

Field:	stype	box	SAB1	SAB2	SAB3	BOOL1	BOOL2	BOOL3	FLD1	FLD2	FLD3	GRP1	GRP2	SEQ	HIST
Table:	qry_SEARCHall	qry_SEARCHall	qry_SEARCHall	qry_SEARCHall	qry_SEARCHall	qry_SEARCHall	qry_SEARCHall	qry_SEARCHall	qry_SEARCHall	qry_SEARCHall	qry_SEARCHall	qry_SEARCHall	qry_SEARCHall	qry_SEARCHall	qry_SEARCHall
Sort:															
Show:	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Criteria:	"ad"	"sab0"	"**"	"**"	"**"	"<>""	"<>""	"<>""	"<>""	"<>""	"<>""	"<>""	"<>""	"<>""	"1"
or:	"ad"	"sab1"	"<>""	"**"	"**"	"<>""	"<>""	"<>""	"<>""	"<>""	"<>""	"<>""	"<>""	"<>""	"1"
	"ad"	"sab12"	"<>""	"<>""	"**"	"<>""	"<>""	"<>""	"<>""	"<>""	"<>""	"<>""	"<>""	"<>""	"1"
	"ad"	"sab123"	"<>""	"<>""	"<>""	"<>""	"<>""	"<>""	"<>""	"<>""	"<>""	"<>""	"<>""	"<>""	"1"
	"ad"	"sab2"	"**"	"<>""	"**"	"<>""	"<>""	"<>""	"<>""	"<>""	"<>""	"<>""	"<>""	"<>""	"1"
	"ad"	"sab23"	"**"	"<>""	"<>""	"<>""	"<>""	"<>""	"<>""	"<>""	"<>""	"<>""	"<>""	"<>""	"1"
	"ad"	"sab3"	" "	"**"	"<>""	"<>""	"<>""	"<>""	"<>""	"<>""	"<>""	"<>""	"<>""	"<>""	"1"
	"ad"	"sab13"	"<>""	"**"	"<>""	"<>""	"<>""	"<>""	"<>""	"<>""	"<>""	"<>""	"<>""	"<>""	"1"

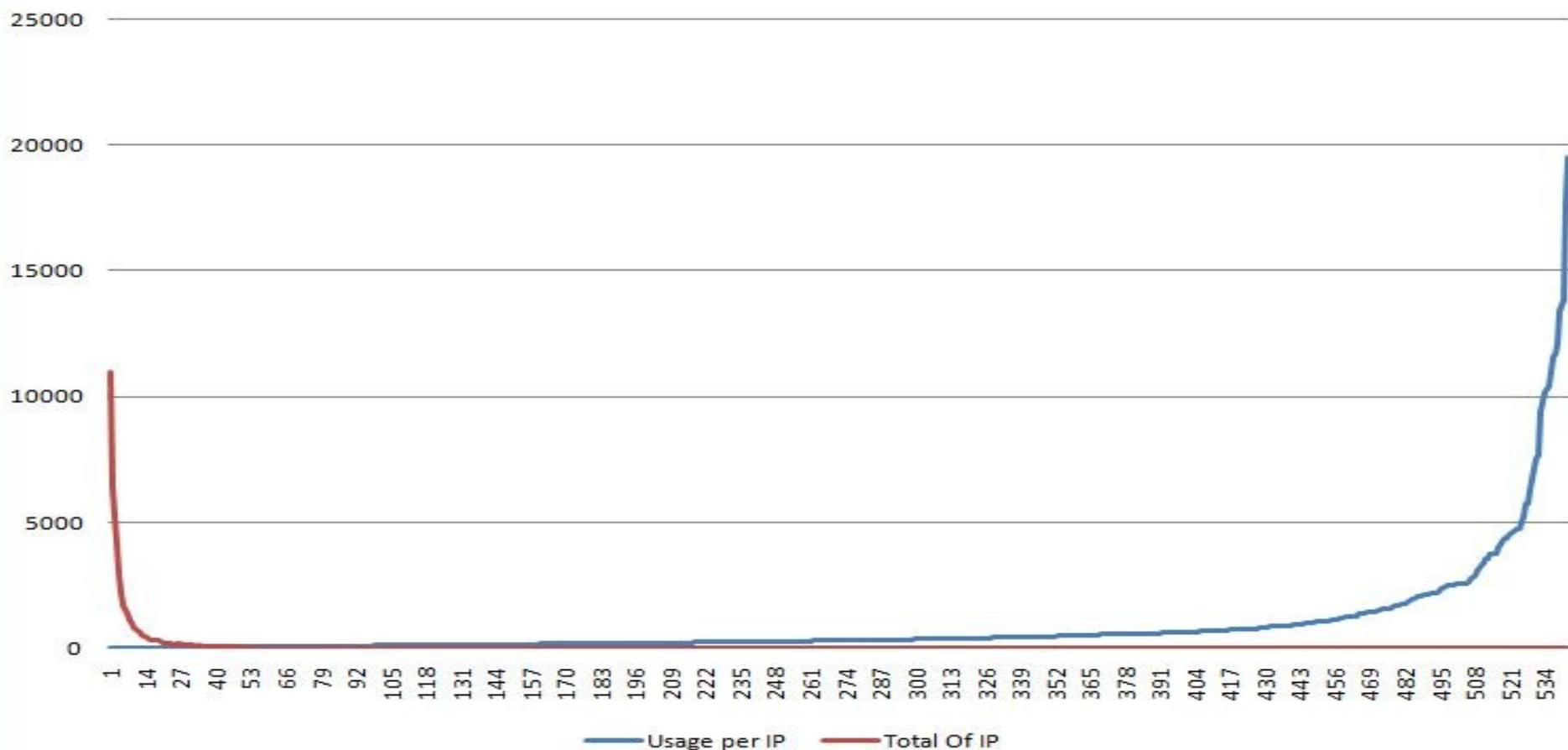
Field:	stype	HIST	SEQ	Search_Arg	Search_Code	SL
Table:	qry_SEARCHall	qry_SEARCHall	qry_SEARCHall	qry_SEARCHall	qry_SEARCHall	qry_SEARCHall
Sort:						
Show:	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Criteria:	"s"	"1"	"<>""	"<>""	"<>""	Is Not Null
or:	"s"	"1"	"<>""	"<>""	"<>""	Is Not Null
	s"	"1"	"<>""	"<>""		Is Not Null

Data Preparation

- Cleaning and parsing the data
 - Extract log data that matches the search query patterns with a Perl script and output the data to XML files by month
- Normalizing searching episodes
 - Import the extracted logs in XML format into MS Access
 - Tag each record with search type and month
 - Extract unique records
 - Separate User groups
 - In-library if possible
 - On-campus if possible
 - Remote
 - Group IPs by searches per IP for each user group
 - LE5 – Users (IPs) with searches less than or equal to 5
 - GT5LE100 – Users (IPs) with searches greater than 5 and less than and equal to 100
 - GT100 – Users (IPs) with searches greater than 100

