

2009

Location, location, location: A transaction comparison of catalog searches originating from the library homepage and Aleph

Jimmy Ghaphery

Virginia Commonwealth University, jghapher@vcu.edu


Thomas McNulty

Virginia Commonwealth University, tmcnulty@vcu.edu

Susan Teague Rector

Truven Health Analytics, seteague@gmail.com

Follow this and additional works at: http://scholarscompass.vcu.edu/libraries_present

 Part of the [Graphics and Human Computer Interfaces Commons](#), and the [Library and Information Science Commons](#)

Downloaded from

http://scholarscompass.vcu.edu/libraries_present/4

This Presentation is brought to you for free and open access by the VCU Libraries at VCU Scholars Compass. It has been accepted for inclusion in VCU Libraries Faculty and Staff Presentations by an authorized administrator of VCU Scholars Compass. For more information, please contact libcompass@vcu.edu.

Location, location, location:

A transaction comparison of catalog searches originating from the library homepage and Aleph

Jimmy Ghaphery, Thomas McNulty,
Susan Teague–Rector
VCU Libraries

ELUNA Annual Conference May 7,
2009

Virginia Commonwealth University

- Largest university in Virginia
- Two downtown campuses here in Richmond
- Enrolls more than 32,000 students
- Ranks among the top universities in the country in research funding



VCU Libraries

- 1.9 million volumes and 16,790 periodical subscriptions
- Aleph version 18
- Other systems: Metalib, SFX, ILLiad, CONTENTdm, DSpace
- <http://www.library.vcu.edu>

Relevance

- Visited 111 ARL Library websites on April 10, 2009
- Findings similar to those reported from Fall 2007 survey in Against The Grain 20(3), June 2008
- Search is common across academic library homepages

ARL website visits: April 10, 2009

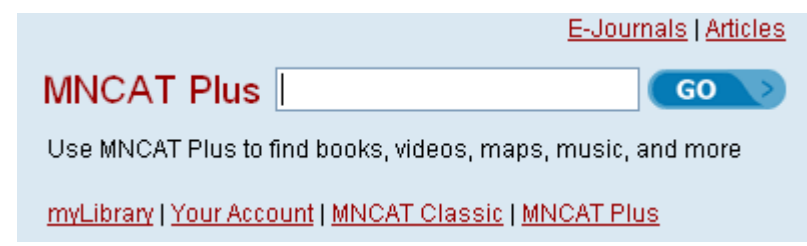
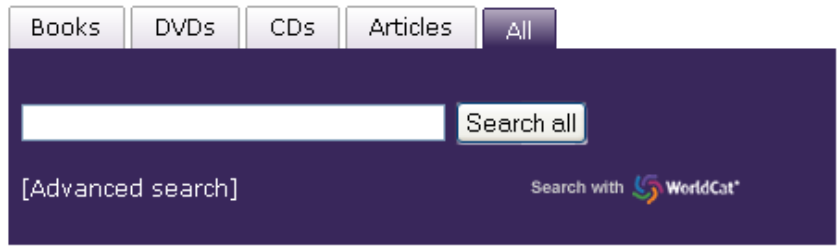
- 94% of sites had a search box
- 85% of sites queried library catalog (including Primo, Encore, WorldCat Local)
- Average number of search targets: 3.45
- Sixteen sites (14%) offered only one search target

ARL website visits: April 10, 2009

- Tabbed search approach most common
- Some momentum toward single searching since 2007
- Tension and rapid development of library search

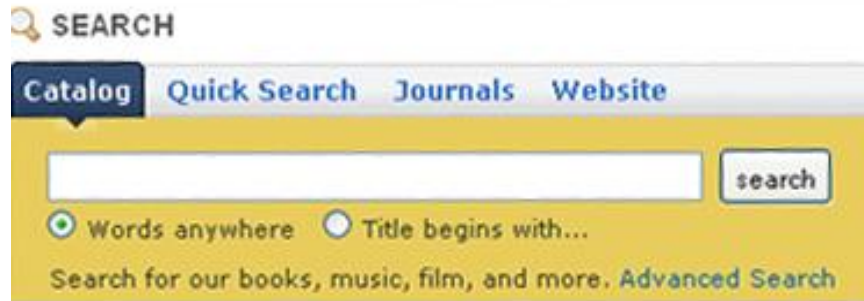


UW WorldCat: Search UW Libraries and beyond 



Our Question

Is there a difference in user search behavior when searching the library catalog from an external search box on the library homepage as opposed to searches in the native interface?



