

# Emerging Sciences of the Internet: Some New Opportunities



**Ron Brachman**

**Vice President,  
Worldwide Research Operations**

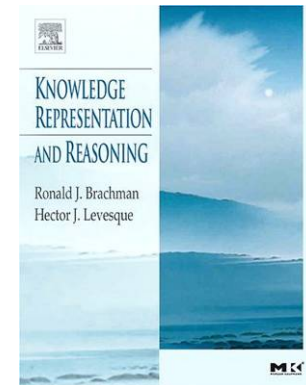
**YAHOO!**  
RESEARCH

4<sup>th</sup> European Semantic Web Conference  
Innsbruck, Austria  
June 5, 2007



# Parallel Paths

- Knowledge representation meets the Web
  - Berners-Lee/Hendler/Lassila *Scientific American* article (2001): “For the semantic web to function, computers must have access to structured collections of information and sets of inference rules that they can use to conduct automated reasoning” – i.e., knowledge representation (KR) technology
    - “Adding logic to the Web”
- Ron meets the Web
  - Long career focused on KR
  - Joined Yahoo! in late 2005





# Today's Remarks

- My original thought about this talk: to show all the great ways the semantic web is transforming the industry
  - and why Yahoo! needs the SW
  - and why KR is great
  - and how we can all profit wildly...
- But...

Major web company imperatives seem to be taking them in a different direction

Key thought from ice hockey\*: the SW community should “not skate to where the puck is, but rather to where it will be”

\*From Wayne Gretzky's father



# Skating to Where the Puck Will be

- I won't be preaching about Web 2.0, 3.0, 3.1416, 98.6 or any other future release, but the ideas behind Web 2.0 are making a huge impact in the industry
- In Yahoo! Research, we're focusing on the deep, underlying knowledge needed to move ahead – the sciences behind Web n.0
- I think there are interesting opportunities here...

See Mark Greaves' article in  
March/April *IEEE Intelligent Systems*





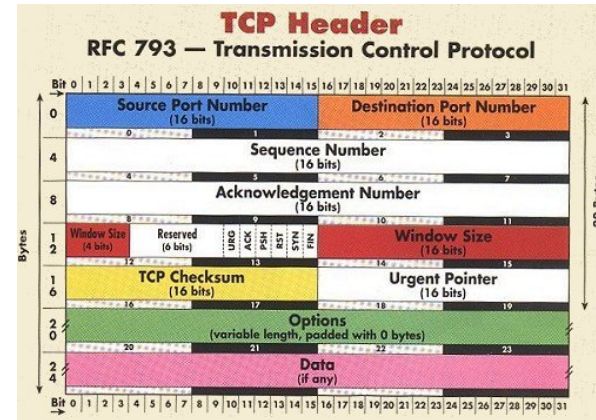
# The Internet

- The **Internet** is a worldwide, publicly accessible network of interconnected computer networks that transmit data by packet switching using the standard Internet Protocol (IP). It is a "network of networks" that consists of millions of smaller domestic, academic, business, and government networks, which together carry various information and services (*Wikipedia*)



# From Technical Infrastructure...

- Started with the invention of the concept of packet-switching and a simple protocol
- ARPANet: focus on collaboration among researchers
  - ftp
    - File transfers
  - telnet
    - Remote logins
  - Email
    - Simple asynchronous text communication

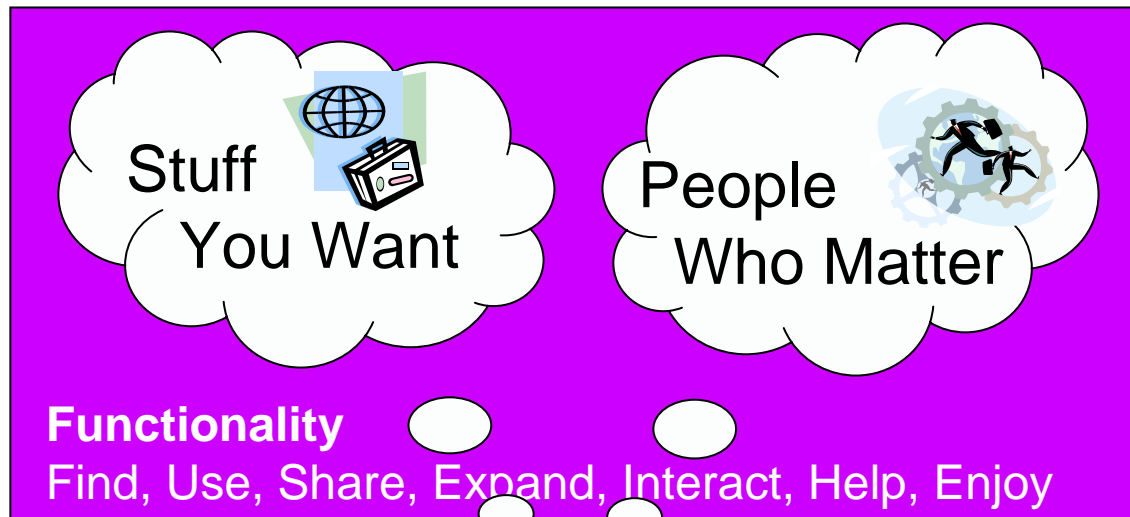




# ...to an Interconnected World

- A distributed, heterogeneous repository of tools, data, and people – all kinds of people

## The user-centric view



## Example views:

- searching/finding
- communicating (email, IM)
- socializing/sharing (groups, communities)
- buying/selling
- matchmaking/recommending
- experiencing media



“Stuff” = data/knowledge, help, things, entertainment



# Far, Far Away from ftp and telnet

- Instant messaging
- Search engines
- E-commerce
- Social media
- Networked games
- Advertising
- Podcasting, UG-video
- Blogging
- Wikis
- VOIP telephony
- The mobile Web
- TB's of data

YAHOO!

