



The Future of Web Search: From Information Retrieval to Information Supply

Andrei Broder

Fellow & VP Emerging Search Technologies
Yahoo! Research

November, 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

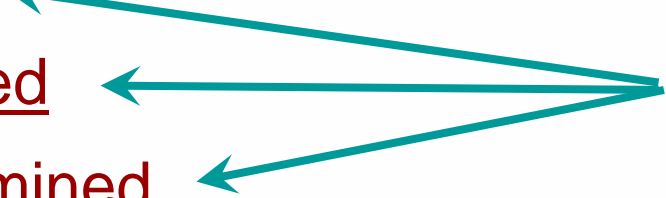


Basic assumptions of Classic Information Retrieval

- **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
- 



User Needs

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



The evolution of commercial web search engines

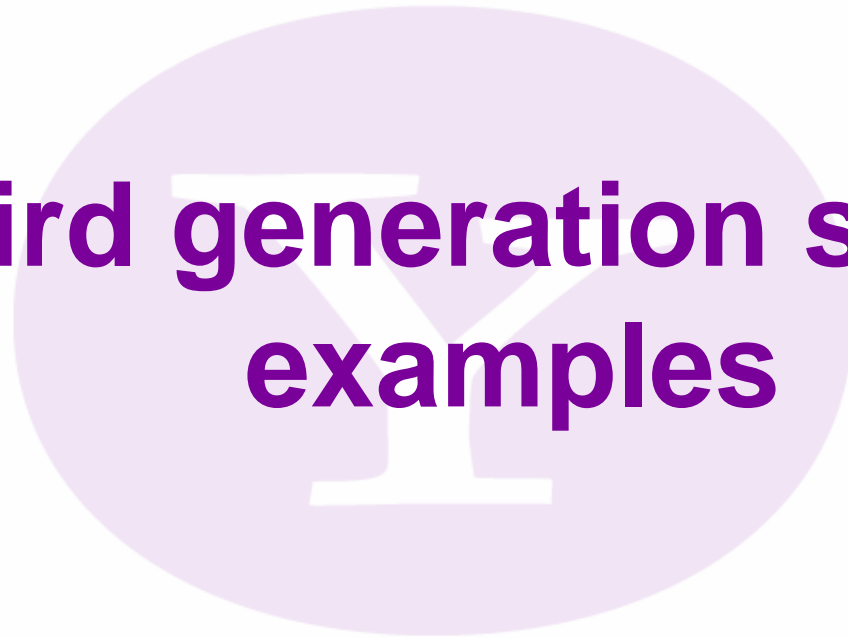
- **First generation** -- use only “on page”, text data
 - Word frequency, language
- **Second generation** -- use off-page, web-specific data
 - Link (or connectivity) analysis
 - Sophisticated mathematical methods
 - Click-through data (What results people click on)
 - Anchor-text (How people refer 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

Third generation search examples





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
 - Tracking numbers (triggered on regular expressions)
 - Etc.
 - Natural Language reformulation
 - Integration of Search and Text Analysis



Yahoo! britney spears

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

britney spears

[My Web](#) BETA

Search Results Results 1 - 10 of about 66,800,000

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 Britnev](#)



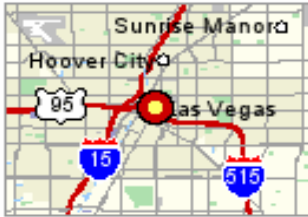
Ask Jeeves: las vegas

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



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 Airport (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 »](#)

andrei broder new york

Search

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 ...



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

LAN Air

Santiago 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.