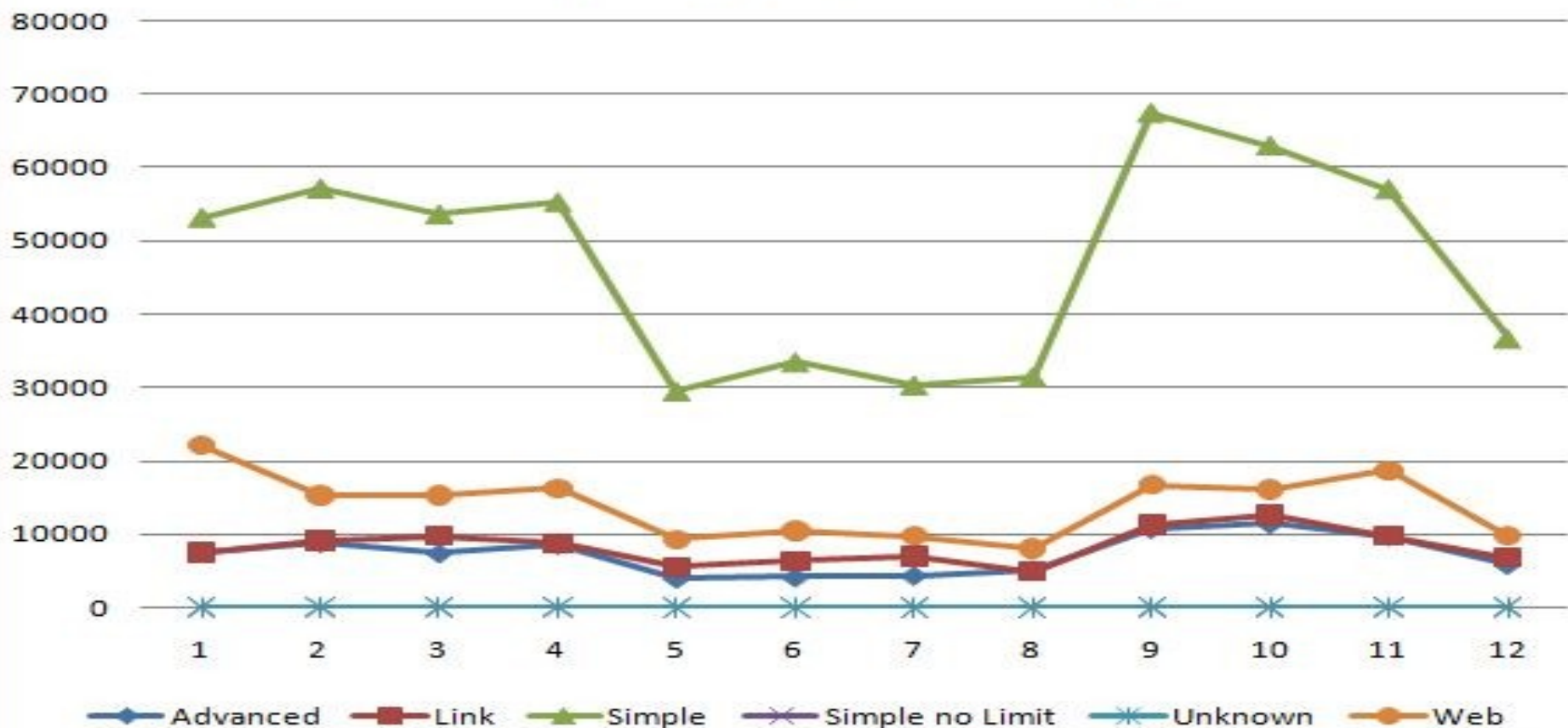
Data Overview – Example - 1

UofL - 2008
OPAC Usages per IP



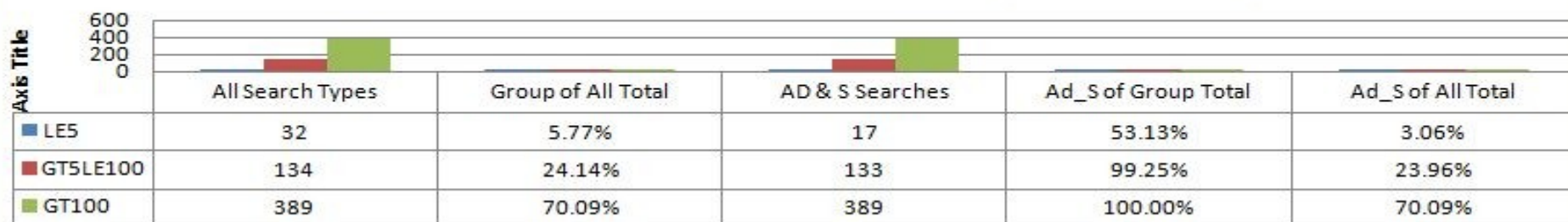
Data Overview – Example - 2

UofL - 2008 Monthly Usage by Search Types

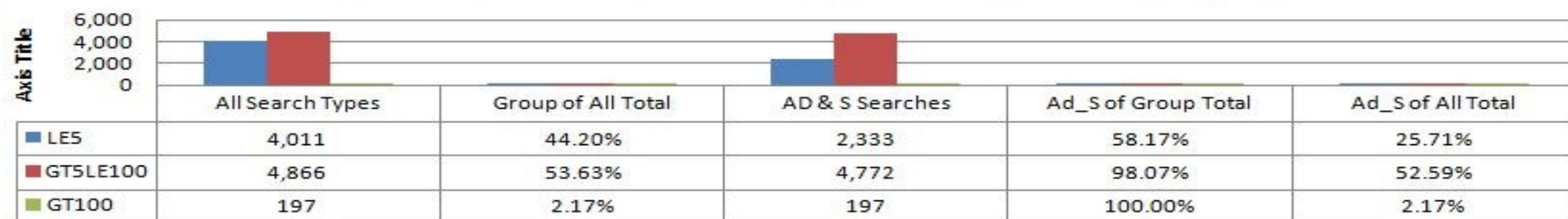


Data Overview – Example 3

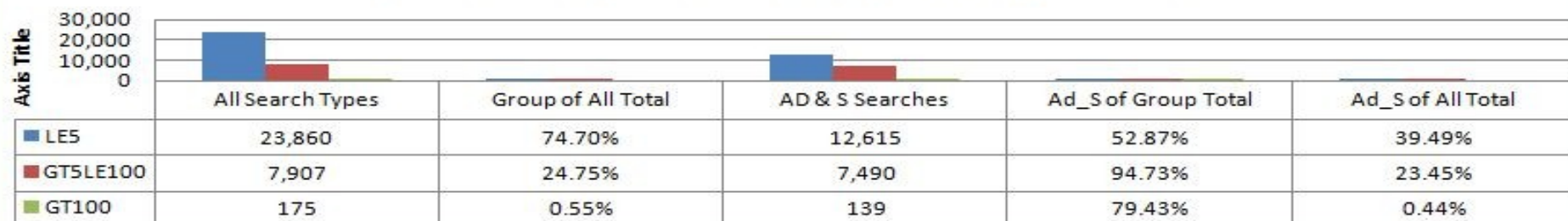
2008 UofL - Searches and User Groups - in-Library



2008 UofL - Searches and User Groups - on-Campus

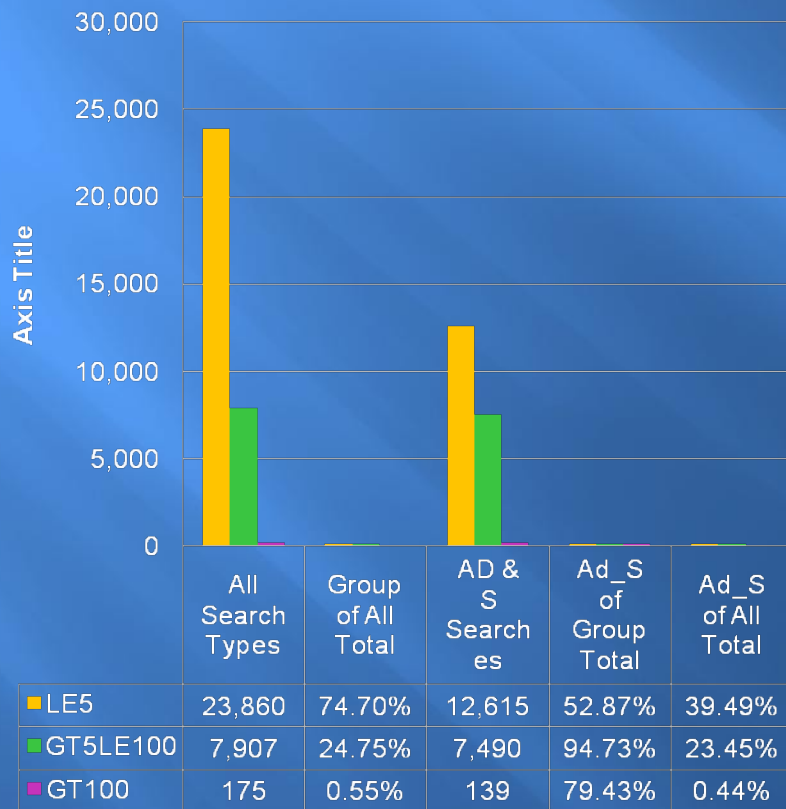


2008 UofL - Searches and User Groups - Remote

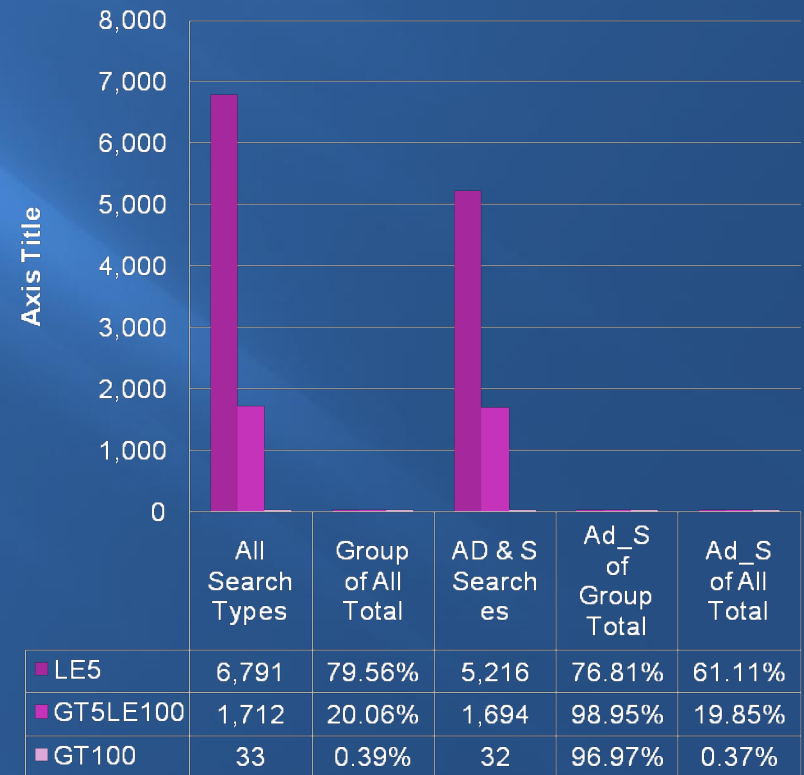


Summary of Remote Users (IPs)