Google™

amazon.com.

ebay

Vonage

SECOND LIFE

del.icio.us  
social bookmarking

[ thefacebook ]

myspace.com  
a place for friends

YouTube

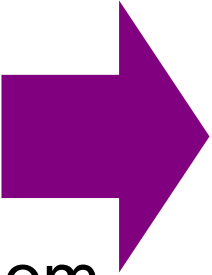
iTunes

AlphaWorld





# Some Buzzwords to Watch for Concepts

- Fragmentation
  - Relationships
  - User-generated
  - Social media/"wisdom of crowds"
  - Rich media
  - CPC/CPM/CTR
- 
- Finding/answering
  - Matching
  - Object identity
  - Integration/synthesis/mashup/retrofit
  - Lightweight

*These concepts are driving large open-Internet providers*

The world has changed from one where the primary concepts were TCP headers, DNS, and telnet, to one where the watchwords are things like "Mashup" and "StumbleUpon"



# We Need to Move From the “Old” CS

- Hardware
- Computer Systems Organization
- Software
- Data
- Theory of Computation
- Mathematics of Computing
- Information Systems
- Computing Methodologies
- Computer Applications
- Computing Milieux

```
A. General Literature
B. Hardware
C. Computer Systems Organization
  C.0 GENERAL
  C.1 PROCESSOR ARCHITECTURES
  C.2 COMPUTER-COMMUNICATION NETWORKS
  C.3 SPECIAL-PURPOSE AND APPLICATION-BASED SYSTEMS (J.7)
  C.4 PERFORMANCE OF SYSTEMS
  C.5 COMPUTER SYSTEM IMPLEMENTATION
  C.m MISCELLANEOUS
D. Software
  D.0 GENERAL
  D.1 PROGRAMMING TECHNIQUES (E)
  D.2 SOFTWARE ENGINEERING (K.6.3)
  D.3 PROGRAMMING LANGUAGES
  D.4 OPERATING SYSTEMS (C)
  D.m MISCELLANEOUS
E. Data
  E.0 GENERAL
  E.1 DATA STRUCTURES
  E.2 DATA STORAGE REPRESENTATIONS
  E.3 DATA ENCRYPTION
  E.4 CODING AND INFORMATION THEORY (H.1.1) ]
  E.5 FILES (D.4.3, F.2.2, H.2)
  E.m MISCELLANEOUS
F. Theory of Computation
G. Mathematics of Computing
H. InFormation Systems
  H.0 GENERAL
  H.1 MODELS AND PRINCIPLES
  H.2 DATABASE MANAGEMENT (E.5)
  H.3 INFORMATION STORAGE AND RETRIEVAL
    H.3.0 General
    H.3.1 Content Analysis and Indexing
    H.3.2 Information Storage
    H.3.3 Information Search and Retrieval
    H.3.4 Systems and Software
    H.3.5 Online Information Services
      Commercial services
      Data sharing
      Web-based services
    ...
  H.4 INFORMATION SYSTEMS APPLICATIONS
    ...
    H.4.3 Communications Applications
      Bulletin boards
      Computer conferencing, teleconferencing, and videoconferencing
      Electronic mail
      Information browsers
      Videotex
    H.4.m Miscellaneous
  H.5 INFORMATION INTERFACES AND PRESENTATION (e.g., HCI) (I.7)
    H.5.0 General
    H.5.1 Multimedia Information Systems
    H.5.2 User Interfaces (D.2.2, H.1.2, I.3.6)
    H.5.3 Group and Organization Interfaces
      Asynchronous interaction
```

[ACM Computing Classification System]

- Collaborative computing
- Computer-supported cooperative work
- Evaluation/methodology
- Organizational design
- Synchronous interaction
- Theory and models
- Web-based interaction
- H.5.4 Hypertext/Hypermedia (I.7, J.7)
- H.5.5 Sound and Music Computing (J.5)
- ...
- I. Computing Methodologies
  - I.0 GENERAL
  - I.1 SYMBOLIC AND ALGEBRAIC MANIPULATION
  - I.2 ARTIFICIAL INTELLIGENCE
    - I.2.0 General
      - I.2.1 Applications and Expert Systems (H.4, J)
        - Cartography
        - Games
        - Industrial automation
        - Law
        - Medicine and science
        - Natural language interfaces
        - Office automation
      - I.2.2 Automatic Programming (D.1.2, F.3.1, F.4.1)
      - I.2.3 Deduction and Theorem Proving (F.4.1)
      - I.2.4 Knowledge Representation Formalisms and Methods (F.)
      - I.2.5 Programming Languages and Software (D.3.2)
      - I.2.6 Learning (K.3.2)
        - Analogies
        - Concept learning
        - Connectionism and neural nets
        - Induction
        - Knowledge acquisition
        - Language acquisition
        - Parameter learning
      - I.2.7 Natural Language Processing
      - I.2.8 Problem Solving, Control Methods, and Search (F.2.2)
      - I.2.9 Robotics
      - I.2.10 Vision and Scene Understanding (I.4.8, I.5)
        - 3D/stereo scene analysis
        - Architecture and control structures [\*\*]
        - Intensity, color, photometry, and thresholding
        - Modeling and recovery of physical attributes
        - Motion
        - Perceptual reasoning
        - Representations, data structures, and transforms
        - Shape
        - Texture
        - Video analysis
      - I.2.11 Distributed Artificial Intelligence
      - I.2.m Miscellaneous
    - I.3 COMPUTER GRAPHICS
    - I.4 IMAGE PROCESSING AND COMPUTER VISION
    - I.5 PATTERN RECOGNITION
    - I.6 SIMULATION AND MODELING (G.3)
    - I.7 DOCUMENT AND TEXT PROCESSING (H.4, H.5)
      - I.7.0 General
      - I.7.1 Document and Text Editing
      - I.7.2 Document Preparation
        - Desktop publishing
        - Format and notation
        - Hypertext/hypermedia
        - Index generation
        - Languages and systems
        - Markup languages
        - Multi/mixed media
        - Photocomposition/typesetting
        - Scripting languages
        - Standards
      - I.7.3 Index Generation [\*\*]
      - I.7.4 Electronic Publishing (H.5.4, J.7)
      - I.7.5 Document Capture (I.4.1)
      - I.7.m Miscellaneous
  - I.m MISCELLANEOUS
- J. Computer Applications
- K. Computing Milieux
  - K.0 GENERAL
  - K.1 THE COMPUTER INDUSTRY
  - K.2 HISTORY OF COMPUTING
  - K.3 COMPUTERS AND EDUCATION
  - K.4 COMPUTERS AND SOCIETY
    - ...
    - K.4.4 Electronic Commerce (J.1)
      - Cybercash, digital cash
      - Distributed commercial transactions
      - Electronic data interchange (EDI)
      - Intellectual property
      - Payment schemes
      - Security
    - K.4.m Miscellaneous
  - K.5 LEGAL ASPECTS OF COMPUTING
  - K.6 MANAGEMENT OF COMPUTING AND INFORMATION SYSTEMS
    - K.6.0 General
      - K.6.1 Project and People Management
      - K.6.2 Installation Management
      - K.6.3 Software Management (D.2.9)
      - K.6.4 System Management
      - K.6.5 Security and Protection (D.4.6, K.4.2)
        - Authentication
        - Insurance [\*\*]
        - Invasive software (e.g., viruses, worms, Trojan horses)
        - Physical security [\*\*]
        - Unauthorized access (e.g., hacking, phreaking)
    - K.6.m Miscellaneous
      - Insurance [\*]
      - Security [\*]
  - K.7 THE COMPUTING PROFESSION
  - K.8 PERSONAL COMPUTING
  - K.m MISCELLANEOUS