SEARCH

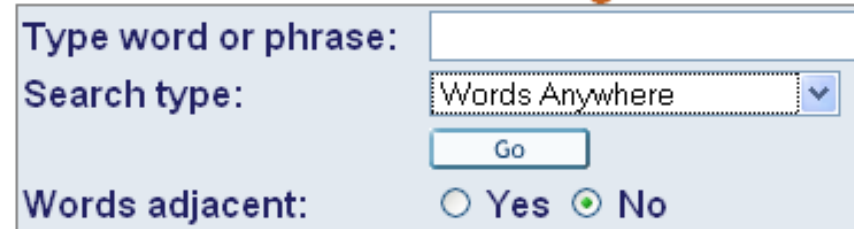
Catalog Quick Search Journals Website

search

Words anywhere Title begins with...

Search for our books, music, film, and more. [Advanced Search](#)

Basic Search of Full Catalog



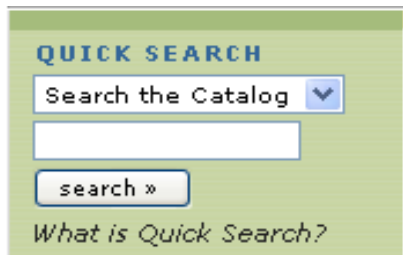
Type word or phrase:

Search type: ▾

Words adjacent: Yes No

2007–2009 Research

- Desire to understand search target usage from homepage drove search log creation
- Search terms are collected using a PHP script & MySQL database
- Data being logged:
 - Location
 - Search Type (Catalog, Articles, Journals, Site)
 - Search Term
 - Referrer
 - Timestamp
 - Scan ID (Words Anywhere, Title Begins with for Catalog searches only)
- Continued to log searches through redesign in 2008 and into 2009



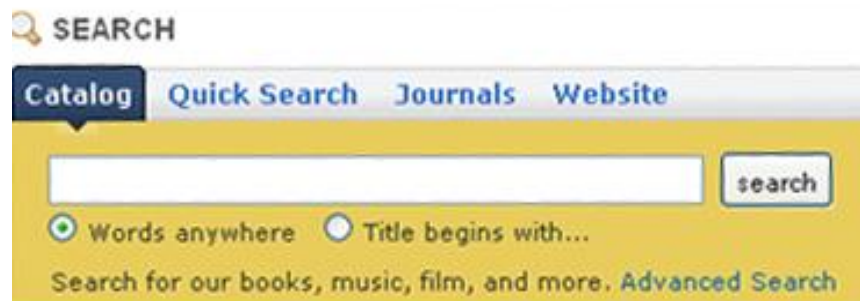
QUICK SEARCH

Search the Catalog ▾

search »

What is Quick Search?

