



From query based Information Retrieval to context driven Information Supply

Andrei Broder

Fellow & VP Emerging Search Technologies
Yahoo! Research
May 20, 2006

YAHOO!



The pre-history of web search ...

- Information retrieval as a modern scientific discipline has been around for 50-60 years
- 1945: Vannevar Bush's "As We May Think"
<http://www.theatlantic.com/doc/194507/bush>
- 1960+: Gerald Salton
- 1978: First ACM SIGIR conference
- 1992: First TREC conference
- See http://www.gslis.org/index.php/Information_Retrieval



The short history of web search – consumer side

- June 11, 1994 – Brian Pinkerton announces WebCrawler (crawled 4000 servers, 200K (?) pages)
- Dec 15, 1995 – Digital announces AltaVista (crawled at 2.5 M pages/day, had 30 M pages(?), claimed to be 100 times faster/bigger than competitors)
- 1998 – Google
- Apr 29, 2004 – Google IPO-mania envelops the world



Without search engines the web would not be possible

1. No incentive in creating content unless it can be easily found – other finding methods failed (taxonomies, bookmarks, etc)
2. The web is both a technology artifact and a social environment
 - “The Web has become the “new normal” in the American way of life; those who don’t go online constitute an ever-shrinking minority.” – [Pew Foundation report, January 2005]
3. Search engines make aggregation of interest possible:
 - Create incentives for very specialized niche players
 - Economical – specialized stores, providers, etc
 - Social – narrow interests, specialized communities, etc
4. The acceptance of search interaction makes “unlimited selection” stores possible:
 - Amazon, Netflix, etc



Web has grown to a mass medium (Pew Foundation studies)

- December 2002: “There is ... abundant evidence that the Internet is now the primary means by which many people get key information.”
 - <http://www.pewinternet.org/reports/toc.asp?Report=80>
- December 2003: “Online activity has consistently grown”
 - Total Internet adult users = 126 M (63% of adult population)
 - Are online on any given day = 66 M
 - Have used Internet to search = 85%
 - http://www.pewinternet.org/pdfs/PIP_Online_Pursuits_Final.PDF
- August 2004: “Getting information is the most highly valued and most popular type of everyday activity done online.”
 - 92% of Internet users say the Internet is a good place to go for **getting everyday information**
 - http://www.pewinternet.org/pdfs/PIP_Internet_and_Daily_Life.pdf
- January 2005: “The Web has become the “new normal” in the American way of life; those who don’t go online constitute an ever-shrinking minority.”
 - Online on any given day = 77 M
 - http://www.pewinternet.org/pdfs/Internet_Status_2005.pdf



Basic assumptions of Classic Information Retrieval

- **Corpus:** Fixed document collection
- **Goal:** Retrieve documents with information content that is relevant to user’s **information need**



Classic IR Goal

- Classic relevance
 - For each query Q and stored document D in a given corpus assume there exists relevance $\text{Score}(Q, D)$
 - Score is average over users U and contexts C
 - Optimize $\text{Score}(Q, D)$ as opposed to $\text{Score}(Q, D, U, C)$
 - That is, usually:
 - Context ignored
 - Individuals ignored
 - Corpus predetermined

Bad assumptions
in the web context



User Needs

- Need [Brod02, RL04]
 - **Informational** – want to learn about something (~40% / 65%)
 - Low hemoglobin
 - **Navigational** – want to go to that page (~25% / 15%)
 - Iberia Airlines
 - **Transactional** – want to do something (web-mediated) (~35% / 20%)
 - Access a service: Barcelona weather
 - Downloads: Mars surface images
 - Shop: Nokia mp3
 - Gray areas
 - Find a good hub: Car rental Spain
 - Exploratory search “see what’s there”



The evolution of commercial web search engines

- **First generation** -- use only “on page”, text data
 - Word frequency, language
- **Second generation** -- use off-page, web-specific data
 - Link (or connectivity) analysis
 - Sophisticated mathematical methods
 - Click-through data (What results people click on)
 - Anchor-text (How people refer to this page)
- **Third generation** -- answer “the need behind the query”
 - Focus on user need, rather than on query
 - Semantic analysis -- what is this about?
 - Integrates multiple sources of data
 - Help the user
 - UI, spell checking, query refinement, query suggestion, syntax driven feedback, context help, context transfer, etc
- **Fourth generation** – this talk!

1994-1997 AV, Excite, Lycos, etc

From 1998. Made popular by Google but everyone now

Still evolving



Search on the Web

- Corpus: The publicly accessible Web
- Goal: Retrieve high quality results that are relevant to user’s need

- Need

- Informational
- Navigational
- Transactional

Low hemoglobin
 Iberia Airlines
 Barcelona weather
 Mars surface images
 Nokia mp3

- Results

- Static pages = text, mp3, images, video, ...
- Dynamic pages = generated on request: mostly data base access, “the invisible web”, proprietary content, etc