## ...to the New

- “Finding” science and systems
- Community science and systems
- Algorithmic advertising
- Computational micro-economics
- Media experiences
- Data science

**We need to bring together existing disciplines...  
...and nurture the creation of new ones**



*This talk:*

A high-level view of some new areas of thinking, key issues for consideration, and research challenges



# Towards a Science of Finding

## Moving beyond “search”

- More than just classical IR
  - Structure
  - Intent
  - Context
- Moving to an information supply model
  - From query-based retrieval to context-driven supply
- Social search
  - People, not search engines



# 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)

1994-1997 AV,  
Excite, Lycos, etc

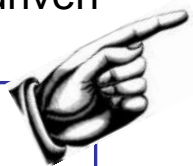
1998 to now. Made  
popular by Google  
but everyone now

- **Third generation:** answer “the need behind the query”
  - *Semantic analysis*
  - *Integrate multiple sources of data*
  - Context determination
    - spatial (user location/target location), query stream (previous queries), personal (user profile), geographic, etc
  - Help the user
    - UI, spell checking, query refinement, query suggestion, syntax driven feedback, context help, context transfer, etc
  - Integration of search and text analysis

Still evolving

I will call out potentially interesting opportunities for SW technology like this

...and some difficult challenges like this





# User Needs

## Results, not Documents

- **Informational** – want *to learn* about something (~40% / 65%)
  - Low hemoglobin
- **Navigational** – want *to go* to a page (~25% / 15%)
  - Austrian Airlines
- **Transactional** – want *to do something* (web-mediated) (~35% / 20%)
  - Access a service
    - Innsbruck weather
  - Downloads
    - MP3 to WMA converter
  - Shop
    - Canon SD550
  - VOIP companies
- Gray areas
  - Find a good hub
  - Exploratory search to “see what’s there”



# Third Generation Search

- Focus on the user need and answer it directly
  - Use text analysis, semantic analysis
    - Hard and “soft” (partial) matches
    - Triggers, e.g., stock symbols, place names, people names, drug names
    - Query language determination
- Deal with transactional queries





# Yahoo!: britney spears

Web | Images | Video | Directory | Local | News | Shopping

**YAHOO!** SEARCH

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](#)



# Yahoo!: salvador hotels

Web | Images | Video | Directory | Local | News | Shopping

**YAHOO!** SEARCH  Search

My Web BETA

Search Results Results 1 - 10 of about 17,900,00

**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](#)

1. [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 [Yahoo! Travel](#).  
[travel.yahoo.com/p-hotel-482915-san\\_salvador\\_hotels-i](#) - [More from this site](#) - [Save](#) - [Block](#)
2. [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”

**Yahoo! Shortcuts**

**Shortcuts Categories**

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

 <b><u>Local</u></b>	 <b><u>News &amp; Information</u></b>	 <b><u>Travel</u></b>	 <b><u>Reference</u></b>	 <b><u>Calculators</u></b>
<a href="#">Maps</a> <a href="#">Weather</a> <a href="#">Local Listings</a> <a href="#">more...</a>	<a href="#">Sports Scores</a> <a href="#">Stock Quotes</a> <a href="#">Images</a> <small>NEW!</small> <a href="#">more...</a>	<a href="#">Airport Information</a> <a href="#">Flight Tracker</a> <a href="#">Hotel Finder</a> <a href="#">more...</a>	<a href="#">Dictionary</a> <a href="#">Definitions</a> <a href="#">Encyclopedia</a> <a href="#">Lookup</a> <a href="#">Synonym Finder</a> <a href="#">more...</a>	<a href="#">Calculator</a> <a href="#">Time Zones</a> <a href="#">Weights and Measures</a> <a href="#">Converter</a> <a href="#">more...</a>



# Google:

andrei broder new york

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

andrei broder new york

**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](http://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 ...



# Third Generation Search Challenges

- How better to take advantage of
  - *Structure*
  - User intent
  - Context
  - The Web's role as a *delivery vehicle for services* and applications



# Current Simple View



Keyword search: seafood san francisco



Buy *San Francisco Seafood* at Amazon



*San Francisco Seafood Cookbook/Guide*



# Structure



“seafood san francisco”



Category: restaurant  
Location: San Francisco



Reserve a table for two tonight at SF’s best Sushi Bar and get a free sake, compliments of OpenTable!

Category: restaurant Location: San Francisco



[Alamo Square Seafood Grill](#) - (415) 440-2828  
803 Fillmore St, San Francisco, CA - 0.93mi - [map](#)

Category: restaurant Location: San Francisco



# Finding Structure



“seafood san francisco”



**CLASSIFIERS**  
(e.g., SVM)

**Category:** restaurant  
**Location:** San Francisco

- Can apply machine learning (ML) to extract structure from **user context** (query, session, ...), **content** (web pages), and **ads**
- Alternative: We can elicit structure from users in a variety of ways



**Problem:** it often won't be pre-specified





# Google Base

http://www.google.com/search?q=toyota+corolla+palo+alto&start=0&ie=utf-8&oe=utf-8&client=f Go toyota corolla palo alto

[Sign in](#)



[Web](#) [Images](#) [Video](#) <sup>New!</sup> [News](#) [Maps](#) [more »](#)

toyota corolla palo alto

Search

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

**Web** Results 1 - 10 of about 126,000 for **toyota corolla palo alto**. (0.23 seconds)

Refine your search for [toyota corolla palo alto](#)

Location

Make

Model

palo alto

toyota

corolla

Go

Remember this location

[Automobiles](#)

**TOYOTA Corolla** CE 2001. Sedan 4D. Stock# 46836A \$8588  
Honda of Stevens Creek ... Professional **Palo Alto** Dental Office.  
Looking for a fun professional person ...  
[www.paloaltdailynews.com/class.html](http://www.paloaltdailynews.com/class.html) - 66k -  
[Cached](#) - [Similar pages](#)