2007



SEARCH

Catalog Quick Search Journals Website

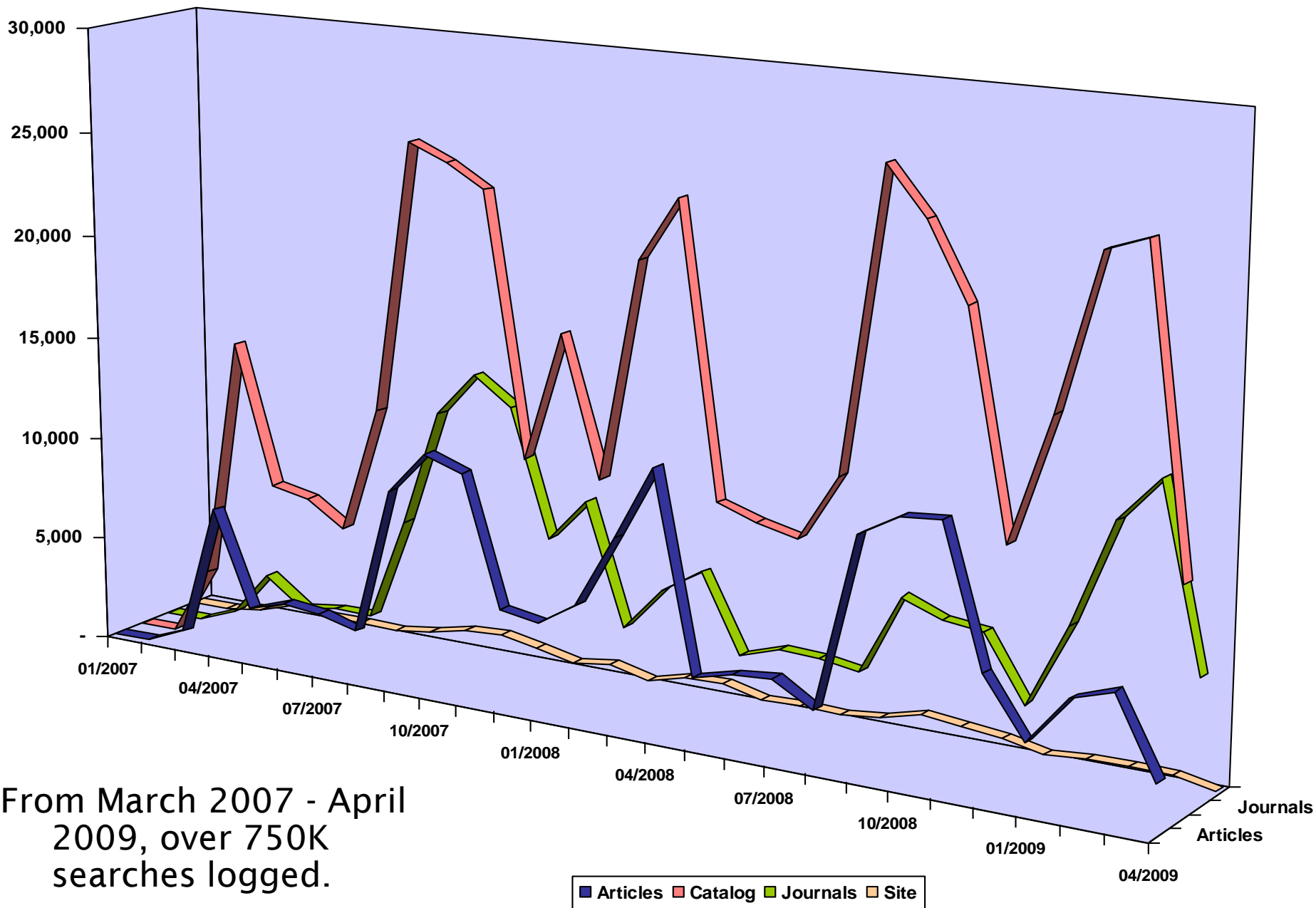
search

Words anywhere Title begins with...

Search for our books, music, film, and more. [Advanced Search](#)

2008/09

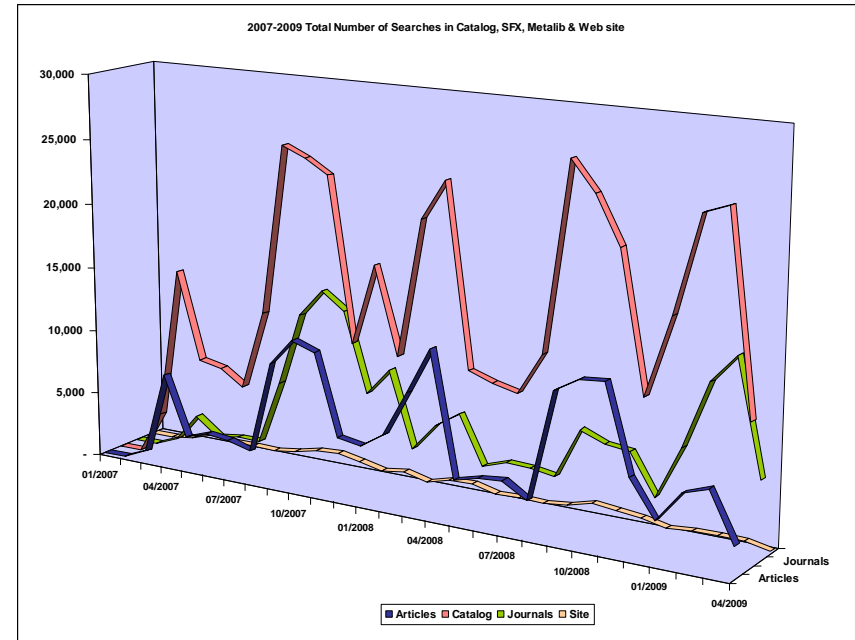
2007-2009 Total Number of Searches in Catalog, SFX, Metalib & Web site