2nd gen. SE

3rd gen. SE



Third generation search: the triple win

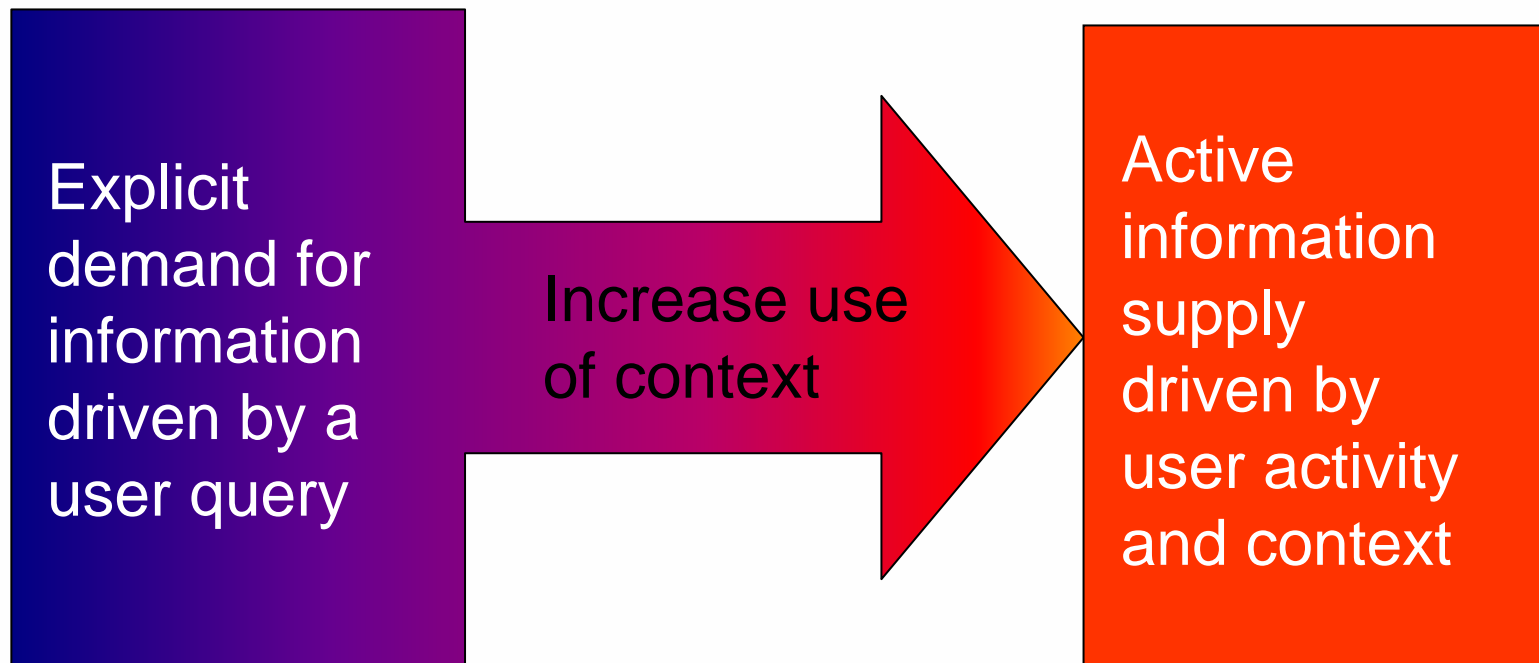
- Answering “the need behind the query” rather than simply returning query matches yields
 - A win for users (better results)
 - A win for content providers (focus)
 - A win for search engines (“monetization of infomediary role”)

Main historical trend in web search

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



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





Historical information supply sources





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...