[Toyota of Palo Alto Palo Alto California, New Car Research, Quotes ...](#)

**Toyota of Palo Alto** new car pricing, free quotes, photos, specs,  
... 2006 **Toyota Corolla**, Download 2006 **Toyota Corolla**  
Brochure in PDF format · View 2006 ...  
[www.tundrasolutions.com/dealers/dealer.php/id/159](http://www.tundrasolutions.com/dealers/dealer.php/id/159) - 48k -  
[Cached](#) - [Similar pages](#)

[PALO ALTO / New housing complex on drawing board / 100 townhomes ...](#)

**PALO ALTO** <BR> New housing complex on drawing board <BR>

Sponsored Links

[Compare Dealer Prices](#)

See dealer and MSRP prices.  
Free car quotes from local dealers.  
[www.Autobytel.com](http://www.Autobytel.com)

[Palo Alto Toyota Dealers](#)

**Toyota** Time is Back - Starts 8/06!  
Get Great Deals on New 2007 Models.  
[www.Toyota.com](http://www.Toyota.com)

[Toyota palo alto](#)

**Toyota** Marin. We are a local dealer  
with internet pricing. Get a quote.  
[www.Toyota-Marin.com](http://www.Toyota-Marin.com)  
San Francisco-Oakland-San Jose, CA

[New Toyotas Quotes](#)

Local Motor Trend New Car Pricing  
Our 60-Second Form is Quick & Easy!  
[MotorTrend.com](http://MotorTrend.com)



# Y! Shortcuts

Search Results 1 - 10 of about 28,400 for toyota corolla palo alto - 0.05 sec. (About this page)

- [Putnam Toyota: Toyota Corolla](#)  
www.putnamtoyota.com | Local Info The Putnam Automotive Group has more than 37 years experience serving the public, and we invite you to become yet another satisfied customer.
- [Buy a New Toyota Corolla below Invoice](#)  
cbi.carsbelowinvoice.com Why pay sticker? You can buy a new car below invoice.

### SPONSOR RESULTS

### SPONSOR RESULTS

[How Much is Your Car Worth at Trade-In?](#)  
Get the trade-in value for your old car and save thousands on a new...  
tiv.trade-in-value.co

- [Used Toyota Corolla Results near Palo Alto, CA - Map All Results](#)  
1999 [Toyota Corolla](#) - \$10,000 - 81,571 miles  
2004 [Toyota Corolla](#) - \$12,991 - 28,158 miles  
2002 [Toyota Corolla](#) - \$9,994 - 54,860 miles  
[Yahoo! Shortcut](#) - [About](#)

[Used Toyota Corollas at AutoExtra.com](#)  
Come to AutoExtra.com for used **Toyota Corolla** and over a million...  
autoextra.com

**Yahools:** [Report](#) bad results or ads. Bucket test: [NONE](#)

1. [Find Used TOYOTA COROLLA Cars In Palo Alto, California on Yahoo!](#)  
Find used **TOYOTA COROLLA** cars in **Palo Alto**, California on Yahoo! Autos. Review used **TOYOTA COROLLA** pictures, price, options and features. Find the best deal and respond to the listings online immediately.  
[autos.yahoo.com/used-cars/ca-palo\\_alto-toyota-corolla.html](#) - [More from this site](#) - [Save](#)
2. [Magnussen's Toyota & Scion of Palo Alto](#)  
New and used cars, parts and service.  
Category: [California](#) > [Palo Alto](#) > [Automotive Dealers](#) > [Toyota](#)

[New Toyota Corolla Price Quotes - Free](#)  
Save thousands on your new car with a