First gen. SE 2nd gen. SE 3rd gen. SE

Third generation search

11



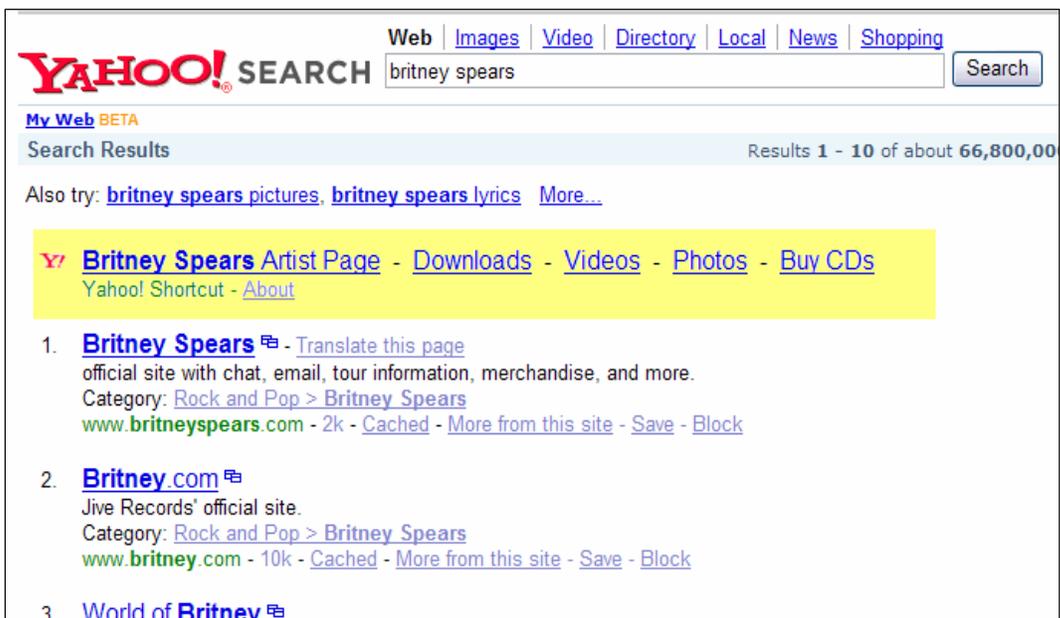
Third generation search engine: answering “the need behind the query”

- Semantic analysis
 - Query language determination
 - Auto filtering
 - Different ranking (if query in Japanese do not return English)
 - Hard & soft (partial) matches
 - Personalities (triggered on names)
 - Cities (travel info, maps)
 - Medical info (triggered on names and/or results)
 - Stock quotes, news (triggered on stock symbol)
 - Company info
 - Etc.
 - Natural Language reformulation
 - Integration of Search and Text Analysis

Third generation search examples

13

 **Yahoo!:** britney spears



The screenshot shows a Yahoo! search results page for the query "britney spears". At the top, there are navigation links for Web, Images, Video, Directory, Local, News, and Shopping. The search bar contains "britney spears" and a "Search" button. Below the search bar, it says "My Web BETA" and "Search Results". The results are displayed as "Results 1 - 10 of about 66,800,00". There are also links for "britney spears pictures", "britney spears lyrics", and "More...". The first result is highlighted in yellow and is a "Yahoo! Shortcut" for the "Britney Spears Artist Page", which includes links for Downloads, Videos, Photos, and Buy CDs. The second result is "Britney Spears" with a description of the official site and category "Rock and Pop > Britney Spears". The third result is "World of Britney".

Web | [Images](#) | [Video](#) | [Directory](#) | [Local](#) | [News](#) | [Shopping](#)

YAHOO! SEARCH

My Web BETA

Search Results Results 1 - 10 of about 66,800,00

Also try: [britney spears pictures](#), [britney spears lyrics](#) [More...](#)

Y! [Britney Spears Artist Page](#) - [Downloads](#) - [Videos](#) - [Photos](#) - [Buy CDs](#)
Yahoo! Shortcut - [About](#)

1. [Britney Spears](#)  - [Translate this page](#)
official site with chat, email, tour information, merchandise, and more.
Category: [Rock and Pop > Britney Spears](#)
[www.britneyspears.com](#) - 2k - [Cached](#) - [More from this site](#) - [Save](#) - [Block](#)
2. [Britney.com](#) 
Jive Records' official site.
Category: [Rock and Pop > Britney Spears](#)
[www.britney.com](#) - 10k - [Cached](#) - [More from this site](#) - [Save](#) - [Block](#)
3. [World of Britney](#) 



Ask Jeeves: las vegas

Ask.com | MyJeeves^{BETA}

Ask Jeeves Web | Pictures | News | Local | Products | More »

las vegas

Web Search: las vegas

[Map of Las Vegas, NV 89101](#) [About](#)

Local Information for Las Vegas, NV
Find: [Map](#) | [Jobs](#) | [Current Weather](#) | [Seasonal Climate](#) | [Local Time](#) | [Yellow Pages](#)
Go To: [Official Site](#) | [Chamber of Commerce](#) | [Facts](#) | [Schools](#)

[Other matches: Las Vegas, NM](#)

Latest News: Las Vegas [About](#)
[Pacific Coast League Standings](#) AP via San Francisco Chronicle 27 minutes ago
[New STACKED Trailer Reveals Exciting Poker Room Drama](#) gamesindustry.biz 1 hour ago