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

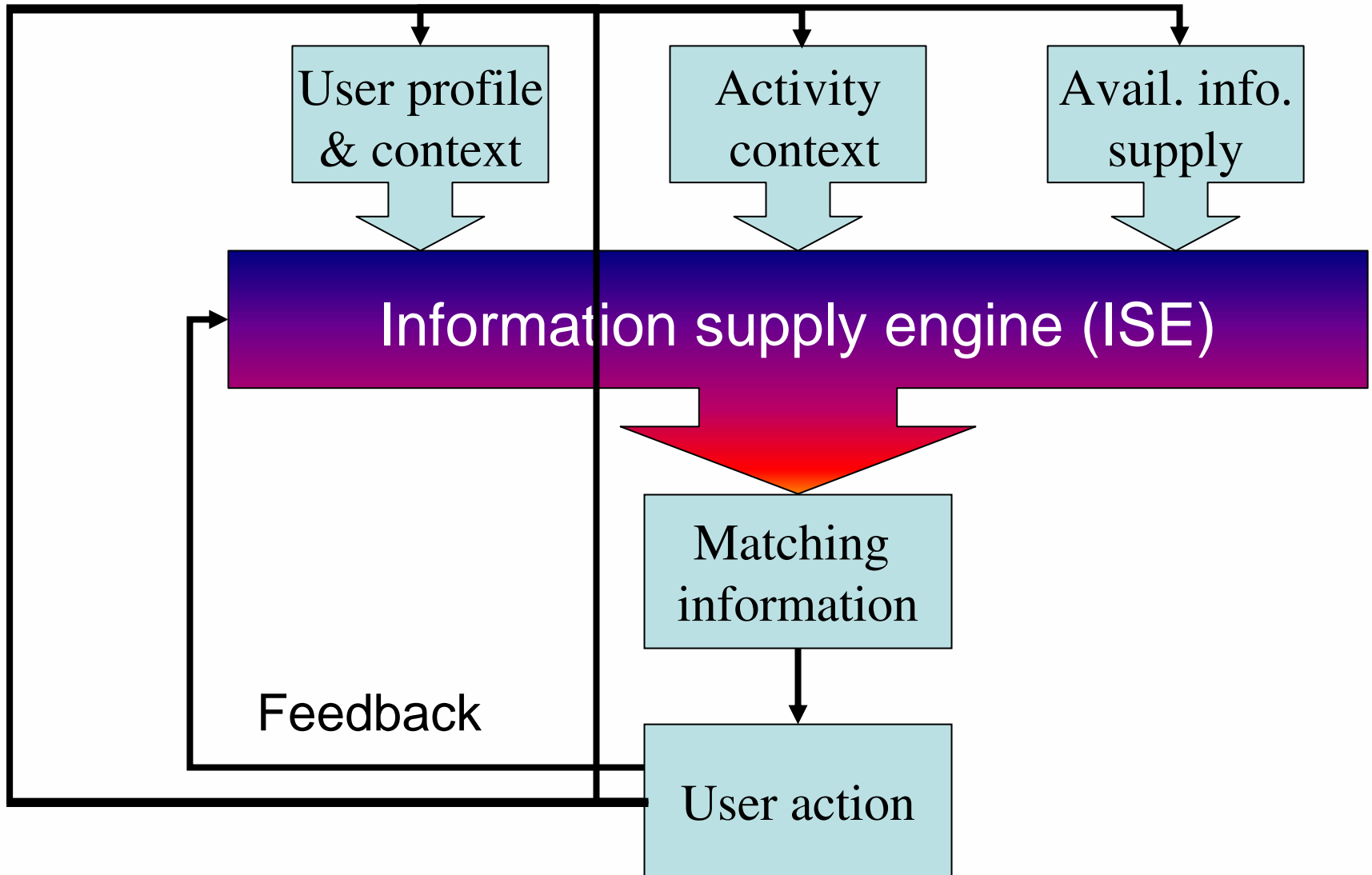


Now...





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 Santiago → 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



YAHOO! NEWS [Web](#) | [Images](#) | [Video](#) | [Audio](#) | [Directory](#) | [Local](#) | **News** | [Shopping](#) | [More »](#)

Israel Innovation Summit


[Subscriptions](#) | [Advanced Search](#) | [Preferences](#)

News Results


NEWS STORIES Results 1 - 10 of about 29 for Israel Innovation Summit. Sort Results by: Relevance | [Date](#)

- [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 TheraVitae 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](#)
- [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](#)

NEWS ALERTS

 Alert me when articles match all these words:
[Israel Innovation Summit](#)

ADD TO MY YAHOO! / RSS

Add your news search for [Israel Innovation Summit](#) to My Yahoo!  [MY YAHOO!](#)

View as RSS: [XML](#)



Alerts

YAHOO! ALERTS

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

email

mobile device

Yahoo! Messenger

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



Automatic Annotations

Yahoo! My Yahoo! Mail

YAHOO! NEWS Welcome, **axbroder**
[Sign Out, My Account]

News Home - U.S. - Business - World - Entertainment - Sports - Tech - Politics - Science - Health - Most Popular - Index

Economy - Stock Markets - Earnings - Personal Finance - Opinion - Press Releases

Search: All News & Blogs Search Advanced

Average American Family Income Declines

By MARTIN CRUTSINGER, AP Economics Writer
Thu Feb 23, 6:16 PM ET

RELATED QUOTES

*DJI	11069.22	-67.95
*DOW	2279.32	0.00
*SP500	1287.79	0.00

Get Quotes

Delayed Data
Provider - Disclaimer
Sponsored By
get 100% commission
Power E*TRADE

NEWS ALERTS
Get an alert when there are new stories about

Federal Reserve
 President Bush
 GDP
 Sen. Charles Schumer

Add Selected Alerts
More Alerts

ELSEWHERE ON THE WEB

CNN.COM
Dow leads Wall Street rebound

ABC NEWS
Unemployed and Tipping the Scales

USATODAY.COM
Largest teachers union, AFL-CIO, step closer

CONCERNED ABOUT YOUR CHILD'S DEVELOPMENT?
FIND OUT WHAT TO EXPECT BETWEEN AGES 3 AND 5.
MAKE INFO

IN YAHOO! FINANCE
2006 Tax Guide
Avoid tax blunders and get tips.
More From Yahoo! Finance

ALSO IN Y! NEWS
Top Tech News
The latest news, features, and more

WASHINGTON - After the booming 1990s when incomes and stock prices were soaring, this decade has been less of a thrill ride for most American families.