UofL – IPs Grouped by Searches per IP



WKU – IPs grouped by Searches per IP



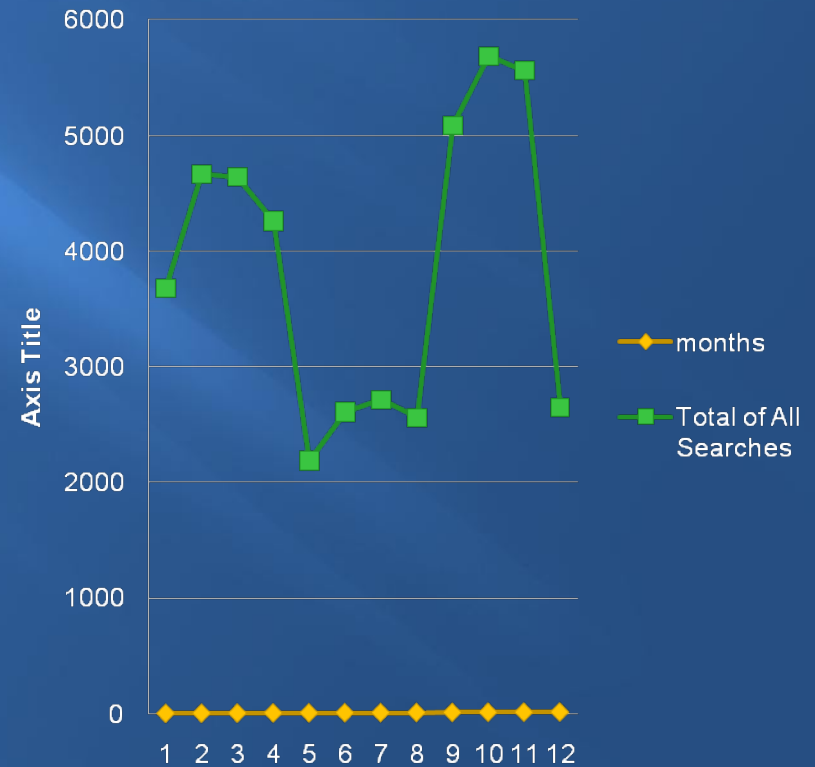
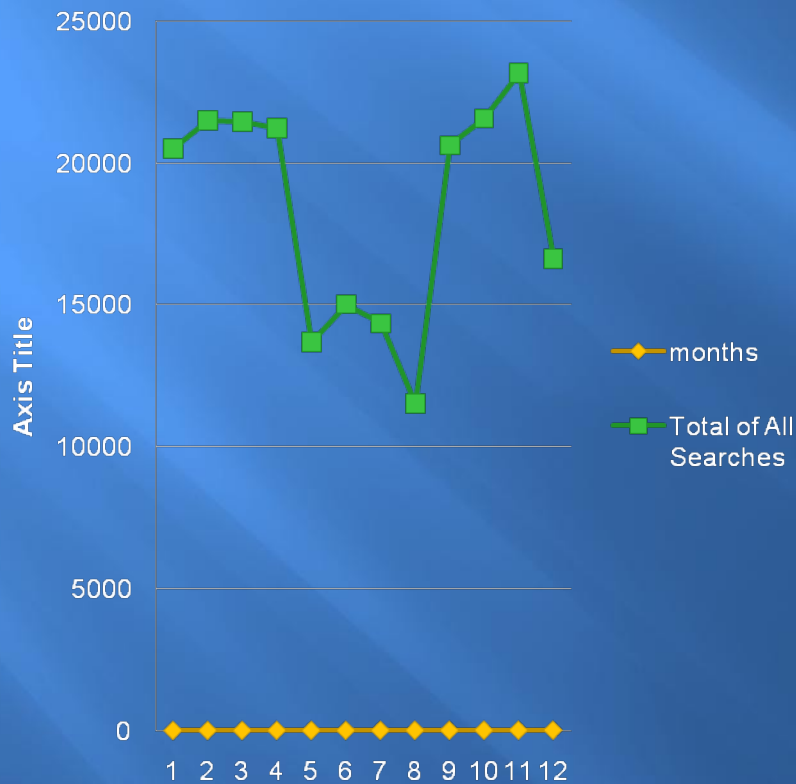
Summary of All Searches

by Remote Users (IPs) - 1

by Month

UofL

WKU



Summary of All Searches

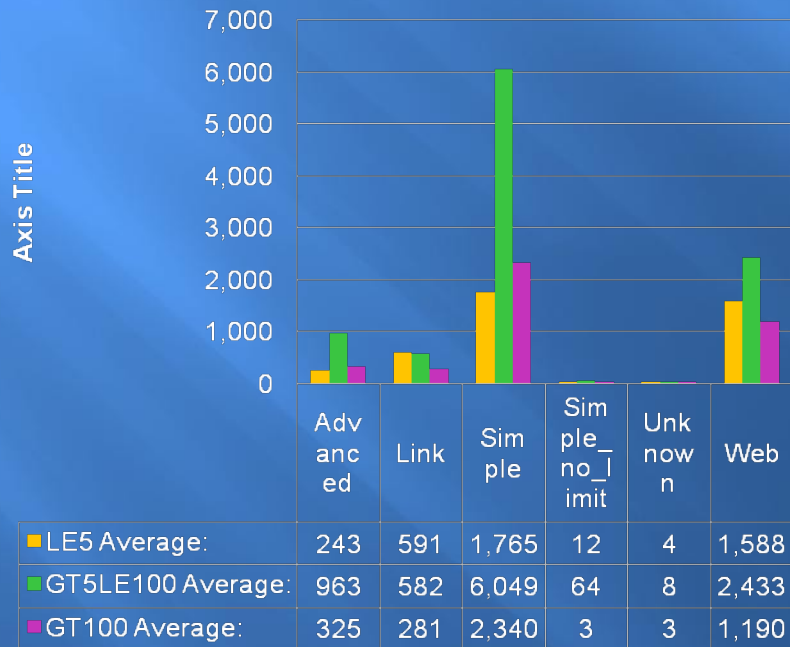
by Remote Users (IPs) - 2

Unique Search Queries

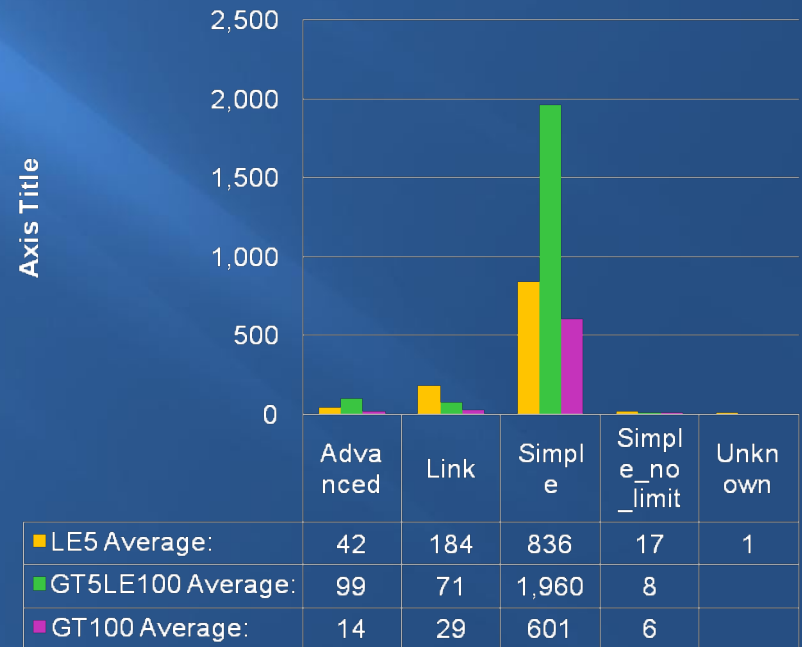
UofL: 236,902

WKU: 46,285

All Searches by Type
by Usage Group



All Searches by Type
by Usage Group



Data Source for This Study

UofL			
months	Total	Advanced	Simple
Average	11,685	1,531	10,154
1	13,046	1720	11,326
2	13,672	1874	11,798
3	13,147	1740	11,407
4	12,982	1772	11,210
5	7,700	1126	6,574
6	9,580	1030	8,550
7	8,631	1114	7,517
8	7,408	993	6,415
9	14,098	1843	12,255
10	13,967	1938	12,029
11	14,957	2019	12,938
12	11,026	1197	9,829
Total	140,214	18,366	121,848

WKU			
months	Total	Guided	Simple
Average	3,551	154	3,397
1	3,415	104	3311
2	4,373	227	4146
3	4,299	241	4058
4	3,923	148	3775
5	1,936	45	1891
6	2,357	83	2274
7	2,414	92	2322
8	2,310	44	2266
9	4,717	260	4457
10	5,254	295	4959
11	5,220	232	4988
12	2,388	77	2311
Total	42,606	1,848	40,758

Data Analysis

- Conducting the data analysis
 - Create a crosstab query for each search type and option to be studied
 - Export the cross tabulated reports as Excel files for summary charts
 - Create report charts in Excel