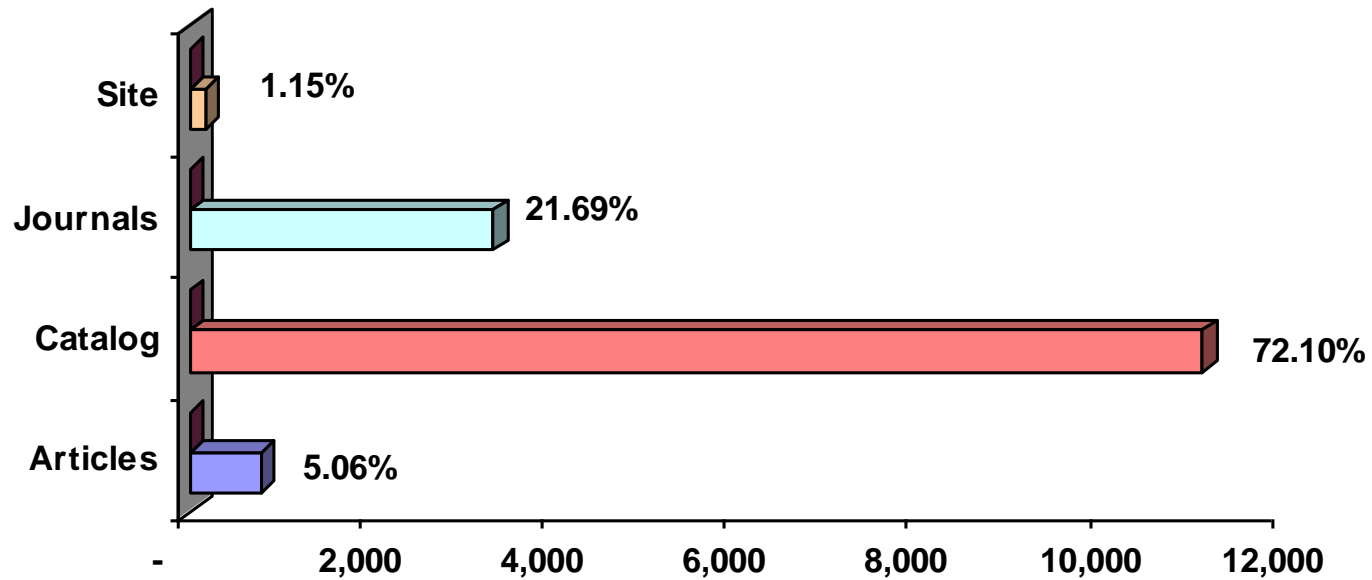
From March 2007 - April 2009, over 750K searches logged.

2009 Sample

- Selected 2 weeks - 2/16/2009
- 3/1/2009 to further investigate
- Selected data from the search Web logs and from Aleph z69 table
- Analyzed data for this time period to look at:
 - Number of searches
 - Location
 - Search Type (Words Anywhere, Title Begins With)



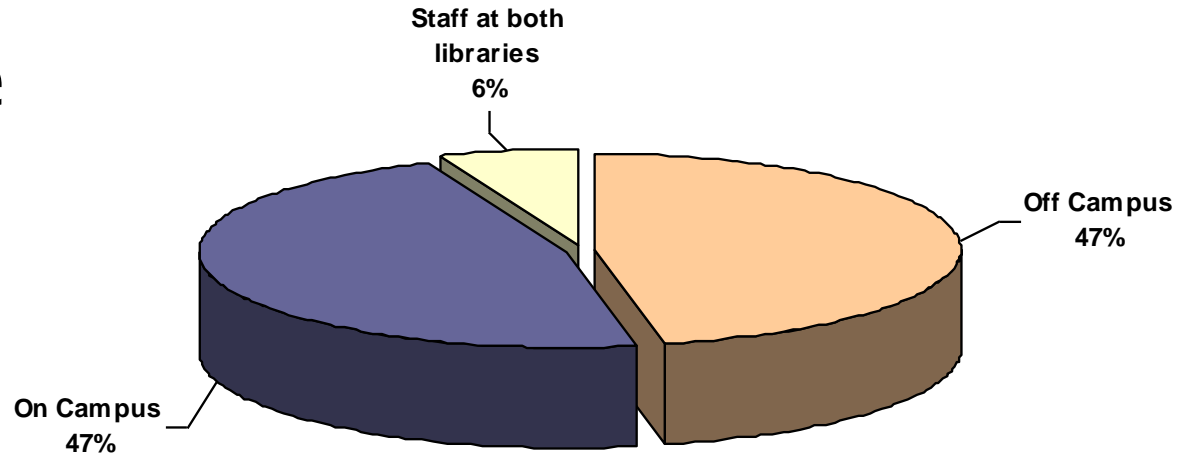
Web Results



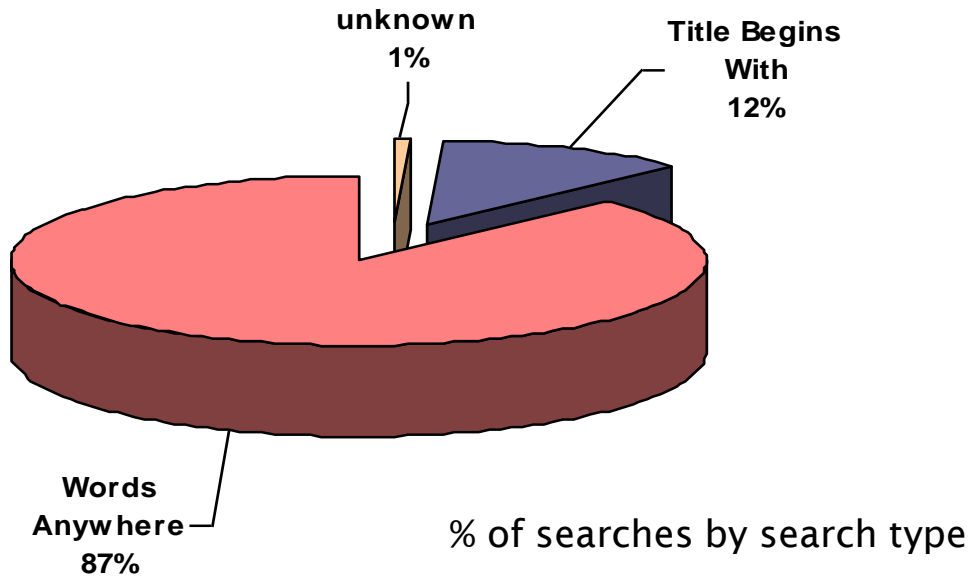
Articles	Catalog	Journals	Site
5.06%	72.10%	21.69%	1.15%
782	11,133	3,349	178