Average incomes after adjusting for inflation actually fell from 2001 to 2004, and the growth in net worth was the weakest in a decade, the **Federal Reserve** reported Thursday.

Many families were struggling in the aftermath of the 2001 recession and the bursting of the stock market bubble in 2000, the Fed's latest "Survey of Consumer Finances" showed. The comprehensive look at household balance sheets comes every three years.

Average family incomes, after adjusting for inflation, fell to \$70,700 in 2004, a drop of 2.3 percent when compared with 2001. That was the weakest showing since a decline of 11.3 percent from 1989 to 1992, a period that also covered a recession.

The average incomes had soared by 17.3 percent in the 1998-2001 period and 12.3 percent from 1995 to 1998 as the country enjoyed the longest economic expansion in history.

The median family income, the point where half the families made more and half made less, rose a tiny 1.6 percent to \$43,200 in 2004 compared with 2001.

Economists said the weakness in the most recent period was understandable given the loss of 2.7 million jobs from early 2001 through August of 2003, when the country was struggling with sizable layoffs caused by the recession, the terrorist attacks and corporate accounting scandals.

The weak income and the stock market decline in the early part of the decade, which wiped out \$7 trillion of paper wealth, had an adverse impact on family balance sheets.

Net worth, the difference between assets and liabilities such as loans, rose by 6.3 percent in the 2001-2004 period to an average of \$448,200, after adjusting for inflation. That gain was far below the huge increases of 25.6 percent from 1995 to 1998 and 25.6 percent from 1998 to 2001, increases that were fueled by soaring stock prices.

The 2001-2004 performance was the worst since net worth actually fell 1.5 percent in the 1989-1992 period.

The median family net worth, the point where half the families had more and half owned less, stood at \$93,100 in 2004, a rise of 1.5 percent from 2001, after adjusting for inflation from 2001.

The report showed that the slowdown in the growth of net worth would have been even more sizable except for the fact that homeowners have enjoyed big gains in the value of their homes in recent years.

The gap between the very rich and other income groups widened during the period.

The top 10 percent of households saw their net worth rise by 6.1 percent to an average of \$3.11 billion while the bottom 25 percent suffered a decline from a net worth in 2001

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.



Recommendations as information supply

amazon.com Prime Andrei's Store See All 32 Product Categories Your Account | Cart | Wish List | Help |

Improve Your Recommendations | Your Amazon Home | Your Profile | Learn More

Search Amazon.com Web Search

Recommended for Andrei Z Broder


Narrow by Event
[Your Watch List](#) (Beta)

Narrow by Category
Your Favorites

- [Books](#)
- [Electronics](#)
- [Kitchen & Housewares](#)
- [Tools & Hardware](#)
- [Software](#)
- [Computers](#)

Recommendations for you are based on [items you own](#) and more.

view: **All** | [New Releases](#) | [Coming Soon](#)

- 

[Blink : The Power of Thinking Without Thinking](#)
by Malcolm Gladwell
Publication Date: January 11, 2005

Our Price: \$15.09
[Used & new](#) from \$14.10

Rate this item I own it Not interested

Recommended because you purchased [The Wisdom of Crowds](#) and more ([edit](#))



Information supply in context

ITEM INFORMATION

Explore this item

- [buying info](#)
- [product specs](#)
- [accessories](#)
- [product description](#)
- [customer reviews](#)

See more by this manufacturer

- [Canon](#)

Discover similar items

- [in Camera & Photo](#)

Customers also bought

- [these other items...](#)

Canon Powershot SD550 7.1MP Digital Elph Camera with 3x Optical Zoom

[Other products by Canon](#)



List Price: ~~\$549.99~~
Price: **\$439.94**
You Save: \$110.05 (20%)

20% off

News and Reviews About This Product [What's this?](#)

Want to learn more? Check out news articles and reviews about this product.
(Links not working? Check your pop-up blocker.)

- [Digital Cameras - Canon PowerShot SD550 Digital Camera Review, Information, Spec ...](#) (from [www.imaging-resource.com](#))
With the PowerShot SD550 and SD450 Digital ELPH, Canon has once again stood its stylish ELPH digicams on end, packing ...
- [Canon PowerShot SD550](#) (from [reviews-zdnet.com.com](#))
The Canon PowerShot SD550 is a very well-designed, 7-megapixel snapshot camera. Just don't look for advanced controls. ...
- [DCRP Review: Canon PowerShot SD550 Digital ELPH](#) (from [dcresource.com](#))
Most of those aren't very exciting, and to be honest I don't know why Canon couldn't put the larger screen ...