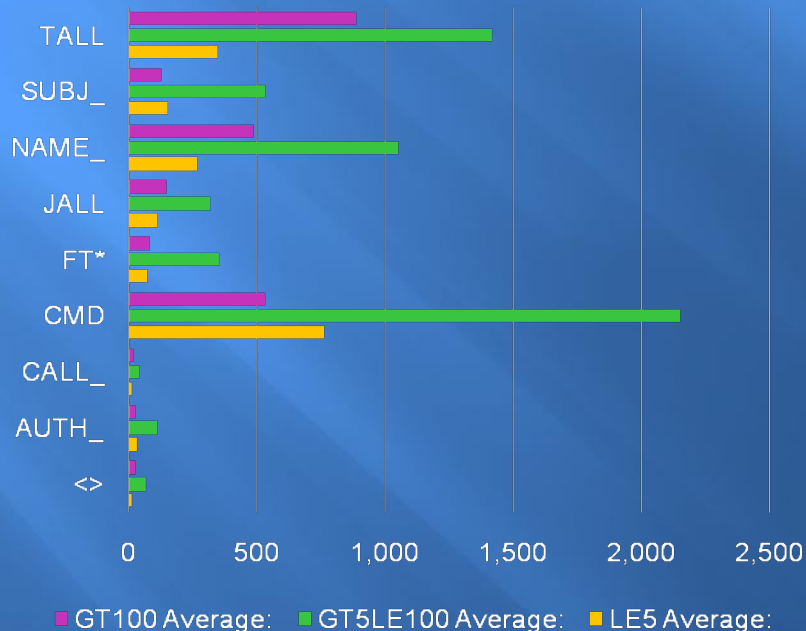
Results: Simple Search by Code - 1

Default Settings

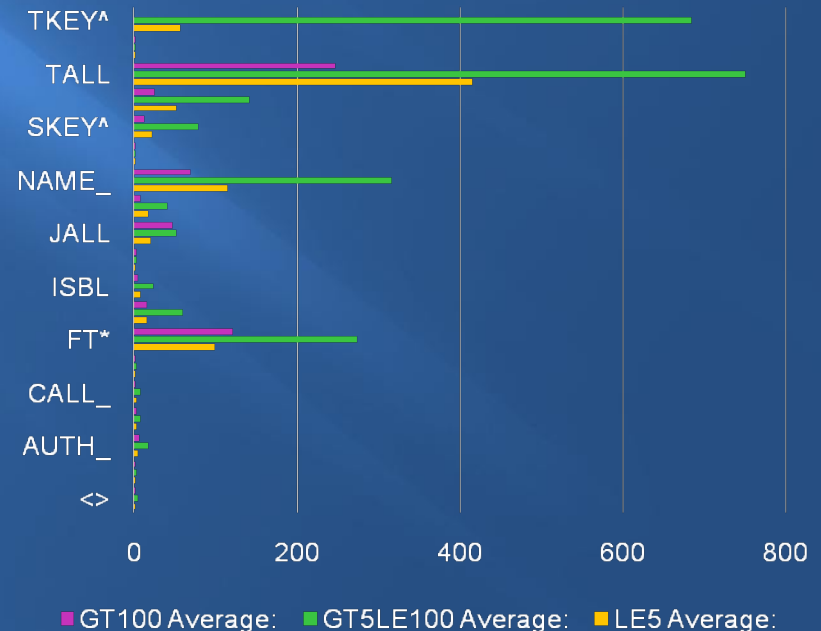
UofL
CMD – Keyword (and, or, not, “phrase”)

WKU
TALL - Title (Alphabetical)

Remote by Usage Group

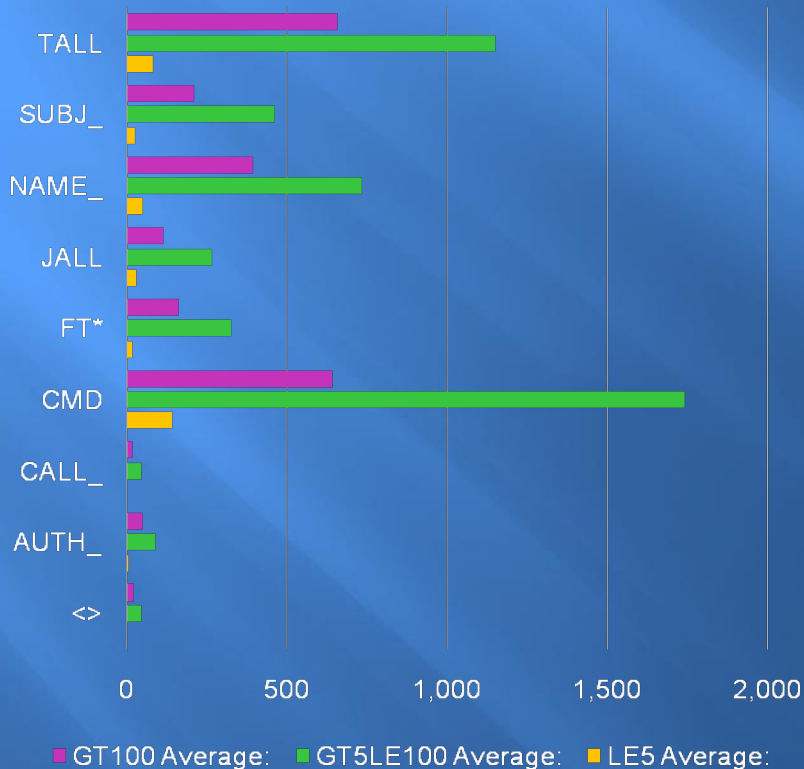


Remote by Usage Group

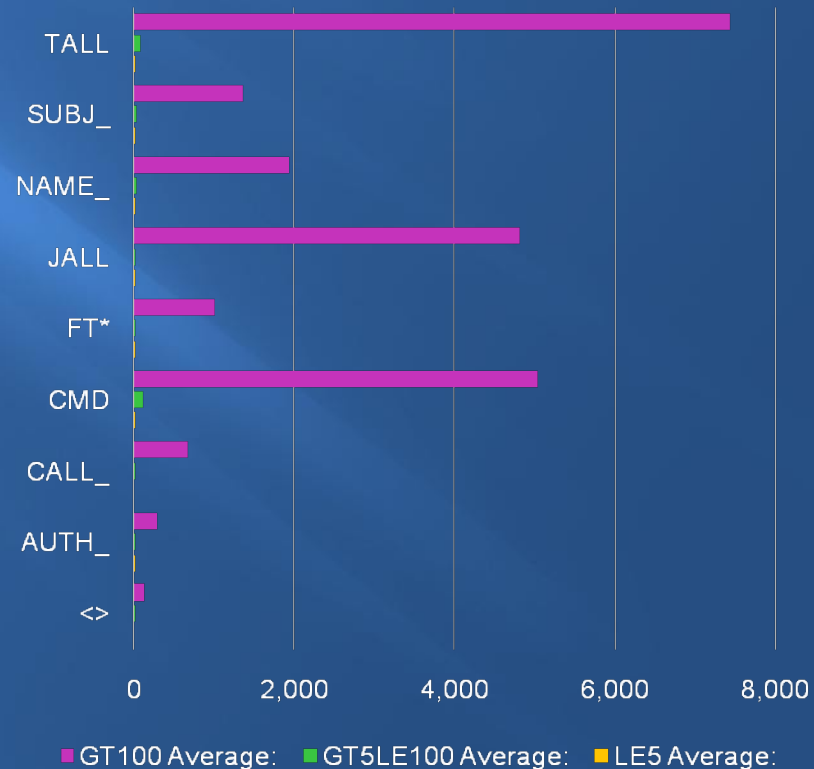


Results: Simple Search by Code - 2

UofL - on-Campus by Usage Group



UofL - in-Library by Usage Group



Results: Simple Search by Limit

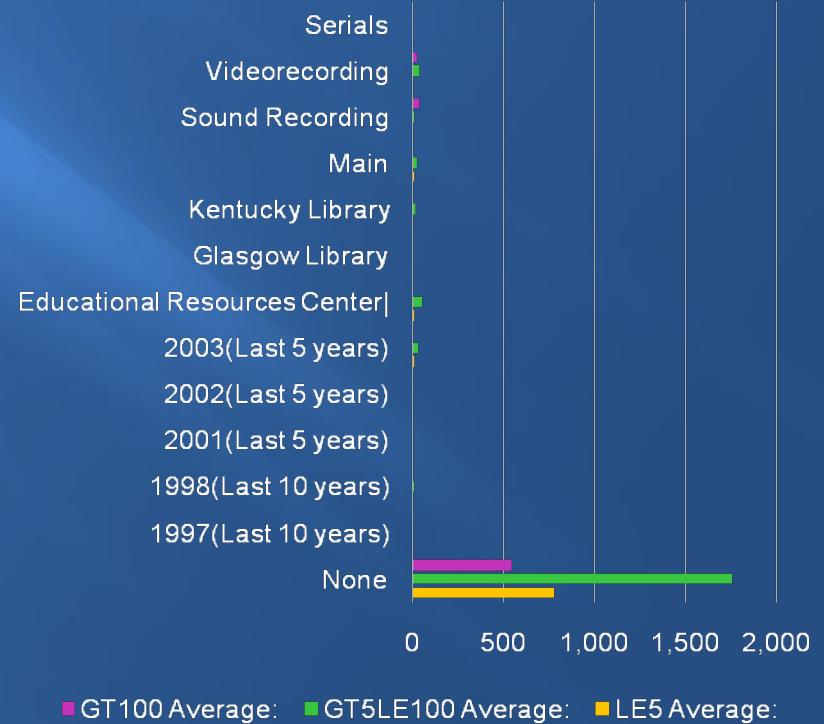
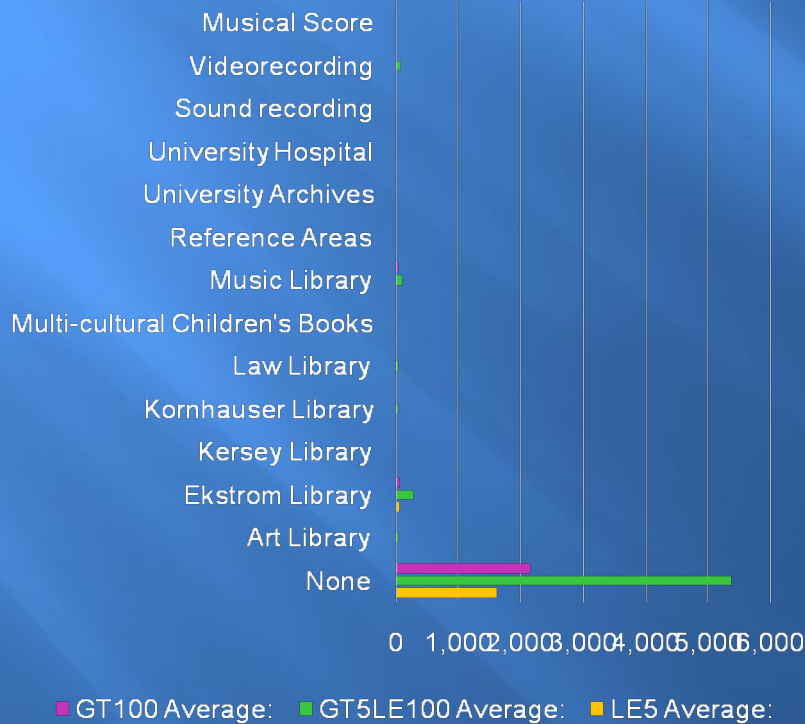
Default Settings