Sponsored Web Results

[Las Vegas Hotel Deals](#)
Up to 70% Off Everything **Vegas**; Rooms, Shows, Golf & Tours.
www.TripReservations.com

[Vegas Hotels & Shows](#)
Great Rates on Hotels, Shows & More **VEGAS.com** Best **Vegas** Rate Guarantee
www.VEGAS.com

[About](#)

[About](#)



Yahoo!: salvador hotels

YAHOO! SEARCH Web | Images | Video | Directory | Local | News | Shopping

salvador hotels

My Web BETA

Search Results Results 1 - 10 of about 17,900,000

Y! [Hotels in Salvador, Brazil - photos, reviews and deals](#)

[Pestana Bahia Hotel - Salvador, Brazil - ★★★★★ - from \\$95.00 - \[availability\]\(#\) - \[rate it\]\(#\)](#)

[Sol Victoria Marina - Salvador, Brazil - ★★★★★ - from \\$85.00 - \[availability\]\(#\) - \[rate it\]\(#\)](#)

[Praia do Forte Resort - Salvador, Brazil - ★★★★★ - from \\$170.00 - \[rate it\]\(#\)](#)

[Yahoo! Shortcut - \[About\]\(#\)](#)

- [San Salvador hotels: Read San Salvador hotel, motel, lodging reviews and compare prices](#)
[hotels in San ...](#)

San Salvador hotels, motels, resorts, inns and bed and breakfast: Find reviews, travel articles, guidebook lists, availability, price, deals, photos, class, amenities and more about hotels in San Salvador, El Salvador at Y.travel.yahoo.com/p-hotel-482915-san_salvador_hotels-i - [More from this site](#) - [Save](#) - [Block](#)

- [Salvador Brazil Hotels ... Jake.com](#)

Jake.com offers a cool new way to search for the perfect hotel for your trip to Salvador, Brazil. If you need a Business, Family or just want to stay in the hippest hotels Jake.com will help you find what your looking for.

[Salvador Hotels](#) [Airmort \(0\) Salvador Hotels](#) [Bahia Othon Palace](#)



Yahoo shortcuts

- Various types of queries that are “understood”

Shortcuts Categories

most popular: [Images](#) - [Maps](#) - [Weather](#) Suggest a new Shortcut. [Click Here](#)

 Local	 News & Information	 Travel	 Reference	 Calculators	 Custom
Maps Weather Local Listings more...	Sports Scores Stock Quotes Images more...	Airport Information Flight Tracker Hotel Finder more...	Dictionary Definitions Encyclopedia Lookup Synonym Finder more...	Calculator Time Zones Weights and Measures Converter more...	Open Shortcuts



Google andrei broder new york

Web [Images](#) [Groups](#) [News](#) [Froogle](#) [Local](#) [Desktop](#) [more »](#)

Google

Web Results 1 - 10 of about 13,200 for **andrei broder new york**. (0.22 seconds)

[Phonebook results for andrei broder new york](#)

 **Andrei Broder**, (718) 432-6973, 630 W 246th St, Bronx, NY 10471
[Google Maps](#) [Yahoo! Maps](#) [MapQuest](#)

[A taxonomy of web search](#)
Publisher. ACM Press **New York**, NY, USA ... Aris Anagnostopoulos, **Andrei Z. Broder**, David Carmel, Sampling search-engine results, Proceedings of the 14th ...
portal.acm.org/citation.cfm?id=792552 - [Similar pages](#)

[Optimal plans for aggregation](#)
Andrei Broder, IBM Research Division. Michael Mitzenmacher, Harvard University ... John Wiley and Sons **New York** 1983 16 M. Shaked and J.G. Shanthikumar ...



Answering “the need behind the query” Context determination

- Context determination
 - spatial (user location/target location)
 - query stream (previous queries)
 - personal (user profile)
 - explicit (user choice of a vertical search,)
 - implicit (use Google from France, use google.fr)
- Context use
 - Result restriction
 - Kill inappropriate results
 - Ranking modulation
 - Use a “rough” generic ranking, but personalize later



The spatial context -- geo- search

- Geo-parsing = the process of identifying geographic context
 - Pages (infer from phone nos, zip, etc). About 10% can be parsed.
 - Queries (use dictionary of place names)
 - Users
 - Explicit (tell me your location, registration, from ISP)
 - From IP data
 - Mobile phones
 - In its infancy, many issues (display size, privacy, etc)



Google: dentists bronx

Web Images Groups News Froogle Local Desktop more »

Google dentists bronx Search

Web

Local results for **dentists near Bronx, NY**

 [Montefiore Medical Ctr](#) - 2.0 miles NE - 3448 Boston Rd, Bronx, 10469 - (718) 547-6111
[Montefiore Medical Group](#) - 3.2 miles SW - 305 E 161st St, Bronx, 10451 - (718) 579-2500
[Bronx Park Dental Group](#) - 0.1 miles NE - 2016 Bronxdale Ave # 303, Bronx, 10462 - (718) 792-7972