The background features a large, light purple watermark of the Yahoo! logo. It consists of a purple oval containing a white 'Y' with a purple shadow, followed by a purple exclamation point and the letters 'TM' in a smaller font.

Web advertising

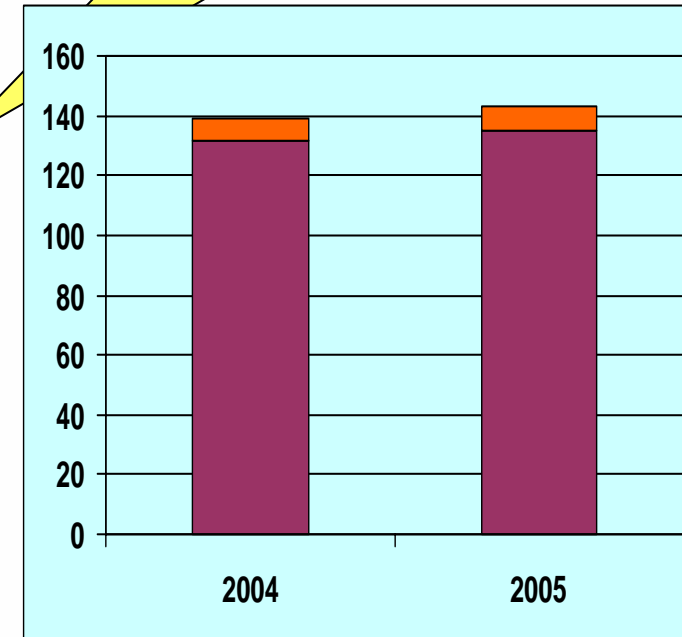


Advertising spend in USA in 2005

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%
FSIs 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

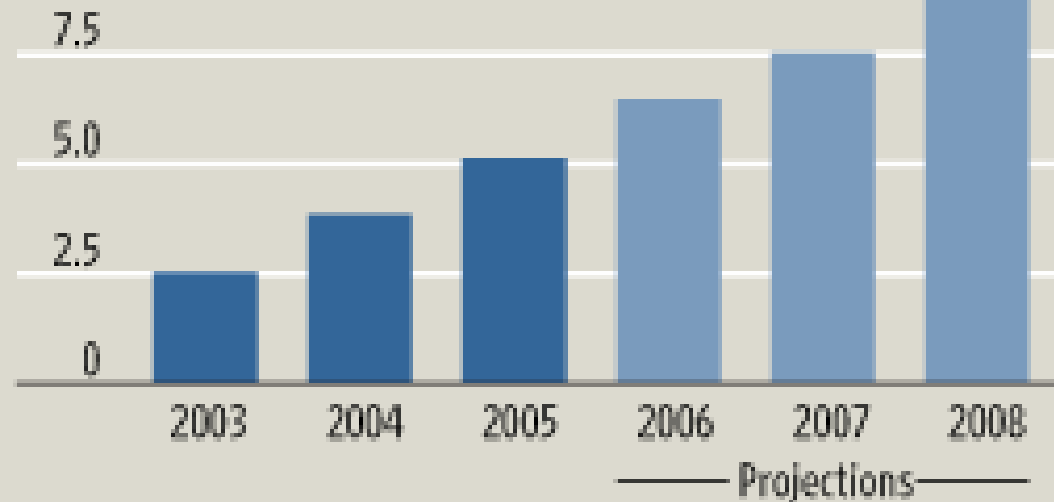


Search advertising spending

Click Costs

Spending on search advertising in the U.S.

\$10.0 billion



Source: eMarketer



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



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

Ads by Yahoo!

Olympus Digital Cameras - Official...
Visit the official Olympus Web site for comprehensive information about our digital cameras and...
www.olympusamerica.com

Latin American Art Galleries Online
Visit Latin American online art galleries. Features images and detailed information. Read...
www.artnexus.com

Discount Prices on Sony Digital...
5-star CNET service rating - find low prices today at BeachCamera.com.
www.beachcamera.com

General

Forum	Topics	Posts
Welcome - come on in! Hi - are you new? Come in and let everyone here know about you. Moderator shanky_pec	29	123
Announcements, Feedback & Suggestions Check this section for board updates, announcements of new features. Leave your kind feedback, suggestions here.. Moderator shanky_pec	18	67
Site Support Report site related problems, broken links, not-working-features here. We will surely resolve problems Moderator shanky_pec	5	18

