

Check 2007

Pittsburg State University

“Google Mini Yields Big Results”
May 24th – 10:45AM

Presented by:

Michael Fienen

Webmaster

Pittsburg State University

mfienen@pittstate.edu

“Google Mini Yields Big Results”

Improving search for campuses on a budget



Introduction



By others, we relied on Google Free (now Co-op) or campus site search Webmaster Tools were useful for detailed results

To improve our ability to monitor visitor's interests and to identify areas not being identified by desired greater visitor information in and our web site

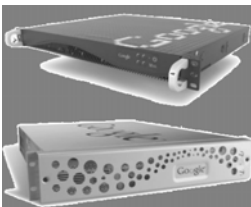
The Problem

- Results page look & feel was very limited
- Crawling frequency was very low
- No control over results
- Visitors weren't always served the newest information
- No way to investigate what people searched for
- Our own content managers couldn't understand why their updates weren't being reflected in results

What to do?

- Identified what we would like to have:
 - Frequent crawls
 - Detailed reports
 - Control over results
- Began looking for improved services
- Investigated options within Google for enhanced service
- A short search turned up the Google Mini and Google Search Appliance

Comparison



- Mini geared for simple use in smaller (<300,000 page) environments
- GSA very powerful but prohibitively expensive (starts at \$30,000)
- Mini had multiple license levels based on site size (starting at \$2,000)
- GSA could handle an enormous library of files (up to 3 million pages)

Let's Just Do It!

A decision was made to buy the Google Mini on the basis that for the cost, the benefits we would get from the box would be quickly apparent. Though made quickly, the decision was far from unwise.

From the time that we decided we needed to address our search problem, to having the server on my desk: less than 2 months.

Now It Gets Interesting

- You get a 1U server
- It's bright blue
- It has one button
- In short, it just *works*
- Oh, you also get a Google shirt. Mine was too big though :'(



We Are Not Alone!

- So, just how many of these are in the wild? Good question, and Google lets us see:

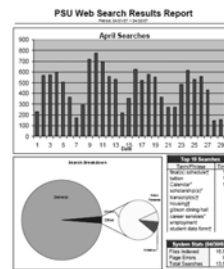


So, What'd We Get?

- Detailed user search data
- Ability to track page errors
- Ability to search multiple domains
- Customized departmental search boxes
- Ability to group indexes into custom collections
- Trickle down technology from the GSA (Google OneBox)
- Customized results pages for multiple collections
- Option to set KeyMatches for certain queries
- Future scalability

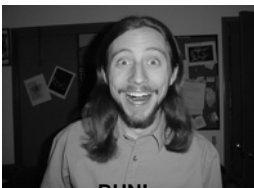


Using the Data



- Report data can be exported as XML
- Contains information on:
 - # of searches
 - Top keywords
 - Top queries
 - Queries that did or did not return results
 - Search frequency by hour

So Now What?



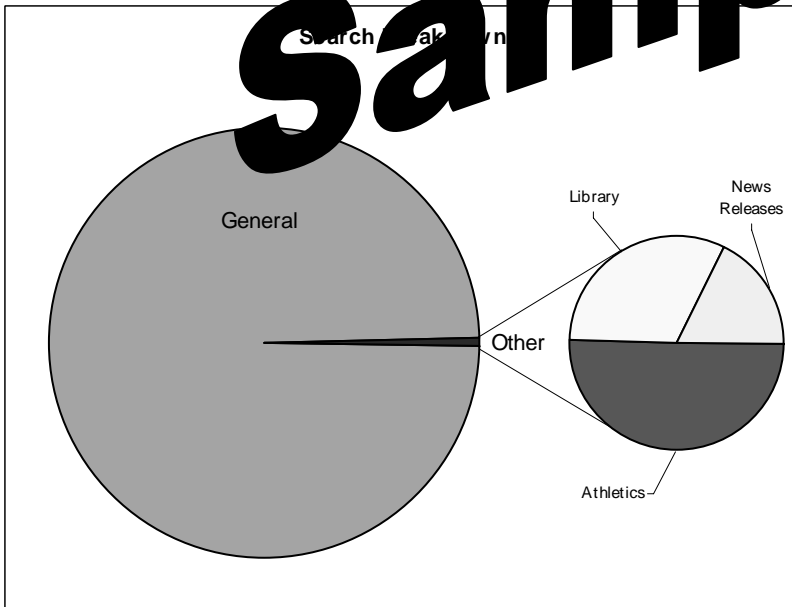
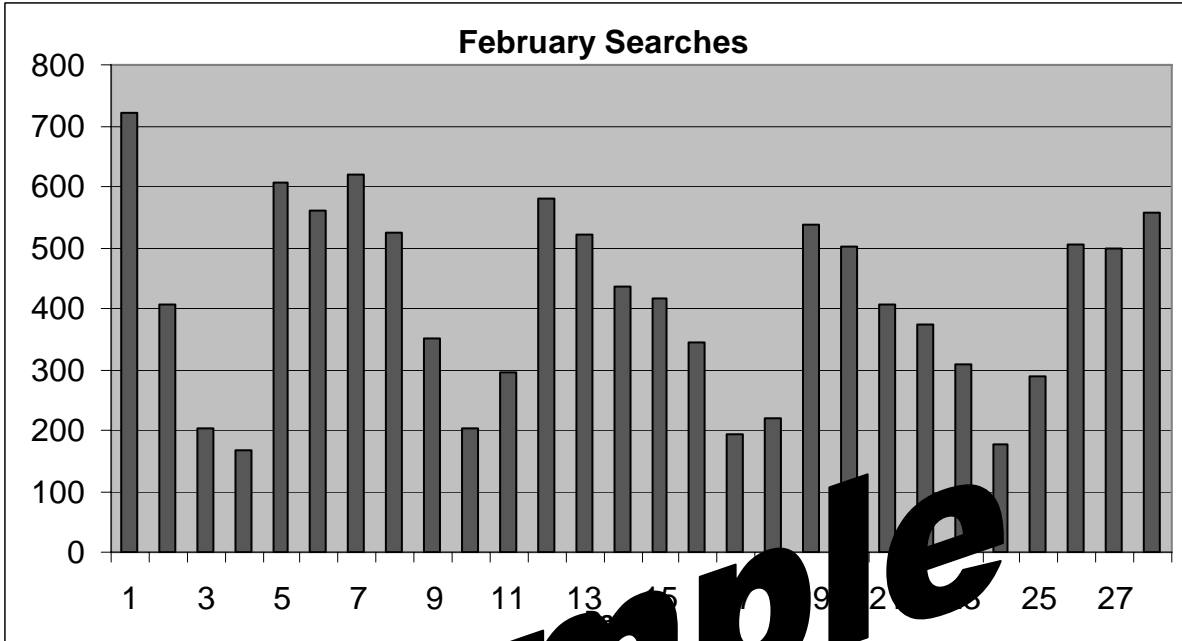
- Using search reports to craft landing pages for certain topics (calendars, dining, athletics, etc)
- Identifying periodic needs and trends (finals schedule, weather cancellations, etc)
- Creating custom search collections and front ends (symphony site, purchasing...)
- Giving departments customized search forms
- Merging search forms (people search vs. traditional)
- OneBox development!
- Wacky high jinks