Bronx Dentists & Dental Insurance @ Dental Plans
Bronx Dentists @ Dental Plans - Individual and Family Discount Dental Plans and Insurance, Affordable Dental Coverage Starting at \$79.95 a Year!
[static.dentalplans.com/newyork/bronx/](#) - 37k - [Cached](#) - [Similar pages](#)

Bronx, NY - Dentist Reviews, Ratings & Recommendations.
 DR. Oogle: **Dentist Reviews**, **Dentist Ratings**, **Dentist Recommendations** and **Dentist Help**. Post your case to receive second opinion from another **dentist** or ...
[new-york.doctooogle.com/dentist_reviews.cfm/pageID/8/index/E](#) - 81k - Aug 6, 2005 - [Cached](#) - [Similar pages](#)

New York, NY - Dentist Reviews, Ratings & Recommendations.
Dentist, New York, NY Abanto, Laarni **Dentist**, **Bronx**, NY Abati, Mario **Dentist**, **Bronx**, NY Abbadessa, Joseph **Dentist**, Staten Island, NY Abbaticchio, Stephen ...
[new-york.doctooogle.com/dentist_reviews.cfm/pageID/8](#) - 120k



Yahoo!: dentists 10471

Web Images Video Directory Local News Shopping

YAHOO! SEARCH dentists 10471 Search

My Web BETA

Search Results

Y! Local Results for **dentists near Bronx** - [Map All Results](#)

[Quinnones, Madelyn - Citident Family Dentists](#) (718) 378-5030 - 1581 Westchester Ave, **Bronx**, NY - 1.31mi [map](#)
[Chumsky, Jay R DDS - Citident Family Dentists](#) (718) 378-5030 - 1581 Westchester Ave, **Bronx**, NY - 1.31mi [map](#)
[Gold, Andrew DDS - Andrew Gold Association Dentists](#) (718) 299-3600 - 505 Claremont Pkwy, **Bronx**, NY - 2.18mi [map](#)
[Yahoo! Shortcut](#) - [About](#)

1. **Bronx, NY Dentists at SuperPages.com** 
 Find the Best **Dentists** in **Bronx**, New York at SuperPages.com. SuperPages from Verizon has listings for many more **Bronx** businesses.
[yellowpages.superpages.com/listings.jsp?C=dentists&CID=493578&...](#) - [More from this site](#) - [Save](#) - [Block](#)

2. **Dentists and Clinics in Bronx, New York - Directory USA** 
 ... **Dentists** and Clinics in **Bronx**, NY ... 729 Burke Ave. **Bronx**, NY 10467-6638 ...

Search for: dentists Address, City & State, or Zip Search Tips
Bronx, NY
 Make this my default Yahoo! location

dentists in ★ Bronx NY

Zoom In
1
2 street
3
4 city
5
6
7
8 state
9
10 country
Zoom Out

Clicking on Map:
Zoom in & Re-Center
Re-Center Only
Map Legend
SmartView™

dentists
Local Results:
1. [Wilson, Fred - Br...](#)
2. [Belluzzi, Felice ...](#)
3. [Schaffer, Samuel ...](#)
4. [Ratner, Ei Dds](#)
5. [Maqideko, Roman D...](#)
6. [Kasparian, Harry ...](#)
7. [Piltser, Yakov Dd...](#)
8. [Chu, Eric K Dds](#)
9. [Kirshbaum, David ...](#)
10. [Peruso, William D...](#)

Results 1-10 of 24
[Next](#)
[View Detailed Results List](#)

Map Nearby
Also find:
e.g. restaurant, museum near current results.

© Yahoo! Research 2006

23



Third generation search: the triple win

- Answering “the need behind the query” rather than simply returning query matches yields
 - A win for users (better results)
 - A win for content providers (focus)
 - A win for search engines (“monetization of infomediary role”)

© Yahoo! Research 2006

24

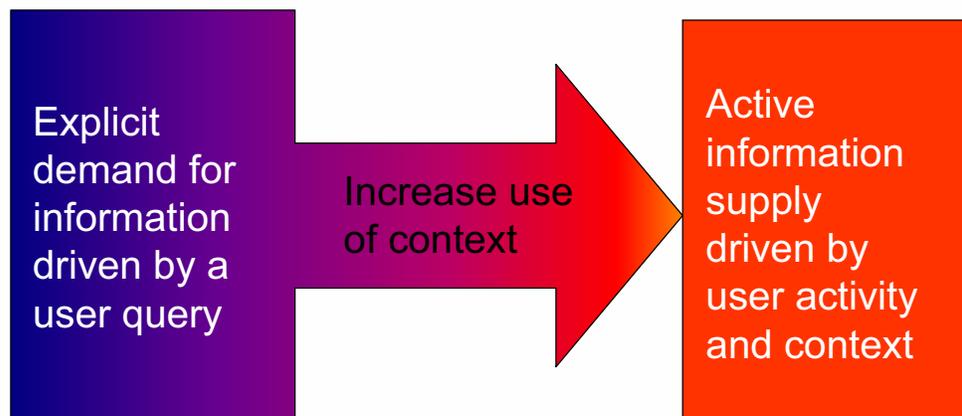
Main historical trend in web search

Move from syntactic matching to (maybe trivial but effective) semantic matching

25



What's next? Fourth generation:
From Information Retrieval to Information
Supply





Historical information supply sources



© Yahoo! Research 2006

27



From Information Retrieval to Information Supply: Buddy presence

Then...

```
> finger karger@CSAIL.MIT.EDU
[CSAIL.MIT.EDU]
KARGER    David Karger          Theory of Computation      Faculty
          <karger@theory.lcs.mit.edu>
          Project: Analysis of Algorithms and/or Information Retrieval
          Work: NE43-321; 258-6167
          Home: 1600 Mass. Ave., Apt. 407, Cambridge, MA 02138; 4919592
>
```

Now...



