

**Super Searching
Seminar:
How to get more
out of web research**

**Sponsored by:
NOLUG**

14 April 2008

**Mary Ellen Bates
Bates Information Services
Colorado, USA
BatesInfo.com**

Super Searching Seminar: How to get more out of web research

NOLUG Seminar
Mary Ellen Bates
14 April 2008



What we'll cover

- Using the New Web for Research
- Advanced and Hidden Features of the Big Three Search Engines
- Building and Using Customised Search Engines



www.BatesInfo.com

2

Using the New Web for Research



www.BatesInfo.com

3

WikiIndex.com

- Index of largest wikis
- Designed as a wiki
- Scroll down to the bottom of the page and use the hierarchical directory
- About 3,800 wikis



www.BatesInfo.com

4

Wikia.com

- Developed by Jimmy Wales' group
- Start at www.wikia.com/wiki/Category:Wikia_categories and drill down



www.BatesInfo.com

5

Wikia.com

- See also the new open source Wikia search engine
 - alpha.search.wikia.com
 - "We are aware that the quality of the search results is low.."



www.BatesInfo.com

6

Technorati

- Comprehensive blog search engine
- Can search for
 - blogs about X
 - blog posts tagged with X
 - blog posts that mention X
 - videos tagged with X



www.BatesInfo.com

7

Technorati

- Results page shows "mentions per day"
- Filter by "authority"!



www.BatesInfo.com

8

Technorati

- Blog Directory of "claimed" blogs
 - technorati.com/blogs
- Use technorati.com/search for advanced search
- Use to analyse not only the "buzz" but what *kind* of buzz



www.BatesInfo.com

9

Google Blog Search

- Searches anything with a feed
- Indexes *only* feed content
- Use the advanced search page
- Use `inpostauthor:` to find what someone's saying, to judge authority

blogsearch.google.com



www.BatesInfo.com

10

50Matches.com

- Social-powered search
- Limits search to sites listed in social bookmarking sites (del.icio.us, digg, reddit, etc)
- Max of 50 results
- Good for limiting to "top" sites



www.BatesInfo.com

11

Yahoo Pipes

- Executes search on specified news feeds
- Example: `snurl.com/1rxy5`
 - News articles about US Presidential candidates from specific news sources
- Example: `snurl.com/1rxya`
 - Library-related news
- [pipes.Yahoo.com](http://pipes.yahoo.com)



www.BatesInfo.com

12

Track Discussion Forums

- BoardTracker.com
- BoardReader.com
- Search message boards and forums
- Can search by word, thread or tag



www.BatesInfo.com

13

Searching podcasts

- Search or browse by category using:
 - PodcastAlley.com
 - Podcast.net
 - pluggd.tv
- Limit words to "series" rather than "episode"



www.BatesInfo.com

14

Podcast chunking

- Moving beyond metadata to speech recognition!
- "Chunking" of the search results
 - Searches *transcripts* of the podcast
- Podscope.com and Everyzing.com



www.BatesInfo.com

15

LinkedIn.com

- Generally, can only view *your* (extended) network
- Can search by name, company, position
- Organisations use this to link employees



www.BatesInfo.com

16

YoName.com

- Search across social networking sites
 - MySpace, LinkedIn, Friendster, Digg, Xanga, YouTube, Bebo, Twitter, etc.
- Not perfect, but a place to start



www.BatesInfo.com

17

Twitter.com

- Micro-blogging
- Track late-breaking news from a conference
- Track initial reactions to an announcement
- Learn more about an individual
- Instant-poll of your network



www.BatesInfo.com

18

Searching Twitter

- Google
 - `search-word site:twitter.com`



www.BatesInfo.com

19

Hidden and Undocumented Features of the Big Three Search Engines



www.BatesInfo.com

20

Google's Secrets



www.BatesInfo.com

21

Google's Experimental Search

- New way to see results
- Add `view:timeline` or `view:info` to the search query
- See results
 - in a timeline
 - with dates highlighted
 - with images featured(!)
 - highlighted on a map



www.BatesInfo.com

22

SearchMash.com

- Unbranded Google site
- Cool interface
- Results sorted by web page, images, blogs, videos, Wikipedia
- Easy search-within-site feature



www.BatesInfo.com

23

Date search

- Never entirely reliable
- Syntax is `daterange:nnnnn-nnnnn`
- Must use Julian calendar
 - Conversion tool at snurl.com/1twze



www.BatesInfo.com

24

Better date limiting

- Add `&as_qdr=XN` to the *search results page URL*
 - X is *d* (days), *w* (weeks) or *y* (years)
 - N is the number of days, weeks or years you want to limit with
- Example: `&as_qdr=w5` limits the results to pages spidered in the last 5 weeks



www.BatesInfo.com

25

Improved news timelines

- Creates a readable page
- Easy to scan, identify trends
- news.google.com/archivesearch



www.BatesInfo.com

26

The .. operator

- Lets you search a range of numbers
- Syntax: `number..number` or `number..` (e.g., `525..` or `..14`)
 - no spaces!
- Only recognizes \$ for currency – not €, £, etc.



www.BatesInfo.com

27

CustomizeGoogle.com

- Firefox fix for Google
- Nice customisation
 - removes ads
 - infinite scroll of results
 - numbers the results
 - lets you repeat search in other search engines
 - add links to archive.org, etc.