Resources

- Compare the Mini and GSA
 - ~ http://www.google.com/enterprise/feature_comparison.html
- Mini User Group
 - ~ <http://groups.google.com/group/Google-Mini>
- Google Mini Feature List
 - ~ <http://www.google.com/enterprise/mini/features.html>
- OneBox Gallery Examples
 - ~ <http://code.google.com/enterprise/oneboxgallery.html>
- User Examples
 - ~ <http://www.pittstate.edu/search/>
 - ~ <http://www.alma.edu/search>
 - ~ <http://www.nwciowa.edu/search/search.aspx>
 - ~ <http://searchuw.wywu.edu/>

PSU Web Search Results Report

Period: 02/01/07 – 02/28/07



Top 10 Searches	
Term/Phrase	Times
<see below>	256
Career services*	149
Calendar*	137
Gibson dining hall	119
Scholarship(s)*†	92
Golden gorilla†	74
Tuition	69
Transcript(s)†	68
Housing†	61
Class cancellation(s)*	47

System Stats (02/28/07)	
Files Indexed	15,755
Page Errors	639
Total Searches	13,225

While individually not in the top 10 searches, the combined total for six different searches varying on the concept of “Gibson Dining Hall” was 256 queries. In the case of “Scholarship(s),” the single version returns bad results and is keymatched, the plural version organically returns a #1 result.

* Search term has a KeyMatch enabled

† Search term appears to bring up relevant results in the top 3 search results



Web Based Search Report View:

[Back to list of reports](#)

Report temp for collection PSU : May 2007

Total Results Pages: 6563 Total Searches: 6451 Distinct Searches: 5839

# Searches Per Day																															
May 2007																															
Date	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
# Searches	566	454	498	266	215	253	509	431	594	513	402	126	260	584	542	238															

Average # Searches Per Hour												
AM	0:00	1:00	2:00	3:00	4:00	5:00	6:00	7:00	8:00	9:00	10:00	11:00
	116	100	75	62	27	26	54	160	290	401	531	513
PM	12:00	13:00	14:00	15:00	16:00	17:00	18:00	19:00	20:00	21:00	22:00	23:00
	478	505	514	460	412	258	255	239	259	271	266	179

Top 100 Keywords	# Occurrences
schedule	534
finals	341
final	323
2007	216
student	179
of	173
center	135
hall	121
tuition	115
graduate	114
and	114
international	91

Top 100 Queries	# Occurrences
finals schedule	194
final schedule	141
finals	78
final exam schedule	44
tuition	43
calendar	41
gibson dining hall	40
faq	38
bookstore	28
health center	27
intensive english	25
Finals Schedule	25

Any of the stats collected here can be output as XML for use elsewhere. Reports can be customized by time, whether or not results were returned, excluding diagnostic terms, or filtering number of keywords returned.

Alternate control panels will also allow you to view specific search and event logs in detail, with similar filtering options.