© Yahoo! Research 2006

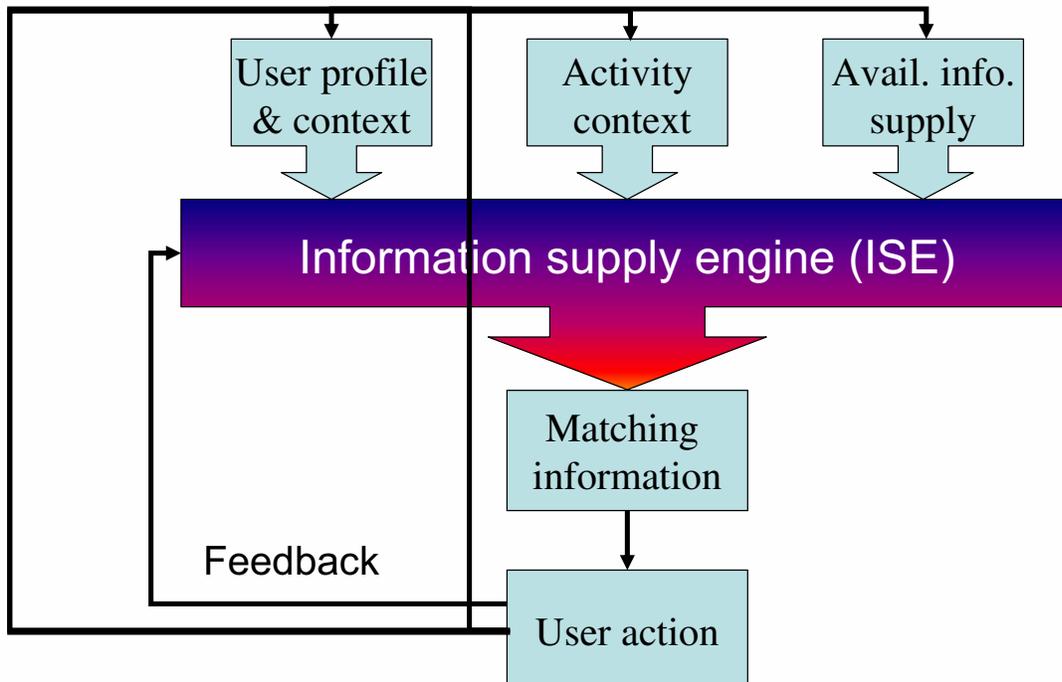
8:54 PM



From Information Retrieval to Information Supply: Car Navigation: Maps → GPS



Information supply picture





Some current information supply approaches

- Recurrent needs
 - Subscriptions (e-mail, RSS, etc)
 - Alerts – News
- Temporary needs
 - E-commerce sites: accessories, commentaries, related purchases, etc
 - Travel sites: Fly to BCN → Book a hotel, rent a car, etc
 - Contextual help → “You seem to be writing a letter”
 - Automatic annotations
 - ...
 - Contextual ads & search driven ads



Subscriptions examples

The screenshot shows a Yahoo! News search results page. At the top, the Yahoo! logo is followed by 'NEWS' and navigation links for Web, Images, Video, Audio, Directory, Local, News, Shopping, and More. A search bar contains the text 'Israel Innovation Summit' and a 'Search' button. Below the search bar are links for Subscriptions, Advanced Search, and Preferences. The main content area is titled 'News Results' and shows 'NEWS STORIES Results 1 - 10 of about 29 for Israel Innovation Summit.' There are two main results listed:

- 1. World Economic Forum Technology Pioneer in Stem Cell Research Invited to Speak at Israel Innovation Summit 2006**
[Press Release] PR Web - Mar 29 12:25 AM
Dr. Valentin Fulga, CEO of the TheraVita Group, the producer of VesCell™ -- Adult Stem Cell Therapy for Heart Disease, has been invited to speak at the inaugural Israel Innovation Summit 2006. The Summit will be held on 4-5 April at the Haifa Congress Center, Haifa, Israel. The World Economic Forum recently named Dr. Fulga as one of its 2006 Technology Pioneers for his work in adult stem cell
[Save to My Web](#)
- 2. Israel Innovation Summit 2006 Opens In April**
Wireless IQ - Mar 20 9:07 AM
The Israel Innovation Summit will be held on April 4-5, 2006. The only...
[Save to My Web](#)

On the right side of the page, there is a 'NEWS ALERTS' section with a bell icon and the text: 'Alert me when articles match all these words: Israel Innovation Summit'. Below this is an 'ADD TO MY YAHOO! / RSS' section with the text: 'Add your news search for Israel Innovation Summit to My Yahoo! MY YAHOO!' and a 'View as RSS: XML' link.