Photography

Forum	Topics	Posts
General Discussion General photography discussions Moderator shanky_pec	64	284
Photo Album Discussion, tips-tricks, suggestions regarding photo gallery Moderator shanky_pec	26	129



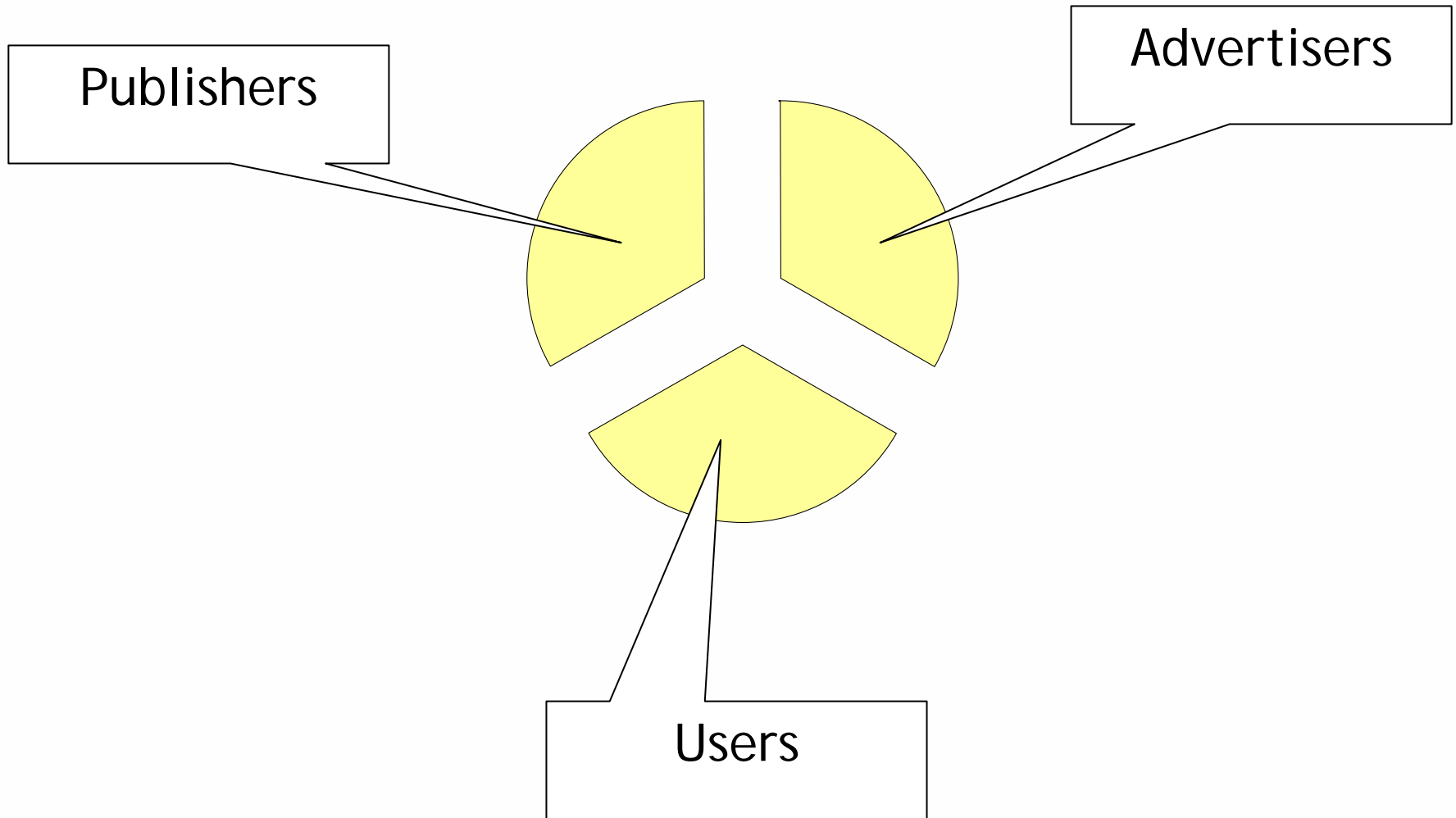
Content match example (II)

The screenshot shows the MSN Entertainment website with an "Artist Spotlight" for J.S. Bach. Several red callout boxes highlight sponsored content:

- Entertainment**: A box pointing to the MSN Entertainment logo.
- J.S. Bach**: A box pointing to the artist's name in the spotlight section.
- Sponsored Sites**: A large box containing three sponsored advertisements:
 - Music by J. S. Bach at Amazon.com**: Amazon.com has a huge selection of merchandise, including CDs, 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, and more customers, 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.