www.BatesInfo.com

28

Simply Google

- Excellent overview site of all Google's features and sites, downloads, blogs
- www.usabilityviews.com/simple_google.htm



www.BatesInfo.com

29

Google sees faces

- images.google.com
- add *&imgtype=face* to the end of the results page URL for pictures of faces
- add *&imgtype=news* to see news images



www.BatesInfo.com

30

Yahoo's Secrets



www.BatesInfo.com

31

Feature: ing

- Looks for page *with that* feature
- feature:index – just home pages
- feature:table
- feature:image – *not* image search
- feature:audio
- feature:video



www.BatesInfo.com

32

Regional limits

- Can limit by region

region:africa

region:asia

region:centralamerica

region:downunder

region:europa

(includes UK)

region:mediterranean

region:mideast

region:northamerica

region:southamerica

region:southeastasia

- Based on top-level domain



www.BatesInfo.com

33

Y!Q


- Contextual searching
- Finds "more like this" paragraph
- Suggests alternate phrases
- yq.search.yahoo.com



www.BatesInfo.com

34

SearchAssist

- On search results page, look for 
 - Displays similar and related concepts, narrower phrases
- Right now, only at yahoo.com, but other countries are coming



www.BatesInfo.com

35

Customized shortcuts

- Create your own shortcuts
 - `!yf crox` – Yahoo Finance search for ticker symbol CROX
 - `!aiip` – the AIIP directory
 - `!wiki mumps` – Wikipedia entry
 - Tied to your Yahoo account
 - search.yahoo.com/osc/help



www.BatesInfo.com

36

Live.com's Secrets



www.BatesInfo.com

37

"Prefer" option!

- Add `prefer:word` to search
 - Ranks these search results higher!
- Note: Exalead also has this feature:
 - `word1 OPT word2` makes word2 "preferred"



www.BatesInfo.com

38

Synonym-suggestion tool

- Intended to help search engine marketers identify similar words
 - snurl.com/zqyd



www.BatesInfo.com

39

Misspelling-suggestion tool

- Also for search engine advertising
- snurl.com/zqyl



www.BatesInfo.com

40

Search for feeds

- `hasfeed:search-term`
- Looks for pages that have your search term AND a link to an RSS or Atom feed



www.BatesInfo.com

41

"Has a link to"

- `contains:file-type`
- Looks for pages that contain a link to a certain file type, such as:
 - `contains:doc`
 - `contains:pdf`
 - `contains:ppt`



www.BatesInfo.com

42

Find product reviews

- search.live.com
- Watch for "smart answer" at top of search results



www.BatesInfo.com

43

Microsoft knows your face

- Live Search Images separates out
 - faces (`filter:face`)
 - portraits (`filter:portrait`)
 - black and white images (`filter:bw`)



www.BatesInfo.com

44

Building and Using Customised Search Engines



www.BatesInfo.com

45

What's a custom search engine?

- You filter the user's search query, limit to certain domains or sites, or tweak the relevance ranking
- Put the library's imprimatur on search results
- Out-google Google
 - www.BatesInfo.com/test.html



www.BatesInfo.com

46

Google Custom Search Engine

- www.google.com/coop/cse/
- Can limit by domain
- Specified sites can be ranked higher, exclusively searched or excluded from the search
- Can include all pages/sites linked from a site



www.BatesInfo.com

47

Google Custom Search Engine

- Can offer a "refinement" button
 - Customised narrowing of results
- Can add logo to search results
- Can allow contributions from others
- Easy to test results while creating



www.BatesInfo.com

48

Google Custom Search Engine

- **Examples of Google CSEs**
 - **Economics:**
cook.rfe.org/ESE_Google.html
 - **Best free reference sites:**
snurl.com/1sLrk



www.BatesInfo.com

49

Yahoo Search Builder

- **Can limit search by domain, append key terms or exclude sites or terms**
- **Can search web or news**
- **Shows tag cloud of search terms**
- **Can insert logo in the search results page**
 - builder.search.yahoo.com



www.BatesInfo.com

50

Swicki

- **Collaborative approach to filtering**
 - **Users can vote a site higher or lower in the results**
- **You supply initial key words**
- **It learns from click-throughs on search results pages**
 - **Modifies relevance ranking**



www.BatesInfo.com

51

Swicki

- Searches text or *video*
- At least 5 search terms required
- eureka.com



www.BatesInfo.com

52

Rollyo.com

- Pretty rudimentary
- Maximum of 25 web sites
- Can see others' "searchrolls"
 - E.g., KM-related topics at snurl.com/16x8o



www.BatesInfo.com

53

Gigablast CTS

- Also fairly rudimentary
- Specify domains to search
- Why use it? Gigablast features

- www.gigablast.com/cts.html



www.BatesInfo.com

54

Live.com

- Can include advanced syntax
 - `inbody:`, `prefer:`, AND/OR/NOT, etc.
- Maximum of 30 sites
- Can search the web or news

- `search.live.com/macros`

###



www.BatesInfo.com

55

Thank you!

Questions? mbates@BatesInfo.com

web site :: www.BatesInfo.com

blog :: www.LibrarianOfFortune.com



www.BatesInfo.com

56

Mary Ellen Bates

Mary Ellen Bates is the owner of Bates Information Services, providing business research to business professionals and special librarians, and consulting and training services to the information industry. She provides half- and full-day workshops on effective online research and other topics related to knowledge work.

She is a frequent keynote speaker, and is the author of six books and innumerable articles on the information industry.

Mary Ellen offers two free email monthly newsletters:

“Bates InfoTip” (BatesInfo.com/tip.html) and

“Info-Entrepreneur Tip of the Month” (BatesInfo.com/ib-tip.html)

If you would like to subscribe to either, just ask
(mbates@BatesInfo.com)

Her latest endeavor is her blog – [Librarian Of Fortune.com](http://LibrarianOfFortune.com)

She, her partner and her dogs live in Colorado, USA.

Mary Ellen Bates

Bates Information Services, Inc.

www.BatesInfo.com

mbates@BatesInfo.com

+1 303.772.7095