UofL: None

WKU: None

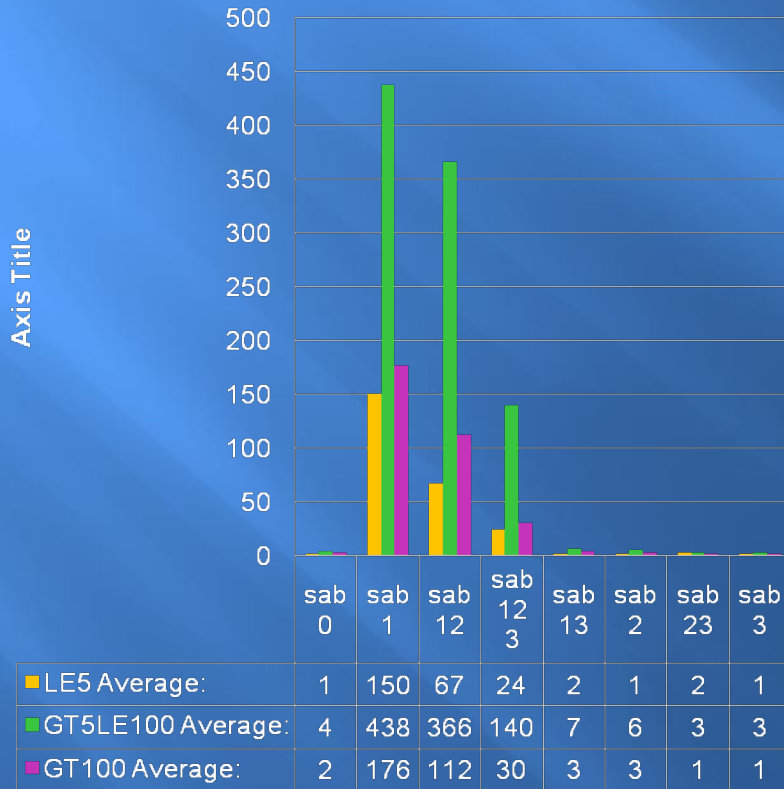
Remote by Usage Group

Remote by Usage Group

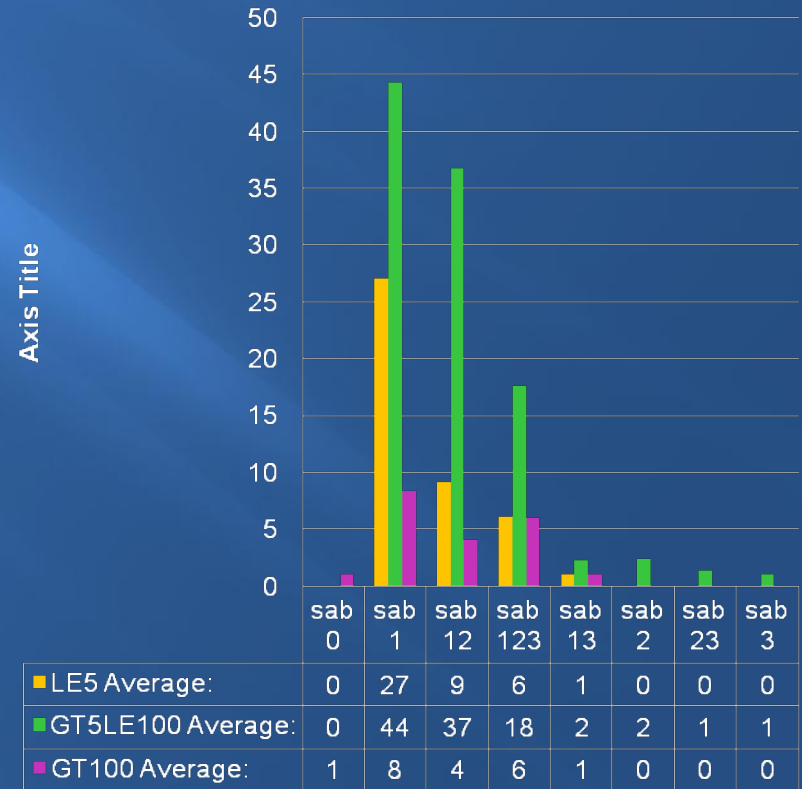


Results – Advanced Search by Search Box

UofL
Remote by Usage Group



WKU
Remote by Usage Group



Results – Advanced Search

used only 1st search box

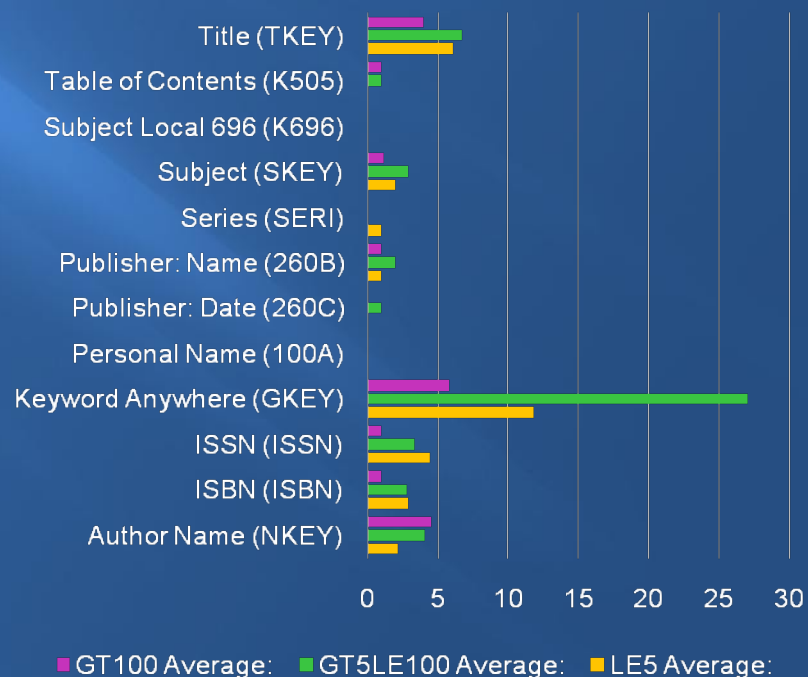
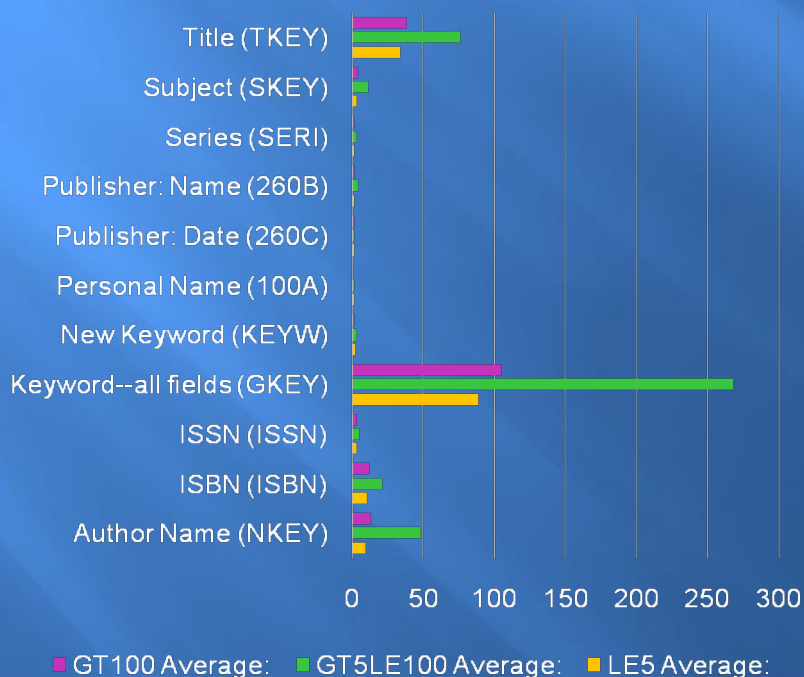
Default Settings