Contextual ads = meeting of Publishers, Advertisers, Users





Quality of matching

STLtoday.com ST LOUIS POST-DISPATCH Thursday, April 29, 2004

Subscribe to the St. Louis Post-Dispatch

HOME NEWS BUSINESS SPORTS ENTERTAINMENT LIFE & STYLE JOBS AUTOS REAL ESTATE AD ZONE NEIGHBORHOODS

#1 ST. LOUIS WEB SITE SITE SEARCH Go! STORY FINDER ADZONE

News > Nation > Story

Bush policies endanger rights, protesters say

By Bernard Mallee
Post-Dispatch Washington Bureau
04/25/2004



WASHINGTON - Armed with bold placards and bright-colored banners, abortion-rights supporters from Missouri and Southern Illinois marched in step with throngs of other protesters Sunday to rail against what they see as an erosion of

YESTERDAY'S MOST E-MAILED STORIES

- "Movin' Out" reinvents musical theater
- Dear Kurt: Da Bears need you
- Seldom are heard discouraging words, "You're fired"
- Washington U. newspaper publishes images of frat party
- Keeping Chief Miniwek could cost U of I funding
- Man is killed for stopping car theft, police say
- Injecting venture capital

Postcards from Iraq



Quality of matching

Ads by Google

[Feminist Fantasies](#)

Essays on feminism in the media, workplace, home, and the military.

www.eagleforum.org

[Borgen Project Meetups](#)

We be changing things! Help make poverty a national priority.

borgenproject.org

[Check Doctor Quality](#)

Get critical quality and experience data on current or new physicians.

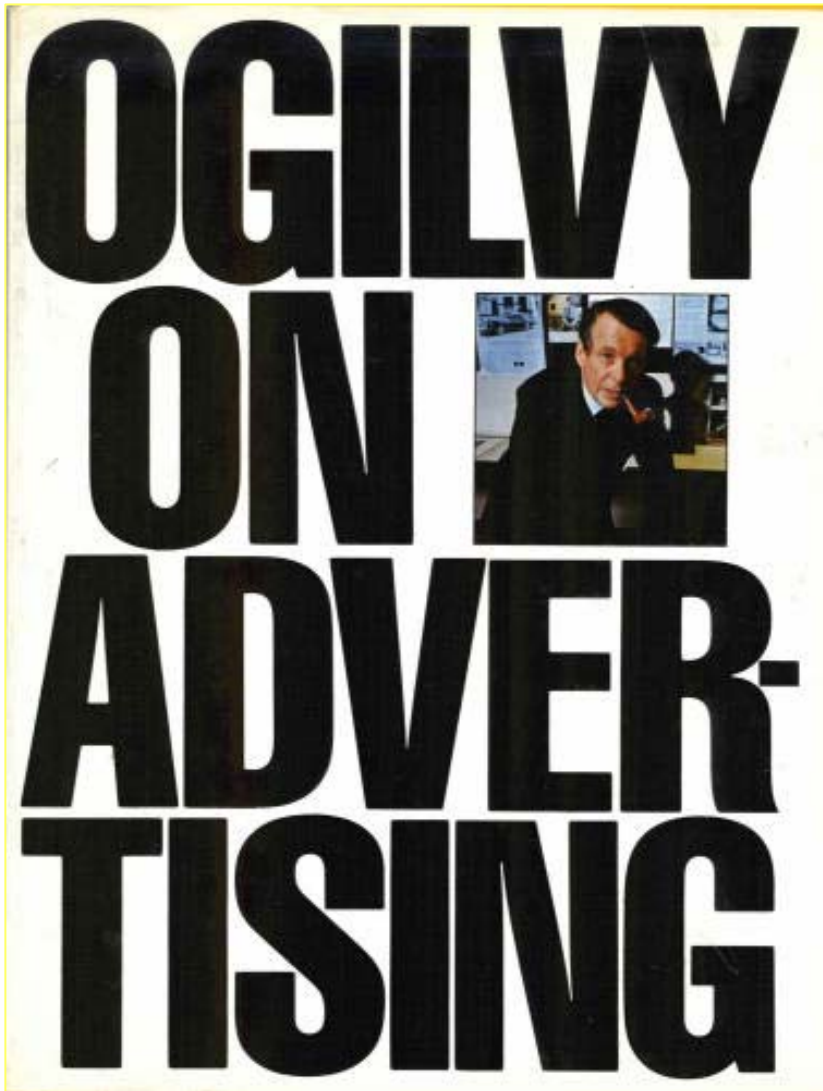
www.HealthGrades.com

[Top of page](#)

[E-mail this page to a friend](#)



Famous book:



by David Ogilvy of
Ogilvy & Mather
[1983]

Recently in paperback!