% of searches from the Homepage to the Catalog

Web Results for Sample Timeframe

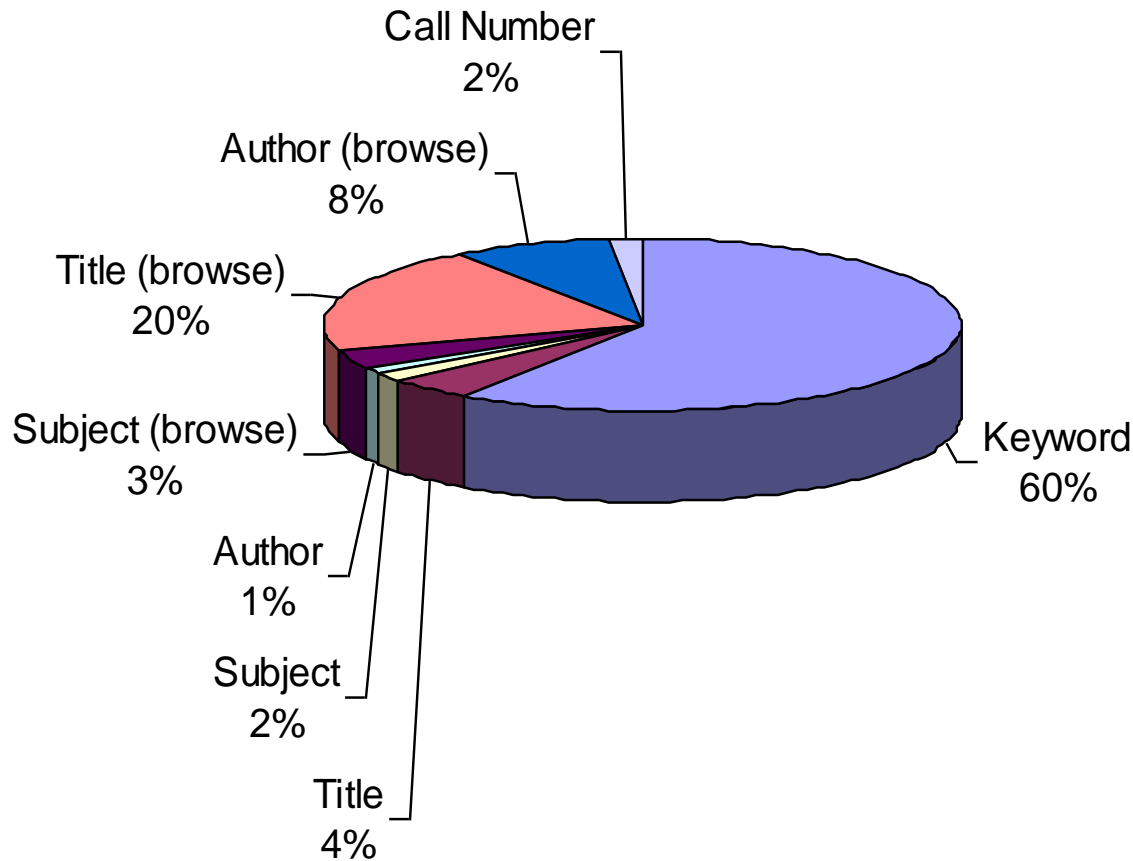


% of searches by location



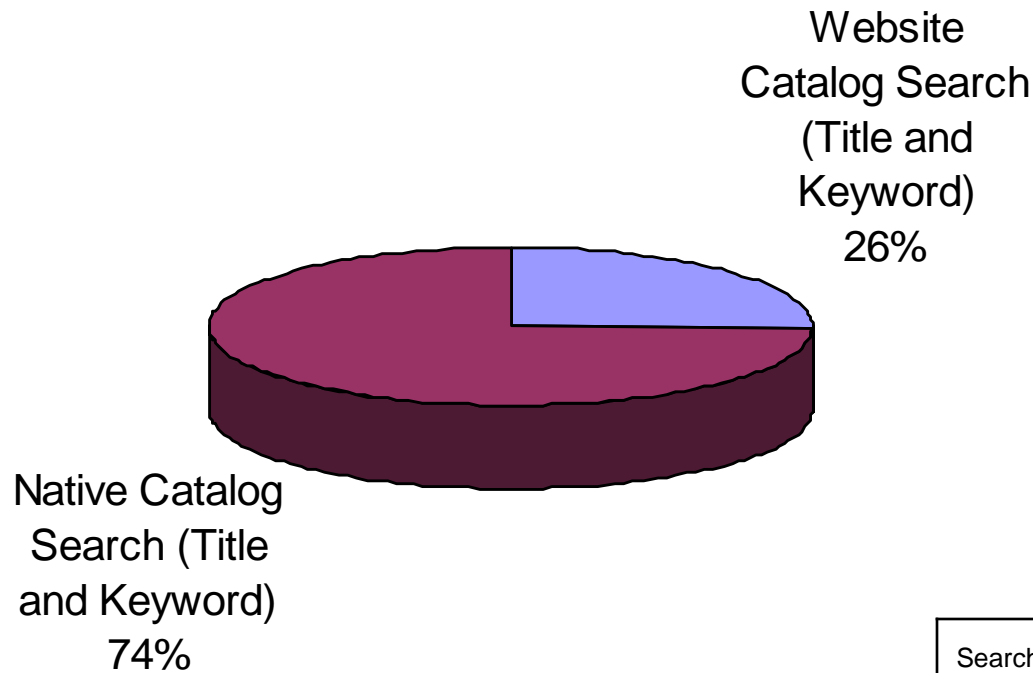
% of searches by search type

Catalog Results



Search Type	Count
Keyword	22,493
Title	1,677
Subject	609
Author	481
Subject (browse)	1,263
Title (browse)	7,670
Author (browse)	2,949
Call Number	682
Total	3,7824

Catalog vs. Web Results



Search type	
Catalog Keyword	22,493
Catalog Title begins with	7670
Website Keyword	9754
Website Title begins with	1300
Total	41,217

Z69 Aleph Table

- Web OPAC Events (KB 5872: Z69 Table What is it how is it cleaned?)
- Z69_TIME_STAMP, Z69_SESSION_ID, Z69_CLIENT_ADDRESS, Z69_EVENT_TYPE
- Z69_SEARCH_BASE, Z69_SEARCH_QUERY, Z69_NO_HITS
- Z69_SCAN_BASE, Z69_SCAN_CODE, Z69_SCAN_TEXT
- Z69_BOR_ID, Z69_BOR_STATUS
- Z69_SDI_REQUEST , Z69_SDI_INTERVAL, Z69_SDI_INTERVAL_TYPE
- Z69_SAVE_TO, Z69_SAVE_FORMAT, Z69_DOC_NUMBER

Sample SQL for Catalog Scans

```
set head off;
select Z69_SCAN_code, count(Z69_SCAN_code) from
  z69 where (
z69_time_stamp like '20090216%' OR
z69_time_stamp like '20090217%' OR
z69_time_stamp like '20090218%' OR
  z69_time_stamp like '20090219%' OR
  z69_time_stamp like '2009022%' OR
  z69_time_stamp like '20090301%')
and z69_scan_base like 'VCU01PUB' and
  Z69_EVENT_TYPE like '29'
group by Z69_SCAN_code;
```