UofL: Keyword-all fields

WKU: Keyword Anywhere

Remote by Usage Group

Remote by Usage Group



Results – Advanced Search

used only 1st & 2nd search boxes

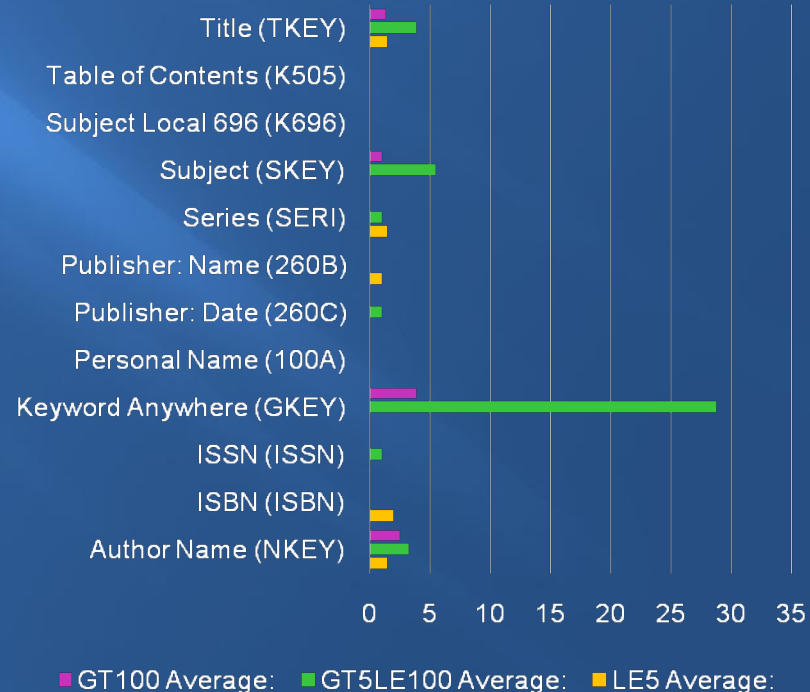
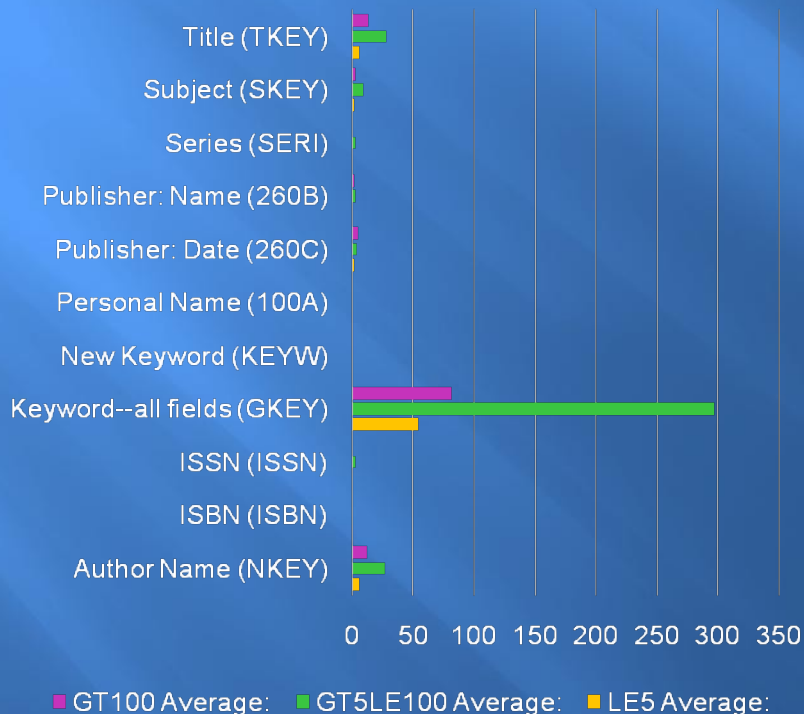
Default Settings

UofL: Keyword-all fields

WKU: Keyword Anywhere

Remote by Usage Group

Remote by Usage Group



Results – Advanced Search

used all three search boxes

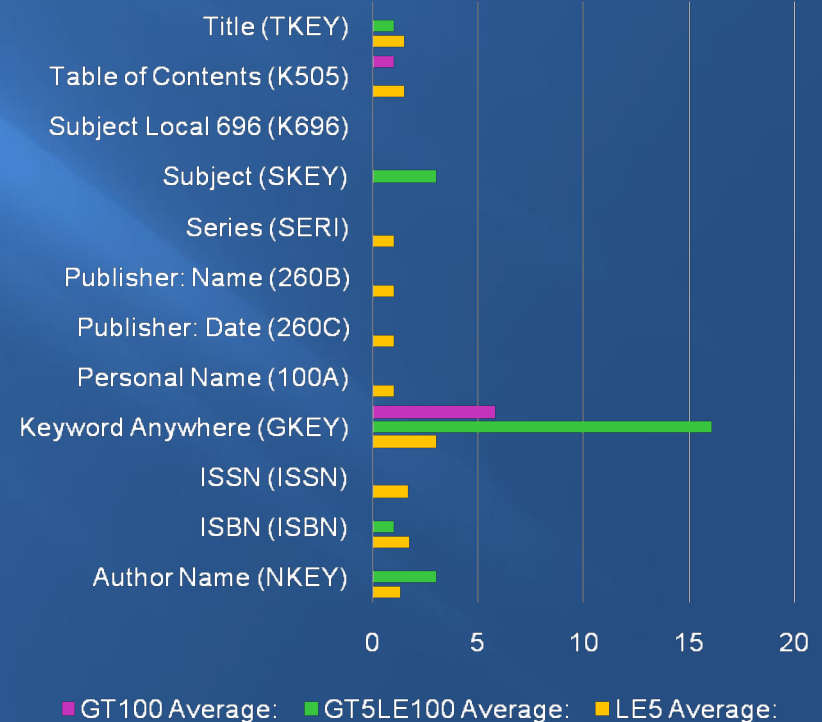
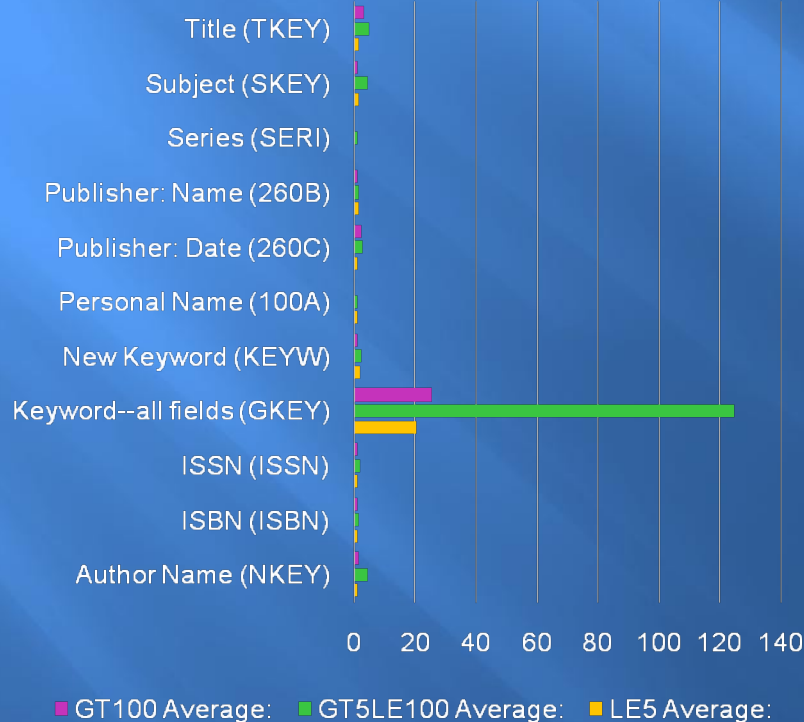
Default Settings

UofL: Keyword-all fields

WKU: Keyword Anywhere

Remote by Usage Group

Remote by Usage Group



Results – Advanced Search Box Grouping

Default Settings

UofL: AND & AND

WKU: AND & AND

UofL 2008

Advanced Search by Grouping - Usage Less Than or Equal to 5 per IP (LE5)

box	GROUP1	GROUP2	Total	1	2	3	4	5	6	7	8	9	10	11	12
sok12	AND with next set	AND with next set	27	21	6										
sok12	AND with next set	NOT with next set	1		1										
sok12	AND with next set	OR with next set	1												
sok12	NOT with next set	AND with next set	3	1											2
sok12	OR with next set	AND with next set	11		1	1			2				3	1	1
sok12	OR with next set	OR with next set	1												
sok123	AND with next set	AND with next set	205	24	21	22	25	27	16	10	12	23	22	25	3
sok123	AND with next set	OR with next set	5		2	1						1		1	
sok123	NOT with next set	AND with next set	1				1								
sok123	OR with next set	AND with next set	3	1	1								1		
sok123	OR with next set	OR with next set	3				1	1					2	1	

WKU 2008

Guided Keyword Search by Grouping - Usage Less Than or Equal to 5 per IP (LE5)

box	GROUP1	GROUP2	Total	1	2	3	4	5	6	7	8	9	10	11	12
sok12	AND with next set	AND with next set	34	2	3	21	10		6	6	2	4	4	11	13
sok12	AND with next set	OR with next set	1			1									
sok12	AND with next set	NOT with next set	4	1									1		2
sok12	OR with next set	AND with next set	1	1											
sok123	AND with next set	AND with next set	41	2	2	2	1	6	8	4	5	4	11	11	1
sok123	AND with next set	OR with next set	1				1								
sok123	OR with next set	AND with next set	1	1											
sok123	OR with next set	OR with next set	1				1								

Advanced Search by Grouping - Usage Greater than 5 and Less Than or Equal to 100 per IP (GT5LE100)

box	GROUP1	GROUP2	Total	1	2	3	4	5	6	7	8	9	10	11	12
sok12	AND with next set	AND with next set	132	34	26	48	48	48	24	24	24	24	41	41	41
sok12	AND with next set	NOT with next set	3	2	1					2					
sok12	AND with next set	OR with next set	7		2	1							1		
sok12	NOT with next set	AND with next set	14			2	3	1						5	1
sok12	NOT with next set	NOT with next set	2				1						1		
sok12	OR with next set	AND with next set	22	2	2	2	2						2	4	5
sok12	OR with next set	NOT with next set	1												
sok12	OR with next set	OR with next set	4	2	1										
sok123	AND with next set	AND with next set	154	134	20	125	187	45	56	33	64	53	140	177	105
sok123	AND with next set	NOT with next set	23	2		3	3			3			2	4	
sok123	AND with next set	OR with next set	24		4	2	1	3	2	3			4	2	
sok123	NOT with next set	AND with next set	4			2	2						4		
sok123	OR with next set	AND with next set	25	1	2	2	2	1			4			3	1
sok123	OR with next set	NOT with next set	1											1	
sok123	OR with next set	OR with next set	23	2	2	1	4		4		2	3	3	5	2