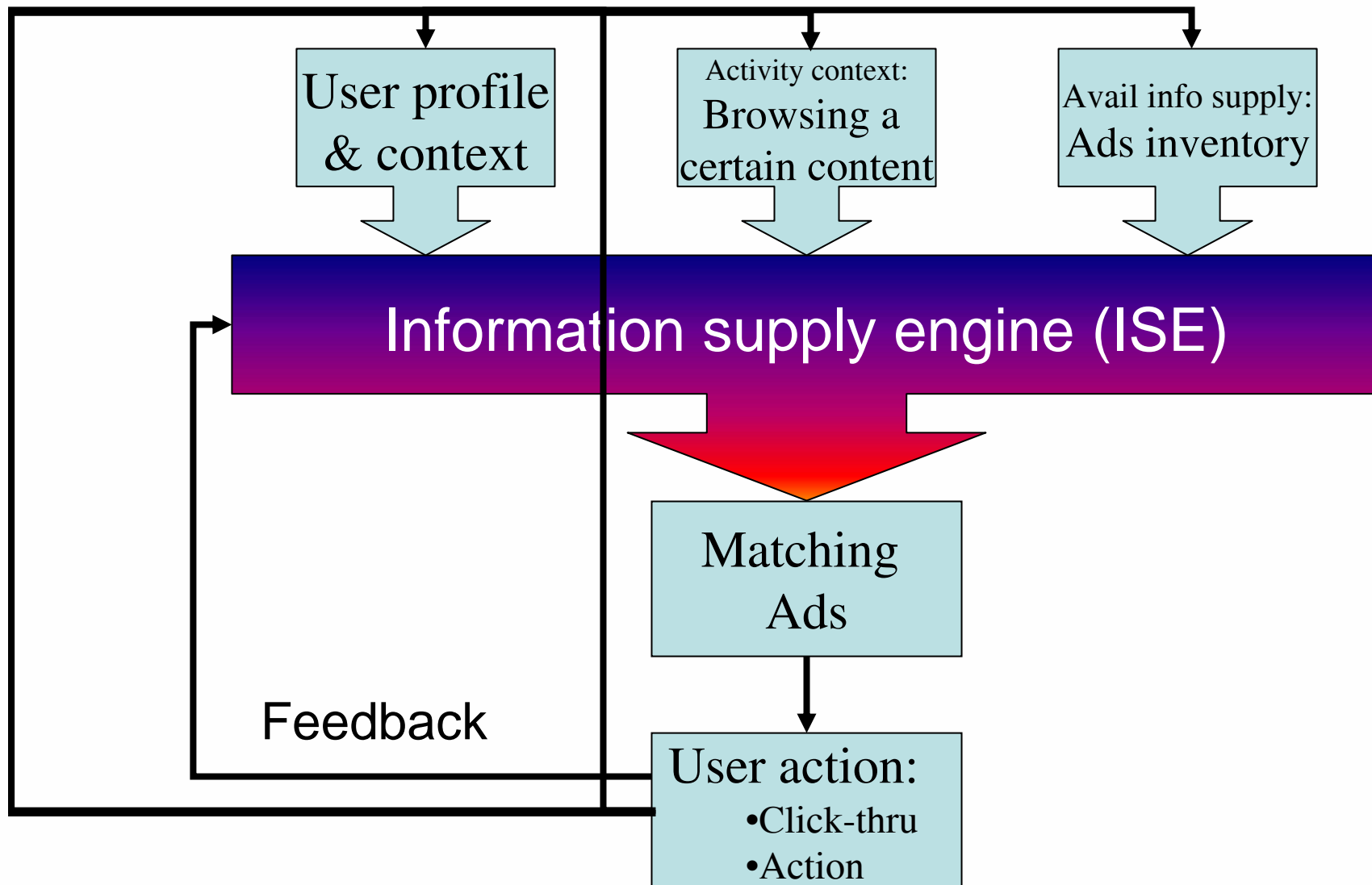


Ads as information supply

- “I do not regard advertising as entertainment or an art form, but as a medium of information....” [David Ogilvy, 1985]
- “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!



LIFE ENGINE™