# From Information Retrieval to Information Supply

Explicit demand for information driven by a user query

Active information supply driven by user activity and context

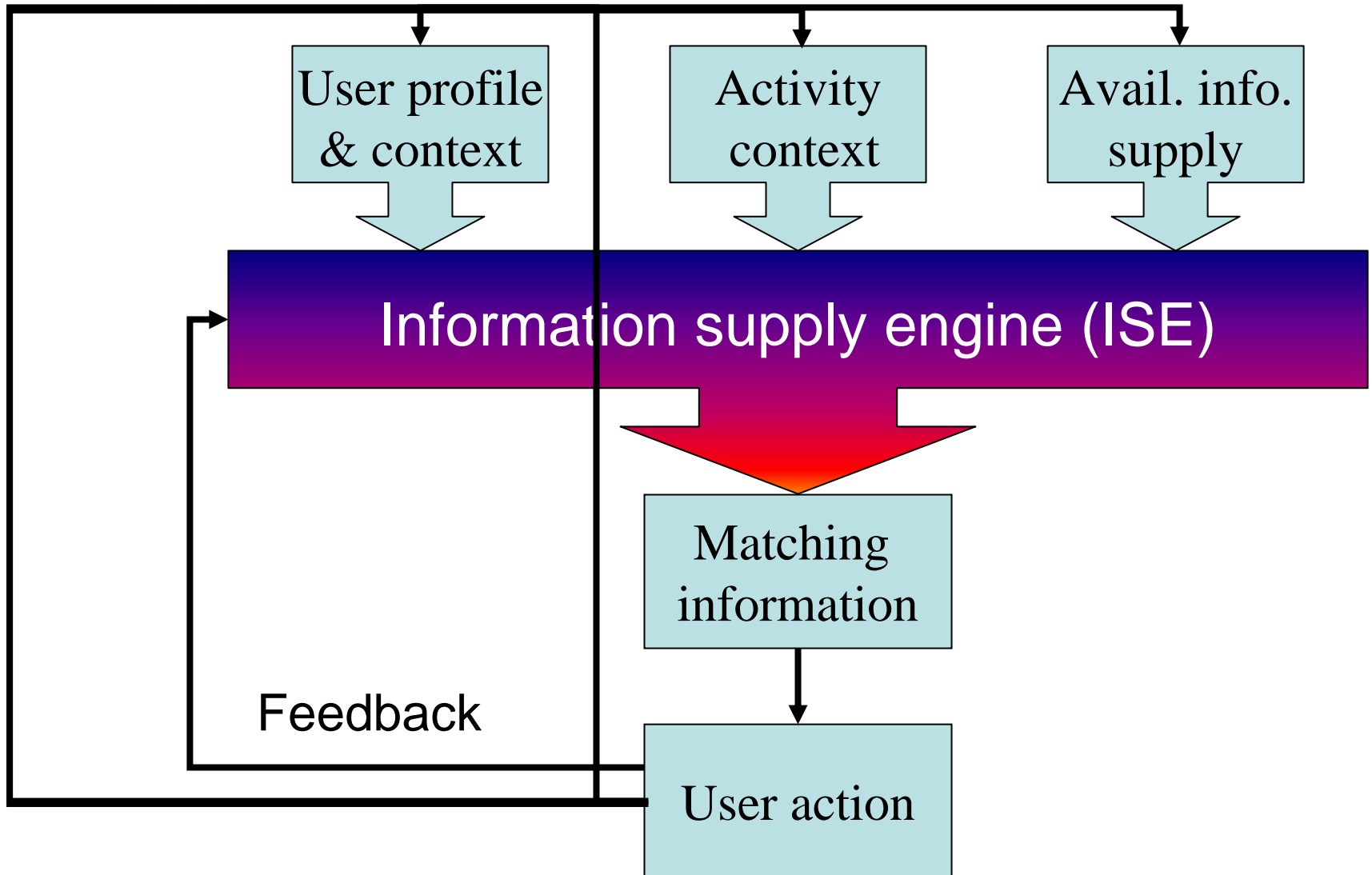


Now...





# Information Supply Architecture





# Some Current Approaches

## Alerts and Subscriptions

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

andrei broder

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

### News Results

NEWS STORIES Results 1 - 3 of about 3 for andrei broder. Sort Results by: Relevance | [Date](#)

**1. Q&A with Andrei Broder, Yahoo VP of Emerging Search Technology**  
[Search Engine Journal](#) - Jan 30 9:23 AM  
Q&A with Andrei Broder, Yahoo VP of Emerging Search Technology The Yahoo Search Blog is opening up its channel of communication with Andrei Broder, Yahoo's VP of Emerging Search Technology and accredited expert on design, analysis, and implementation of algorithms for information retrieval and applications. From Tara @ Yahoo : "Broder is co-winner ...  
[Save to My Web](#)

**2. Andrei Broder Taking Questions At Yahoo**  
[WebProNews](#) - Jan 27 4:32 PM  
Visitors to the Yahoo Search Blog can post questions for Broder, who recently joined the company to serve as a research fellow and vice-president of emerging

### NEWS ALERTS

Alert me when articles match all these words:  
[andrei broder](#)

ADD TO MY YAHOO! / RSS

Add your news search for [andrei broder](#) to My Yahoo!

View as RSS: [XML](#)

## YAHOO! ALERTS

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

email

Yahoo! Messenger

mobile device

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





# Annotations

Yahoo! My Yahoo! Mail

Welcome, **axbroder**  
[Sign Out, My Account]

**YAHOO! NEWS** News Home - 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
*GSPC	1287.79	0.00

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 28.7 percent from 1990 to 2001, increases that were fueled by soaring stock prices.

The 2001-2004 performance was the worst since net worth actually declined by 1.1 percent in the 1989-1992 period.

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

The report showed that the slowdown in the economy and the loss of jobs would have been even more sizable except for the fact that the top 10 percent of households have enjoyed big gains in the value of their homes in recent years.

The gap between the very wealthy 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

ADVERTISEMENT

Refinance \$200,000 for as low as \$833 a month\*

Refinance today with LendingTree®

See how much you can save!

Learn More

LendingTree®  
www.LendingTree.com

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


"This is the continuing story of the rich getting richer," said David Wyss, chief economist at Standard & Poor's in New York. "Clearly, the gains in wealth are going to the top end."

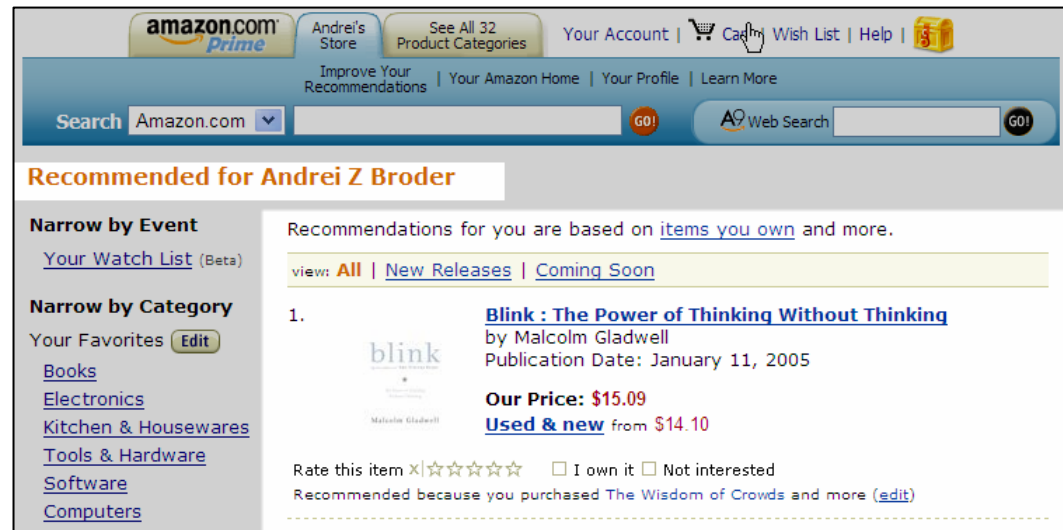
Democrats used the new report to blast **President Bush**'s economic policies, contending it would be wrong to make permanent his tax cuts which primarily benefited the wealthy.

"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.



# Others

- “Interested in the following accessories for your iPod?:....”
- “Now that you’ve booked your flight, would you like to rent a car?”
- “You seem to be writing a business letter...” 
- Contextual and search-driven ads
- Recommendations



The screenshot shows the Amazon.com interface. At the top, there are navigation links for 'amazon.com Prime', 'Andre's Store', 'See All 32 Product Categories', 'Your Account', 'Cart', 'Wish List', and 'Help'. Below this is a search bar with 'Amazon.com' entered and a 'GO' button. To the right is a 'Web Search' bar with a 'GO!' button. The main content area is titled 'Recommended for Andrei Z Broder'. It includes a 'Narrow by Event' section with a link to 'Your Watch List (Beta)'. Below that is a 'Narrow by Category' section with links for 'Books', 'Electronics', 'Kitchen & Housewares', 'Tools & Hardware', 'Software', and 'Computers'. The main recommendation is for the book 'Blink: The Power of Thinking Without Thinking' by Malcolm Gladwell, published on January 11, 2005. The price is listed as '\$15.09' with a note 'Used & new from \$14.10'. There is a star rating system and a checkbox for 'I own it' or 'Not interested'. A note at the bottom says 'Recommended because you purchased The Wisdom of Crowds and more (edit)'.