Results

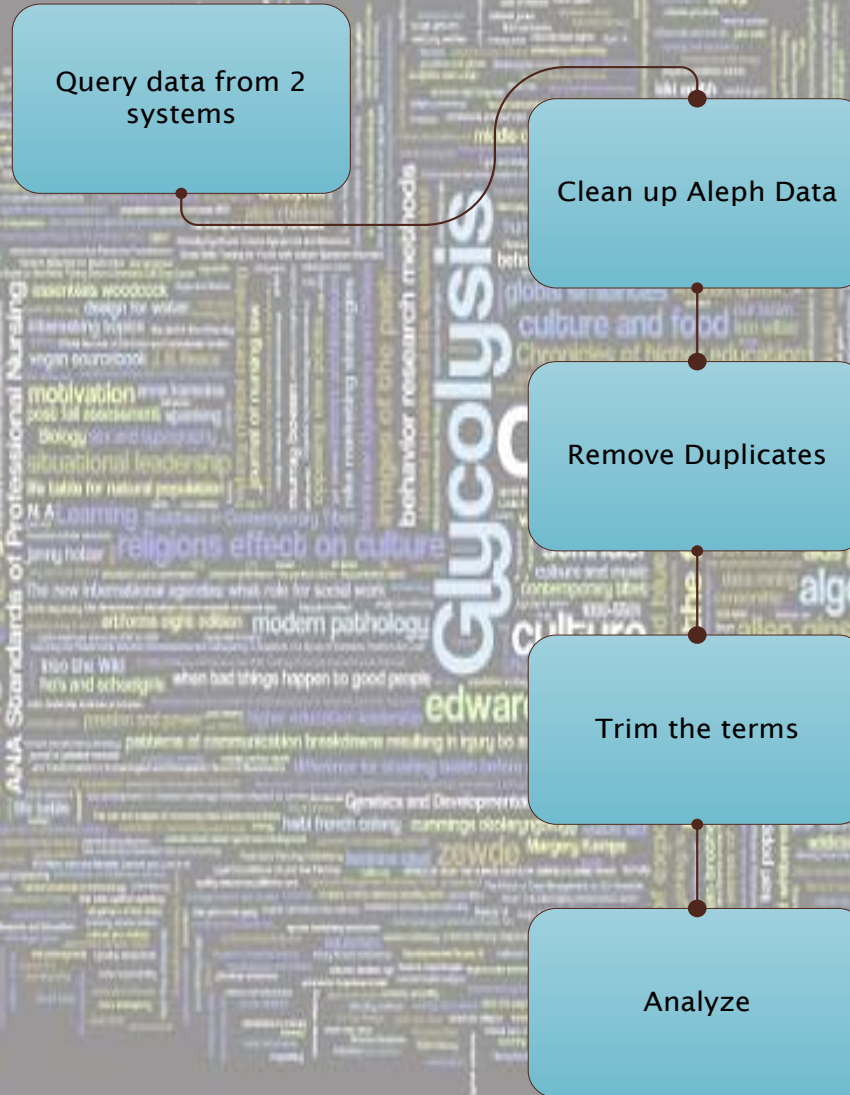
• 020	35
• 022	58
• ACO	1
• AUT	2949
• BAT	24
• CDA	5
• CDL	95
• CDS	62
• CMD	20
• CND	682
• CTA	1
• DVD	1900
• FLM	155
• GBA	5
• GCB	18
• IRS	20
• NEW	1
• PER	46
• SRS	14
• SUB	1263
• SUL	82
• SUM	9
• TRT	8970
• TTL	171
• TUT	6
• XYZ	10

A Smaller Sample

Analyzed 1 day of searching (2/18/2009)
to look at:

- Average length of search term
- Common search terms
- What terms appeared in Aleph but didn't originate from the Web
- What terms from the Web error out in Aleph
- Average number of hits in Aleph when the search originates from the Web vs. when it originates from Aleph

Data Massage



Results from February 18, 2009

Homepage

Of the searches from the homepage (679), 23% returned 0 results in the catalog

The average number of results from a query was 734 records

The largest number of results from a query was 74,459 for the search term 'Science'

Native Interface

Of the searches in the native interface (993), 22% returned 0 results in the catalog

The average number of results was 602 records

The largest number of results for query was 85,422 for the search term 'html'

Results from February 18, 2009

Homepage

The average search term length was ~3 words

The longest term was 19 words (“Exploring the Relationship Between Homelessness and Delinquency A Snapshot of a Group of Homeless Youth in San Jose California”) and the shortest was 1 word

52% of the searches contained 1 or 2 words, 85% contained 4 words or less!