YAHOO! ALERTS

Stay in the know wherever you go!
Get real-time updates delivered to you instantly via:

Yahoo! Alerts connects you to the latest news, weather, stocks, and more!
[Learn More, Take a Tour](#)

Y! Automatic Annotations



Democrats used the new report to blast **President Bush**'s economic policies, contending it would be wrong to make permanent his tax cuts which primarily benefited the wealthy.

"These statistics show why, even though **GDP** is rising, most people do not feel better off," said Sen. Charles Schumer (news, bio, voting record), D-N.Y.

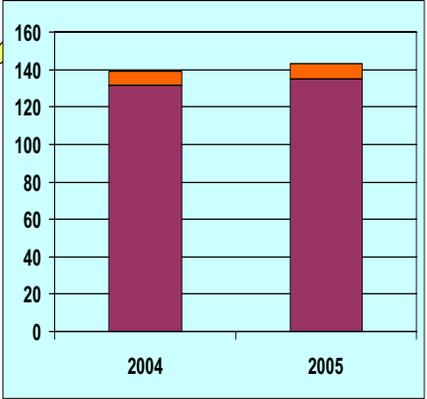


Advertising spend in USA in 2005

Excludes search advertising [TNS Media Intelligence]

Media	Full Year 2005 (Millions)	Full Year 2004 (Millions)	% Change
NEWSPAPERS (LOCAL)	\$25,090.50	\$24,814.40	1.10%
NETWORK TV	\$22,455.20	\$22,523.40	-0.30%
CONSUMER MAGAZINES	\$21,688.00	\$20,167.40	7.50%
CABLE TV	\$15,874.10	\$14,248.80	11.40%
SPOT TV ²	\$15,529.20	\$17,158.70	-9.50%
INTERNET³	\$8,322.70	\$7,343.00	13.30%
LOCAL RADIO ⁴	\$7,364.90	\$7,273.40	1.30%
B-TO-B MAGAZINES	\$4,471.00	\$4,364.60	2.40%
SYNDICATION - NATIONAL	\$4,222.50	\$3,930.90	7.40%
SPANISH LANGUAGE MEDIA ⁵	\$4,219.20	\$3,976.10	6.10%
OUTDOOR	\$3,528.80	\$3,213.00	9.80%
NATIONAL NEWSPAPERS	\$3,466.70	\$3,303.50	4.90%
NATIONAL SPOT RADIO	\$2,604.10	\$2,616.50	-0.50%
SUNDAY MAGAZINES	\$1,619.50	\$1,497.40	8.20%
FSI's ⁶	\$1,441.50	\$1,391.90	3.60%
NETWORK RADIO	\$1,009.90	\$1,027.80	-1.70%
LOCAL MAGAZINES	\$385.50	\$317.70	21.30%
TOTAL⁷	\$143,293.40	\$139,168.60	3.00%

Internet 2005: \$ 8.3B
(+13.3% vs 2004)

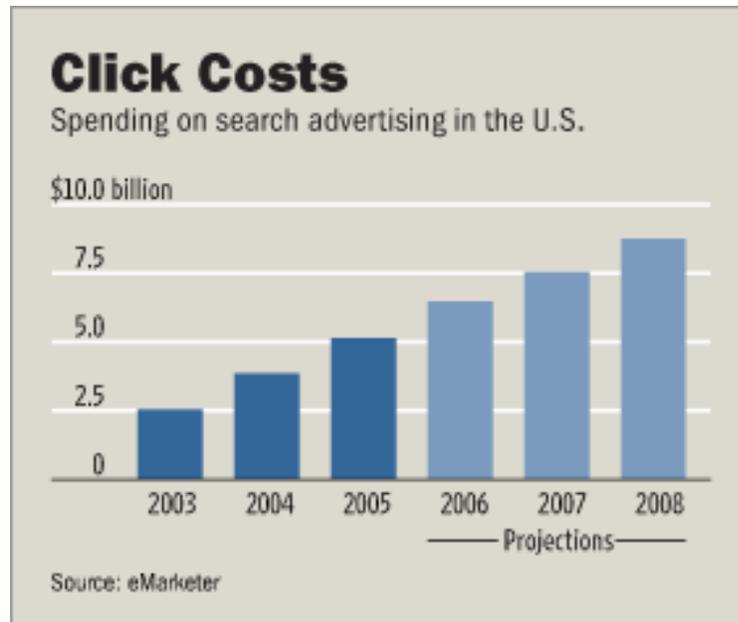


Total US 2005: \$ 143.3B
(+3% vs 2004)

Source: TNS Media Intelligence



Search advertising spending



© Yahoo! Research 2006

39



Search as mass medium (97% of revenue of web search)

- Sell advertising/audience reach
- As opposed to classic media can measure
 - Clickthrough rate (CTR)
 - Conversion rate (from browsers to buyers)
- Concepts
 - CPM = cost per *mille* (thousand) impressions
 - CPC = cost per click
 - CPT/CPA = cost per transaction/cost per action a.k.a. referral fees or affiliate fees