Guided Keyword Search by Grouping - Usage Greater than 5 and Less Than or Equal to 100 per IP (GT5LE100)

box	GROUP1	GROUP2	Total	1	2	3	4	5	6	7	8	9	10	11	12
sok12	AND with next set	AND with next set	407	24	70	50	30	6	19	23	7	51	20	23	14
sok12	AND with next set	NOT with next set	3		1	1								1	
sok12	AND with next set	OR with next set	2							2					
sok12	NOT with next set	AND with next set	3			3						5			
sok12	OR with next set	AND with next set	10		4	4	1	2					2	4	1
sok12	OR with next set	NOT with next set	2											2	
sok12	OR with next set	OR with next set	1												1
sok123	AND with next set	AND with next set	64	12	31	27	14		5	2	10	23	23	23	3
sok123	AND with next set	OR with next set	4	1	2	1								1	1
sok123	OR with next set	OR with next set	21	1		1							1	13	5

Advanced Search by Grouping - Usage Greater 100 per IP (GT100)

box	GROUP1	GROUP2	Total	1	2	3	4	5	6	7	8	9	10	11	12
sok12	AND with next set	AND with next set	131	125	6	125	125	22	26	24	22	23	157	173	94
sok12	AND with next set	OR with next set	4	4			1								
sok12	AND with next set	NOT with next set	3	2	1										

Guided Keyword Search by Grouping - Usage Greater 100 per IP (GT100)

box	GROUP1	GROUP2	Total	1	2	3	4	5	6	7	8	9	10	11	12
sok12	AND with next set	AND with next set	43	3	5	3	4		3	2	4	3	3		
sok12	OR with next set	AND with next set	1				1								
sok123	AND with next set	AND with next set	23	1	1	1		3				21	3		
sok123	AND with next set	OR with next set	1				1								

Results – Advanced Search Boolean Options

Default Settings

UofL: all of these

WKU: as a phrase

UofL 2008

Advanced Search by Boolean - Usage Less Than or Equal to 5 per IP (LES)																
box	BOO1	BOO2	BOO3	Total	1	2	3	4	5	6	7	8	9	10	11	12
bob 2	all of these	as a phrase	as a phrase	749	62	30	63	76	70	65	43	52	34	74	73	42
bob 2	all of these	as a phrase	as a phrase	3	1											
bob 2	all of these	as a phrase	as a phrase	2												
bob 2	all of these	as a phrase	as a phrase	9	1		1						1			2
bob 2	as a phrase	as a phrase	as a phrase	6			1									
bob 2	as a phrase	as a phrase	as a phrase	2							1					
bob 2	as a phrase	as a phrase	as a phrase	12	4									4		2
bob 2	as a phrase	as a phrase	as a phrase	6	3								1			
bob 2	as a phrase	as a phrase	as a phrase	2				1	1							
bob 2	as a phrase	as a phrase	as a phrase	0	1	4										
bob 20	all of these	as a phrase	as a phrase	262	25	3	23	31	20	16	11	13	27	23	24	16
bob 20	all of these	as a phrase	as a phrase													
bob 20	all of these	as a phrase	as a phrase													
bob 20	as a phrase	as a phrase	as a phrase	2										2		
bob 20	as a phrase	as a phrase	as a phrase	16				1						2		2
bob 20	as a phrase	as a phrase	as a phrase													

UofL 2008

Advanced Search by Boolean - Usage Greater than 5 and Less Than or Equal to 100 per IP (GTSLC100)																
box	BOO1	BOO2	BOO3	Total	1	2	3	4	5	6	7	8	9	10	11	12
bob 2	all of these	as a phrase	as a phrase	4076	303	436	313	300	279	237	214	173	400	484	541	239
bob 2	all of these	as a phrase	as a phrase	6						2						
bob 2	all of these	as a phrase	as a phrase	15	3		1	1						5		2
bob 2	all of these	as a phrase	as a phrase	9												
bob 2	all of these	as a phrase	as a phrase	20			7	1	2	4		4			4	
bob 2	all of these	as a phrase	as a phrase	44	3		11	1	2	2			3	4		5
bob 2	as a phrase	as a phrase	as a phrase	26	3	4	3					3	1		2	3
bob 2	as a phrase	as a phrase	as a phrase	32	2	1	3		2	3			3	3	1	
bob 2	as a phrase	as a phrase	as a phrase													
bob 2	as a phrase	as a phrase	as a phrase	4												
bob 2	as a phrase	as a phrase	as a phrase	74	5	0	1	0	3	2	5	4	0	0	0	6
bob 2	as a phrase	as a phrase	as a phrase	6	1		1									
bob 2	as a phrase	as a phrase	as a phrase	6												
bob 2	as a phrase	as a phrase	as a phrase	60		3	4	2	9	1	26	5		4		
bob 2	as a phrase	as a phrase	as a phrase	2	1											
bob 20	all of these	as a phrase	as a phrase	1406	102	136	163	130	43	64	105	73	162	175	107	74
bob 20	all of these	as a phrase	as a phrase													

WKU 2008

Guided Keyword Search by Boolean - Usage Less Than or Equal to 5 per IP (LES)																
box	BOO1	BOO2	BOO3	Total	1	2	3	4	5	6	7	8	9	10	11	12
bob 12	all of these	as a phrase	as a phrase	3												
bob 12	all of these	as a phrase	as a phrase	3												
bob 12	as a phrase	as a phrase	as a phrase	7	1			6								1
bob 12	as a phrase	as a phrase	as a phrase	2												
bob 12	as a phrase	as a phrase	as a phrase	7	1						1			1		
bob 12	as a phrase	as a phrase	as a phrase	1												
bob 12	as a phrase	as a phrase	as a phrase	01	4	7	10	2	4	5	2	4	0	3	11	
bob 12	as a phrase	as a phrase	as a phrase	1												
bob 12	as a phrase	as a phrase	as a phrase	46		7	2			4	5		0	3	11	
bob 12	as a phrase	as a phrase	as a phrase	2				2								
bob 12	as a phrase	as a phrase	as a phrase	21	3		2		1	1	3	3	2	2	3	

WKU 2008

Guided Keyword Search by Boolean - Usage Greater than 5 and Less Than or Equal to 100 per IP (GTSLC100)																
box	BOO1	BOO2	BOO3	Total	1	2	3	4	5	6	7	8	9	10	11	12
bob 12	all of these	as a phrase	as a phrase	3												
bob 12	all of these	as a phrase	as a phrase	29	2	2			6					1	3	
bob 12	all of these	as a phrase	as a phrase	3						1					1	
bob 12	all of these	as a phrase	as a phrase	11		2		2			1			3	1	
bob 12	as a phrase	as a phrase	as a phrase	7												
bob 12	as a phrase	as a phrase	as a phrase	7				6								1
bob 12	as a phrase	as a phrase	as a phrase	1												
bob 12	as a phrase	as a phrase	as a phrase	3												
bob 12	as a phrase	as a phrase	as a phrase	372	22	70	70	24	7	10	17	0	50	40	33	10
bob 12	as a phrase	as a phrase	as a phrase	5					1						3	
bob 12	as a phrase	as a phrase	as a phrase	1												
bob 12	as a phrase	as a phrase	as a phrase	3	1				1							
bob 12	as a phrase	as a phrase	as a phrase	11	2	2			2							1
bob 12	as a phrase	as a phrase	as a phrase	2												
bob 12	as a phrase	as a phrase	as a phrase	2	1				3							2
bob 12	as a phrase	as a phrase	as a phrase	1										1		
bob 12	as a phrase	as a phrase	as a phrase	162	7	23	26	5		1	5		0	48	21	6