# More Technical Challenges

- 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
- Models for incorporation of feedback and mixed-initiative control



Key issue: non-categorical information





# Adding People to the Mix

http://flickr.com/photos/tags/jaguar/clusters/

Home | Sign Up Sign In | Help

Photos: [Explore Flickr](#) \* [Learn More](#)

**flickr** BETA

Tags / [jaguar](#) / [clusters](#)

(Or, try an [advanced search](#).)

[car](#), [cars](#), [auto](#), [etype](#), [automobile](#), [classic](#), [vintage](#), [autoshow](#), [red](#), [show](#)

→ [See more in this cluster...](#)

[zoo](#), [animal](#), [cat](#), [animals](#), [bigcat](#), [seattle](#), [woodlandparkzoo](#), [sleep](#), [edinburgh](#), [caged](#)

→ [See more in this cluster...](#)

[guitar](#), [fender](#)

→ [See more in this cluster...](#)

[aircraft](#), [raf](#)

→ [See more in this cluster...](#)

These are the most recent photos tagged with [jaguar](#). [See more...](#)

“Social search”



# The Power of Social Media

- Flickr – community phenomenon
- Millions of users share and tag each others' photographs (interesting question: why?)
- The “wisdom of the crowd” can be used to search
- The principle is not new – anchor text used in “standard” search

But tags are

- Structureless
- Organic
- Uncontrolled/able



"Social Search"

# YAHOO! ANSWERS

[Sign In, My Account]

Home - Help - Forum  
What's going on in Answers?

## ask.

Ask a question on any topic and get answers from real people.

(you have 110 characters to work with)

Post Question

## answer.

Share what you know and you might make someone's day.

Featured Question



Can I buy tile to match my 1950s-era kitchen countertop?

## discover.

10 million answers and counting. Learn something new today.

Featured Topic



See what people are asking about in:

Travel

Search Yahoo! Answers:

Search

Advanced

My Q&A

10 million answers. [Thanks to all the world's Answerers.](#)

Vast breadth of coverage



### Categories

- Arts & Humanities
- Business & Finance
- Cars & Transportation
- Computers & Internet
- Consumer Electronics

Share what you know. Answer open questions.

[How do you get a bleach spot out of your pants?](#)

Asked by [girly\\_antonist](#) - [Cleaning & Laundry](#) - 1 second ago

[is there a home remedy forgetting rid of ants out side with out harming my dogs or plants?](#)

Asked by [cheetarajade](#) - [Other - Home & Garden](#) - 2 seconds ago

[which poker site is the most profitable for a tournament player?](#)

Asked by [judas](#) - [Card Games](#) - 12 seconds ago

[what is the website where you can play all those games? there is a new TV commercial about it....?](#)



# Social Search Challenges

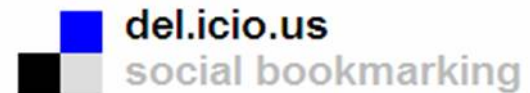
- Improve web search by
  - Learning from shared community interactions, and leveraging community interactions to create and refine content
    - Enhance and amplify user interactions
  - Expanding search results to include sources of information (e.g., experts, sub-communities of shared interest)
- Big issues
  - Finding answers
  - *Coalescing answers*
  - Incentives for contributions
  - Coping with spam
  - Irrational crowds

Reputation, Quality, Trust, Privacy



# Community Science and Systems

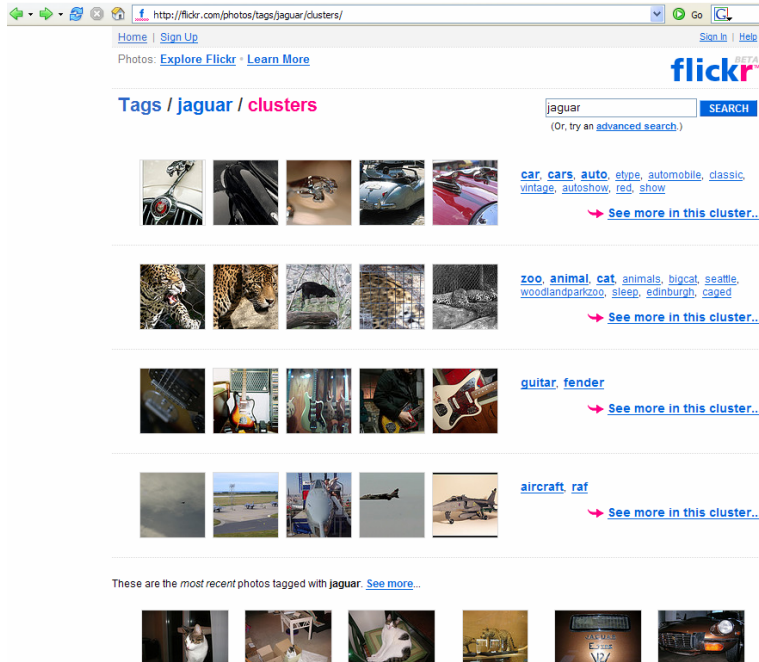
- “Social Media”
  - Different types of communities with different needs
  - Sociology, Network Theory
- Distributed/collaborative systems
  - Marriage of automatic extraction with structured augmentation
  - Imposing unifying views







# The Power of Community



Millions of users share and tag each others' photographs

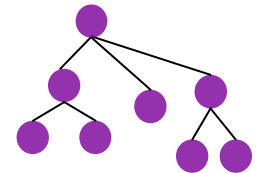
Surpassed 60 million users in twelve months

Yahoo! Groups, MySpace, FaceBook have tens of millions of users



# Understanding Online Communities

- Exploratory analysis of massive datasets of online community usage
- Challenges
  - Analyze evolving social networks of users, content and interactions to
    - Learn models of community structure and dynamics
    - Learn models of individual characteristics and preferences
    - Develop robust frameworks for evolution of authority and trust





# Four Types of Communities

## Social Networks

Communication &  
Expression

*Facebook, MySpace*

360/Groups

## Enthusiasts / Affinity

Hobbies & Interests

*Fantasy Sports,  
Custom Autos*

Music

## Knowledge Collectives

Find answers & acquire  
knowledge

*Wikipedia, MyWeb,  
Flickr, Answers, CIM*

Social Search

## Marketplaces

Trusted  
transactions

*eBay, Craigslist*





# Powering Online Communities

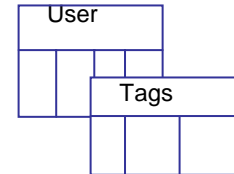
- Powering online communities

- Goal:

- Self-tuning platforms to support *stylized database services* and applications on a planet-wide scale

- Challenges:

- Performance
    - Federation
    - Reliability
    - Maintainability
    - Application-level customizability
    - Security
    - Varied data types & multimedia content
    - Extraction and exploitation of structure from web content*



# YAHOO! GROUPS

Welcome, **sihemameryahia**  
[[Sign Out](#), [My Account](#)]

[Groups Home](#) - [Help](#)

### My Groups [Manage](#)

**yg-research**  
(No new activity)



[Simplify](#)

New = within last 7 days  
[Don't see all your groups?](#)

### Get the Story

Yahoo! Groups is making it even easier for people to meet others and make a difference together. But don't take our word for it. [Read their stories here.](#)

### Send us Feedback

We want to hear from you! If you have something to say, [let us know.](#)

### Find a Yahoo! Group


[Search](#)

(Example: [meditation](#), [jokes](#), [conservation](#))


Easy as 1-2-3.  
[Start your group today.](#)

### Inside Yahoo! Groups


See how people are connecting and sharing:




**Babies**  
Feeling overwhelmed? Find solace in parents.



**Roadside Fun**  
Find a unique location and take a road trip!



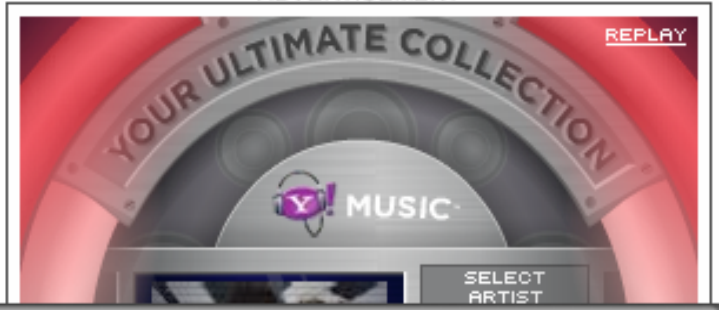
**Diaper Days**  
New moms get a chance to leave the nest.



**Garden Messenger**  
Get your fingers in the dirt.

- ### Browse Groups
- [Business & Finance](#)
  - [Computers & Internet](#)
  - [Cultures & Community](#)
  - [Entertainment & Arts](#)
  - [Family & Home](#)

ADVERTISEMENT





# Another Kind of Search

- What is a relevant group?
  - A group whose content is relevant to the query keywords
  - A group to which many of my buddies belong
  - A group where many of my buddies post messages
  - A group with some of my preferred characteristics: traffic, membership
- Search within a group
  - Messages in a group stored in one mbox file distributed across 20 machines. Each mbox is at most 2MB. Large groups have 1000 messages and large messages are 2KB.
  - Search on
    - Message: author (name, email address, Y! alias, YID), body, subject, is-spam, is-special-notice, is-topic
    - Thread: Returned if its first message is on the input topic
  - Messages returned sorted by date.



# Yahoo! My Web 2.0

My Web 2.0 Home Page



My Web (beta)

Search the Web

Search buttons. My Web search includes search for both My Stuff and community stuff

Sun. 11/13/05 2:10pm PST

Now Serving 319,070 pages and 84,650 tags

Personal knowledge

**My Web:** [308 Saved Pages](#) [add] | [171 Tags](#) | [Search History](#) [OFF - On]

Community knowledge

**My Community's Web:** [4,935 Saved Pages](#) | [3,879 Tags](#) | [53 Contacts](#) [invite]

Recent Pages | [Popular Pages](#)

[» Scoble Pulls Anti-Google Post InsideGoogle » part of the Blog News Channel](#)

Note: scoble pulls post

Shared by: [Tim](#) 😊 Tags: [scoble google](#) 19 minutes ago - [Details](#) - [Save](#)

Community vitality  
• Statistics  
• Recent pages  
• Popular pages

Public knowledge

**Everyone's Web:** [319,070 Saved Pages](#) | [84,650 Tags](#)

Recent Pages | [Popular Pages](#)

[Tips for Working Securely from Hotspots](#)

Shared by: [hermosa1001](#) Tags: [security](#), [wireless](#) 14 minutes ago - [Details](#) - [Save](#)

Public vitality  
• Statistics  
• Recent pages  
• Popular pages

(Courtesy: Raymie Stata)



# Save/Tag Pages You Like

You can save / tag pages you like into My Web from toolbar / bookmarklet / save buttons

You can pick tags from the suggested tags based on collaborative tagging technology

Type-ahead based on the tags you have used

You can specify a sharing mode

You can save a cache copy of the page content

Enter your note for personal recall and sharing purpose

Save To My Web

**YAHOO! SEARCH** Welcome [Sign out, My Account]

**My Web** [View My Saved Pages](#)

Save this page to My Web:

URL: <http://blog.lisaconference.org/>

Title: The LISA ~05 Conference Blog

Add Note: Great place to get updated info about Lisa'05

Suggested: 2005 + blog - lisa + usenix

Tags: 2005, lisa, bl  
\* Separate multiple tags with commas

Access:  Me  My Community  Everyone

Store a copy of this page.

Save Cancel

(Courtesy: Raymie Stata)



# Unaugmented Results

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

[My Web BETA](#) | [My Search History](#) OFF | [On](#)

[Search Services](#) | [Adv](#)

Search Results

Results 1 - 10 of about 129,000,000 for [Lisa](#) -

Also try: [lisa lynnette clark](#), [lisa loeb](#), [lisa raye](#), [mona lisa](#) [More...](#)

**News Results for Lisa**

[Lisa Lynn Sargeson \(Olsen\) Marsolek](#) - Independent Record - 4 minutes ago  
[UCLA's Lisa Willis Named Women's Basketball Pac-10 Player of the Week](#) - Pac  
 By [LISA MEYER TRIGG](#) Editor - Banner Graphic - Nov 23 11:37 AM  
[Yahoo! Shortcut](#) - [About](#)

Latest news results for "Lisa". Mostly about people because Lisa is a popular name

**My Web Results for Lisa (41)**

41 results from My Web!

1. [The Localization Industry Standards Association : home page](#)

Remember me. Quick Links. Welcome to LISA. Becoming a global enterprise is one of the most important challenges that your organization will ever face. There is no one right way to do it, but you should not have to reinvent the wheel. ... LISA is the leading international forum for organizations doing business globally ...