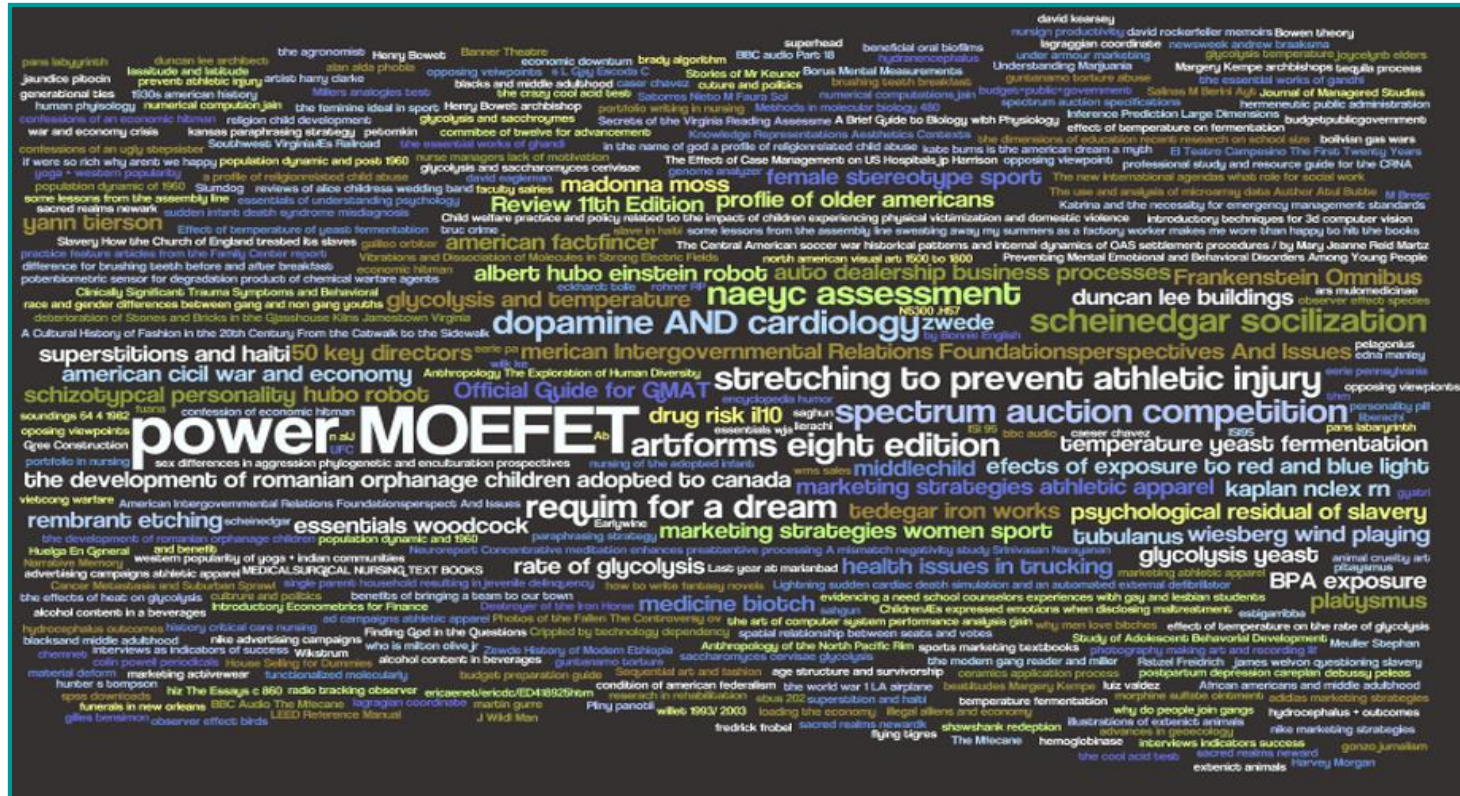
Native Interface

The average search term length was ~3 words

The longest term was 23 words (“some lessons from the assembly line sweating away my summers as a factory worker makes me wore than happy to hit the books”) and the shortest was 1 word

61% of the search terms contained 1 or 2 words – 87% contained 4 words or less!

Zero Results



Search Terms from the Web to the catalog for 2/18/2009 with 0 hits in the catalog

Created with Wordle: <http://www.wordle.net/create>

Zero Results

	<i>Total</i>	<i>Misspellings</i>	<i>Boolean operators</i>	<i>Punctuation</i>
Originating in Aleph	214	35.05 %	8.41%*	5.61%
Originating on the Web and sent to Aleph	156	25.00%	5.77%*	n/a

*100% Boolean operators were 'and' operands

Observations

Phrases were used quite often in lieu of using Boolean searching

Many of the search terms returned results in WorldCat, Google or Amazon.

Specific editions of books were requested: For example: "artforms eight edition," "Official Guide for GMAT« Review 11th Edition," "essentials of understanding psychology 7e"

Conclusions

- Location doesn't seem to matter
 - A search box is a search box no matter where it is
- Library website is a significant portal into the catalog
- The catalog needs to function more like a search engine
 - Getting something is better than nothing

Next Steps

- Perform more extensive text analysis of the search results
- Investigate better error trapping
 - Spell check
 - Citation recognition
 - Punctuation
- Investigate NextGen catalog interfaces

Questions for Discussion

1. How many entry points do you offer into the library catalog?

- How do you evaluate their effectiveness?
- Is consistency of search options important?

Questions for Discussion

2. How do you determine what types of search options (if any) to offer on the library homepage?

- How do you evaluate these search options?
- Is this external search featured in library promotion and instruction?
- Some have questioned the relevance of the library homepage, why bother with external search at all?

Questions for Discussion

3. In surveying ARL libraries the number of search silos varied from ten (site, catalog, catalog plus articles, Google web:scholar:books, staff directory, University site:people:departments) to zero.

Sixteen sites (14%) offered only one search silo.

- What does this say about the current state of library search?

Contacts

- Jimmy Ghaphery: jghapher@vcu.edu
- Tom McNulty: tmcnulty@vcu.edu
- Susan Teague-Rector: seteague@vcu.edu