Findings

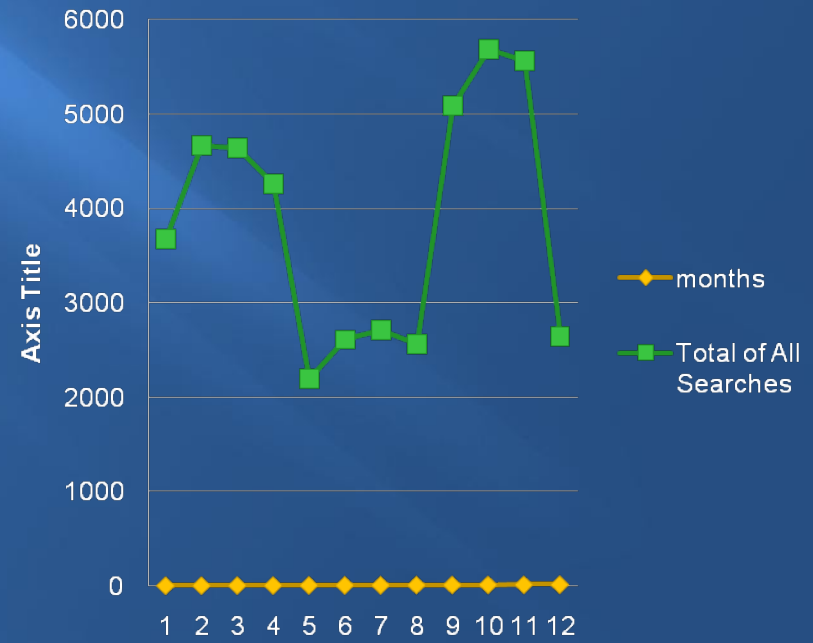
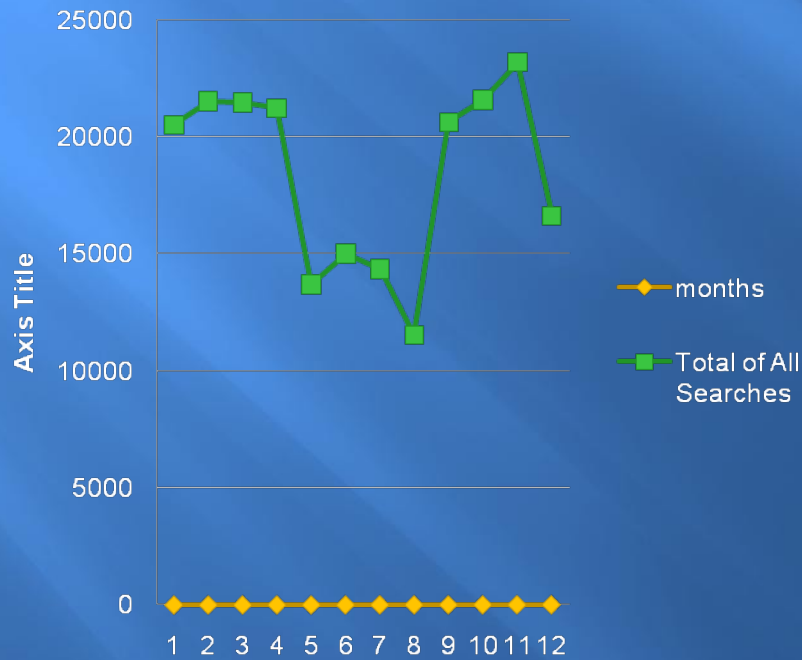
- The overall results for both UofL and WKU are the same:
 - Simple Search is the most heavily used interface
 - Default options (except one for WKU) are used the most, no matter what they are, quick limits in Simple Search, or Boolean operators and box grouping options in Advanced/Guided Keyword Search

Other Findings - 1

The usage by month shows that the traffic of OPAC use through out the year seem to follow the semester's schedules.

UofL

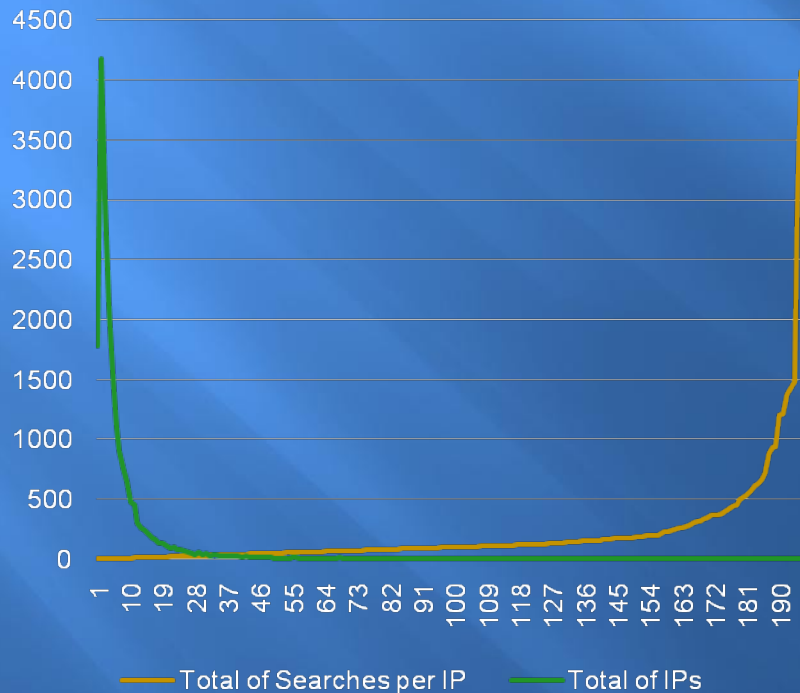
WKU



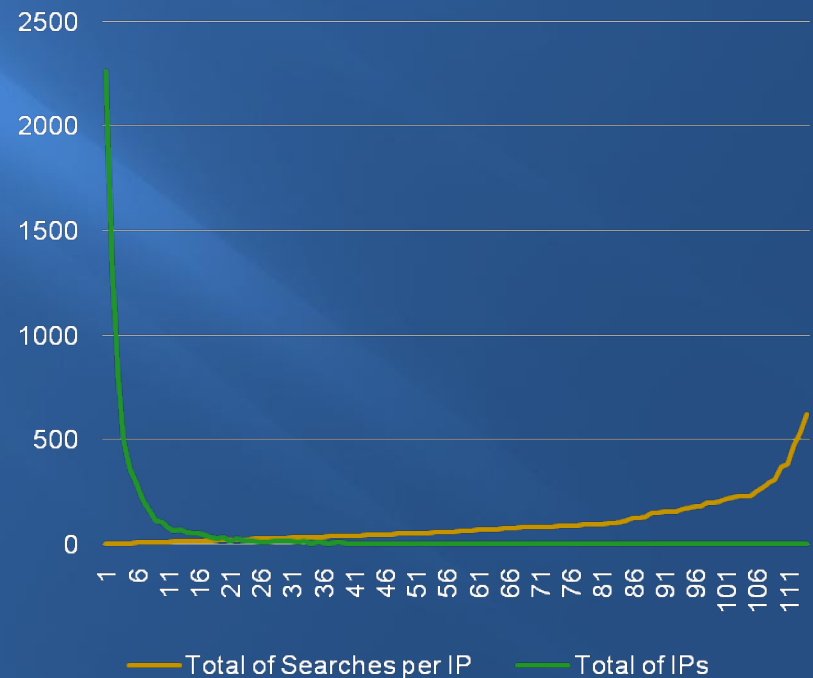
Other Findings - 2

The majority of remote users rarely or occasionally use OPAC

UofL - Searches per IP



WKU - Searches per IP



Other Findings - 3

- The next two mostly used options for UofL are Title and Author, and for WKU are Keyword or Title Keyword, and Author
- For UofL's in-library user group, Simple Search by "Title" shows higher usage than by the default, "Keyword"

Conclusions

- Users seem to use what is provided. No matter what the defaults are, “Keyword” or “Title,” “All of these” or “As a phrase”.
- Users seem to prefer simple search whether the interface is one search box (Simple Search) or multiple search boxes (Advanced Search)
- The most frequently used options are keyword, title and author
- For the in-library group of UofL (including library staff), search by “Title” shows higher usage than by “Keyword”
- The majority of OPAC users (IPs) are occasional users
- The traffic of OPAC use follows the semester’s schedules
- The number of IPs in the transaction logs are not equivalent to the number of actual users

Things Learned and Future Studies

- Overall picture of the OPAC Search
- What to look for in statistics application configuration
- Session level and term level analysis
- TLA vs. Web Analytics

References

- [1] Jansen, B. J. (2006). Search log analysis: What it is, what's been done, how to do it. *Library & Information Science Research*, 28, 407-432.
- [2] Jansen, B. J, I. Taksa, & A. Spink. (2009). Research and Methodological Foundations of Transaction Log Analysis. In B.J. Jansen, Spink A. & Taksa I. (Eds.), *Handbook of Research on Web Log Analysis* (pp.1-16). Hershey, PA : Information Science Reference.
- [3] Villén-Rueda, L., Senso, J. A., & Moya-Angón, F. de. (2007). The Use of OPAC in a large academic library: a transactional log analysis study of subject searching. *The Journal of Academic Librarianship*, 33(3), 327-337.
- [4] Moukdad, H. & Large, A. (2001). Users' perceptions of the Web as revealed by transaction log analysis. *Online Information Review*, 25(6), 349-358.
- [5] Rubin, J. H. (2004). Workshop - Log Analysis Pays Off - Analyzing your Web logs can help you improve your site's design and efficiency. Here's how to make the best use of that data. *Network Computing*, 15 (18), 76-78.
- [6] Huntington, P., Nicholas, D. & Jamali, H. R. (2007). Site navigation and its impact on the content viewed by the virtual scholar: a deep log analysis. *Journal of Information Science*, 33(5) , 598-610.

References (cont.)

- [7] Moulaison, H. (2008). OPAC Queries at a Medium-Sized Academic Library: A transaction log analysis. *LRTS*, 52(4), 230-237.
- [8] Wang, P., Berry, M., & Yang, Y. (2003). Mining longitudinal Web queries: Trends and patterns. *Journal of the American Society for Information Science and Technology*, 54 (8), 743–758.
- [9] Wallace, M. A. (1993). How do patrons search the online catalog when no-one's looking: Transaction log analysis and implications for bibliographic instruction and system design *RQ*, 33, 239-252.
- [10] Malliari, A . & Kyriaki-Manessi, D. (2007). Users' behaviour patterns in academic libraries' OPACs: a multivariate statistical analysis. *New Library World*, 108(3/4), 107-122. doi: 10.1108/03074800710735311
- [11] Knieval, J. E., et. al., (2009). Does Interface Design Influence Catalog Use? A Case Study. *College & Research Libraries*, 70(5), 446-458.
- [12] Kurth, M. (1993). The limits and limitation of transaction log analysis. *Library Hi Tech*, 11(2), 98-104.

Acknowledgements

- I would like to thank Calvin Miracle for his assistance in programming the extraction Perl script and extracting the logs
- I would like to thank Nelda Sims for sharing WKU's logs and providing me some related background information

Questions?

Thank you!

Weiling Liu
Director, the Office of Libraries Technology
University Libraries
University of Louisville
w.liu@louisville.edu