© Yahoo! Research 2006

40



A sponsored search ad

Google [Web](#) [Images](#) [Groups](#) [News](#) [Froogle](#) [Maps](#) [more »](#)

www2006 [Advanced Search](#) [Preferences](#)

Web Results 1 - 10 of about 178,000 for www2006. (0.06 seconds)

WWW2006
World Wide Web Conference. Edinburgh, Scotland. 22-26 May 2006.
[www2006.org/](#) - 15k - May 5, 2006 - [Cached](#) - [Similar pages](#)

WWW2006 - Workshops
The role of the workshop track in WWW2006 is to further research, development and ...
Additional questions about the WWW2006 Workshops can be emailed to ...
[www2006.org/workshops/](#) - 27k - May 5, 2006 - [Cached](#) - [Similar pages](#)
[[More results from www2006.org](#)]

WWE 2006 | 3rd Annual Workshop on the Weblogging Ecosystem ...
WWW2006 3rd Annual Workshop on the. Weblogging Ecosystem: Aggregation, Analysis and Dynamics · Call for Papers · Workshop Program. Important Dates ...
[www.blogpulse.com/www2006-workshop/](#) - 6k - [Cached](#) - [Similar pages](#)

WWW2006
WWW2006. 14 Mar 2006. Volume 9, Issue 29. Developers, start your engines! ... #The dev-track program at <http://www2006.org/developers/> has now been updated ...

Sponsored Links

[WWW2006 Tutorial](#)
Foundations and challenges of web advertising -- Edinburgh, May 24
[www2006.org/tutorials/#T01](#)



A content match ad

Content match ad

Photographer's Stop

Portal | Forum | Album | Search | Memberlist | Profile | Private Messages | FAQ | Register | Log in

SPONSORED LINKS	General									
<p>Ads by Yahoo!</p> <p>Olympus Digital Cameras - Official... Visit the official Olympus Web site for comprehensive information about our digital cameras and... www.olympusamerica.com</p> <p>Latin American Art Galleries Online Visit Latin American online art galleries. Features images and detailed information. Read... www.artnexus.com</p> <p>Discount Prices on Sony Digital... 5-star CNET service rating - find low prices today at BeachCamera.com. www.beachcamera.com</p>	<p>Forum</p> <p>Welcome - come on in! Hi - are you new? Come in and let everyone here know about you. Moderator shanky_pec</p> <p>Announcements, Feedback & Suggestions Check this section for board updates, announcements of new features. Leave your kind feedback, suggestions here.. Moderator shanky_pec</p> <p>Site Support Report site related problems, broken links, not-working-features here. We will surely resolve problems Moderator shanky_pec</p>	<table border="1"> <thead> <tr> <th>Topics</th> <th>Posts</th> </tr> </thead> <tbody> <tr> <td>29</td> <td>123</td> </tr> <tr> <td>18</td> <td>67</td> </tr> <tr> <td>5</td> <td>18</td> </tr> </tbody> </table>	Topics	Posts	29	123	18	67	5	18
Topics	Posts									
29	123									
18	67									
5	18									
	Photography									
	<p>Forum</p> <p>General Discussion General photography discussions Moderator shanky_pec</p> <p>Photo Album Discussion, tips-tricks, suggestions regarding photo gallery Moderator shanky_pec</p>	<table border="1"> <thead> <tr> <th>Topics</th> <th>Posts</th> </tr> </thead> <tbody> <tr> <td>64</td> <td>284</td> </tr> <tr> <td>26</td> <td>129</td> </tr> </tbody> </table>	Topics	Posts	64	284	26	129		
Topics	Posts									
64	284									
26	129									



Content match example (II)

The screenshot shows an MSN Entertainment page with an 'Artist Spotlight' for J.S. Bach. A 'Sponsored Sites' box is overlaid on the right side of the page, listing several e-commerce sites:

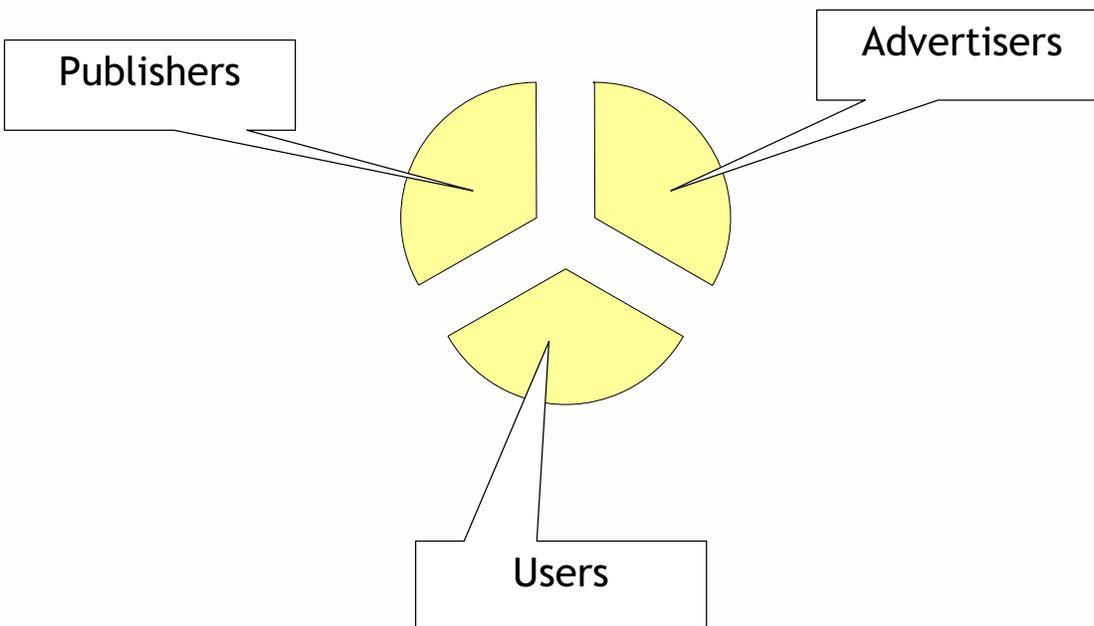
- Music by J. S. Bach at Amazon.com**
Amazon.com has a huge selection of merchandise, including videos and DVDs at great savings. Free Super Saver Shipping.
- Find "J. S. Bach" from \$55.00 at Buy.com**
Buy now at Buy.com. With over 1 million products to choose from, you can buy with confidence at Buy.com.
- If It Makes Music, It's on eBay**
You can find J. S. Bach music and collectibles right here today, you'll find the artists you're looking for on eBay.

© Yahoo

43



Contextual ads = meeting of Publishers, Advertisers, Users



© Yahoo! Research 2006

44



Quality of matching



Quality of matching



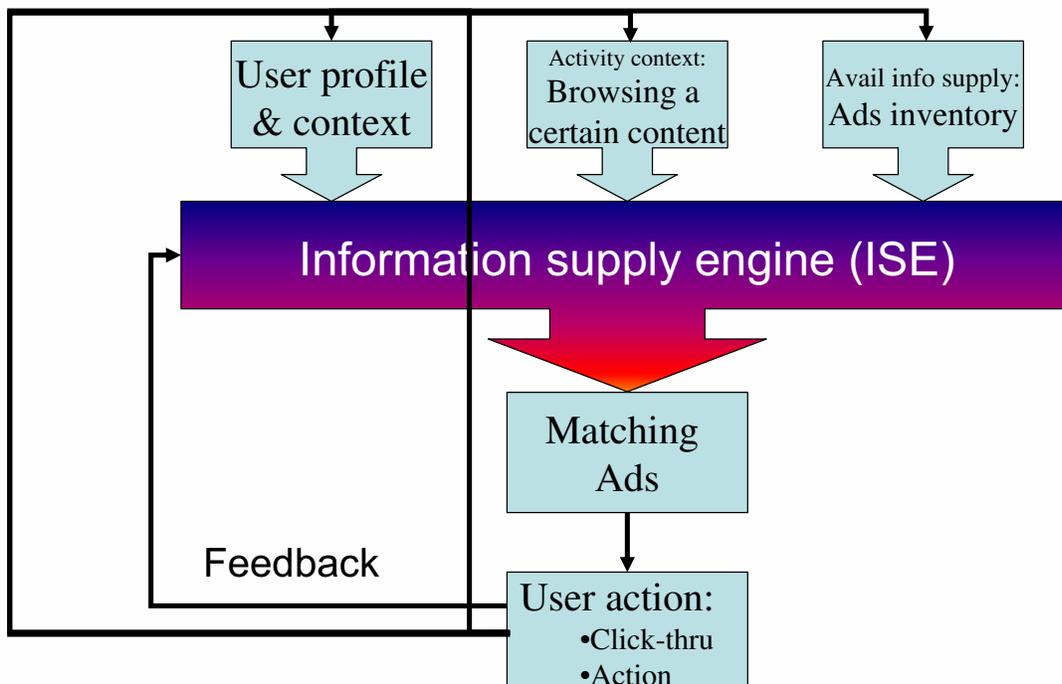


Ads as information supply

- “Advertising as Information” [Nelson, 1974]
- Irrelevant ads are annoying; relevant ads are interesting
 - Vogue, Skiing, etc are mostly ads and advertorials
- Even for keyword ads context is more than the keyword
 - User profile
 - Query stream
 - Location
 - Previous impressions



Ads as Information supply





Technical challenges to the transition to information supply

- A theory of information supply
- Representation of context
 - Vector model with lots of parameters?
- Representation of information
 - Certainly bag of words is not enough ...
- Representation of user
 - Probabilistic data
- Matching the three above

broder@yahoo-inc.org

Thank you!

HOW BIG CAN

Users served **429 million**
Yahoo's worldwide **9800**
Yahoo! Big Thinkers at WWW2006 **22**

Will you be the Yahoo! Big Thinker at WWW2007?

YOU THINK?

Interested? Email your resume to: thinkbig@yahoo-inc.com

51



LIFE ENGINE™