Category: [Software](#) > [Translation](#)  
 RSS: [View as XML](#) - [Add to My Yahoo!](#)  
[www.lisa.org](#) - [More from this site](#) - [Save](#) - [Block](#)

Web search results are very diversified, covering pages about organizations, projects, people, events, etc.

2. [Laser Interferometer Space Antenna](#)

The Laser Interferometer Space Antenna is a mission that will detect and study gravitational waves from sources involving massive black holes and galactic binaries. ... Download new LTS (PDF file) LISA is a joint mission between the European Space Agency and NASA (Structure and Evolution ...

[lisa.jpl.nasa.gov](#) - 19k - [Cached](#) - [More from this site](#) - [Save](#) - [Block](#)

3. [ESA Science & Technology: LISA](#)

... THE MISSION: LISA is an ESA-NASA mission involving three spacecraft flying approximately 5 million kilometres apart in ... Letter of Intent to Participate in LISA data processing study ...

[sci.esa.int/science-e/www/area/index.cfm?fareaid=27](#) - 31k - [Cached](#) - [More from this site](#) - [Save](#) - [Block](#)

4. [a modern girl](#)





# Community-/Personally-Enhanced Search

YAHOO! SEARCH

Lisa

My Web (beta)

Search the Web

My Web 2.0 BETA

[Home](#) [Invite](#) [Contacts](#) [Add Page](#) [Tools](#)

Pages Tags Contacts

[History](#) [Blog](#) [FAQ](#) [Discuss](#) [Preferences](#)

Search Results

Results 1 - 10 of 41 for Lisa shared by your contacts and their contacts.

Related tags: [lisa](#), [2005](#), [usenix](#), [blog](#) [More...](#)

## My Community's Results (41)

1. [SAGewire | LISA 2005 Blog](#)

system administrator n.a system administrator is one who, as a primary job function, manages computer and network systems on behalf of another, such as an employer or client. LISA 2005 Blog. David N.  
 Saved by: [Zhichen Xu](#) and 1 other 10:05 AM PST - [Details](#)  
 Tags: [blog](#), [lisa](#), [usenix](#)  
<http://sagewire.sage.org/article.pl?sid=05/08/16/0237249> - [Save](#)
2. [The LISA '05 Conference Blog](#)

Behind the Scenes at the Largest System Administration Conference in the World. November (here's another burning-hot news nugget that will be posted on the official conference site with days)  
 Saved by: [Jianchang \(JC\) Mao](#) 10:09 AM PST - [Details](#)  
 Note: **Official LISA conference blog. A lot of good stuff behind the scene**  
 Tags: [2005](#), [blog](#), [lisa](#), [usenix](#)  
<http://blog.lisaconference.org/> - [Save](#)
3. [19th Large Installation System Administration Conference ...](#)

Administrators of all specialties and levels of expertise meet at LISA to exchange ideas, sharpen old skills, learn new techniques, debate current issues, and meet colleagues and friends.  
 Saved by: [Jianchang \(JC\) Mao](#) on November 26, 2005 - [Details](#)  
 Note: **Qi is giving a keynote speech. Excellent system conferece. Dates: Dec 4-9, 2005**  
 Tags: [conference](#), [lisa](#), [systems](#)  
<http://www.usenix.org/events/lisa05> - [Save](#)

Excellent set of search results from my community because a couple of people in my community are interested in Usenix Lisa-related topics



# Community Information Management

- Many real-life communities have a web presence
  - Database researchers, movie fans, stock traders
- Each community = many disparate data sources + people
- Members want to *query and track at a semantic level*:
  - Any interesting connection between researchers X and Y?
  - List all courses that cite this paper
  - Find all citations of this paper in the past one week on the Web
  - What is new in the past 24 hours in the database community?
  - Which faculty candidates are interviewing this year, where?





# DBLife Portal – A CIM Prototype

- Cimple is a joint project between University of Wisconsin at Madison and Yahoo! Research
- Software platform that manages data-rich online communities
- Prototype system (DBLife) up and running since 2005
- 1164 sources, crawled daily, 11000+ pages / day
- 160+ MB, 121400+ people mentions, 5600+ persons

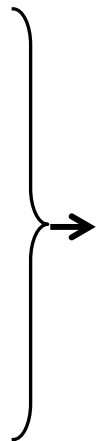
Researcher  
Homepages

Conference  
Pages

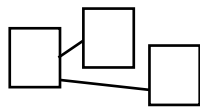
Group Pages

DBworld  
mailing list

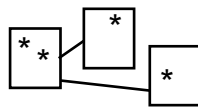
DBLP



Web pages

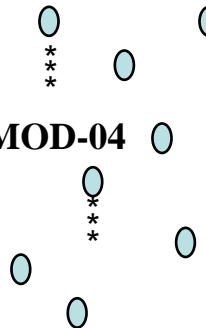


Text documents



SIGMOD-04

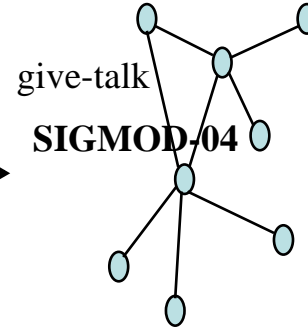
Jim Gray



give-talk

SIGMOD-04

Jim Gray



Search, query

Mining

Alerts

News summary

...





# DBLife Portal

- Integrated information about a real-world community
- Collaboratively built and maintained by the community



- Potentially important connection to bottom-up, extracted structure – but how?

DBLife: Superhomepage of Gerhard Weikum - Mozilla Firefox

File Edit View Go Bookmarks Yahoo! Tools Help

**DBLife** weikum Search Login Help The Cimple Project Wiki

## Gerhard Weikum

Check all *non-related* images and click the submit button. [Show All Images](#)

0 votes 0 votes 0 votes 0 votes 0 votes 0 votes 0 votes

Submit [Check All](#) [Clear All](#)

Mentions 1 - 10 out of 137

Tuesday Oct 31, 2006 [PC member of SIGMOD 2007](#)  
...Jeff UllmanStanford University Patrick ValduriezINRIA Wei WangUniversity of North Carolina **Gerhard Weikum**Max-Planck Institute of Computer Sc. Till WestmannBEA Kyu Young WhangKAIST Jun YangDuke University...  
<http://sigmod07.rii.tsinghua.edu.cn/program.shtml> [Cached](#) [Annotated](#)

Monday Oct 30, 2006 ["PUBLICATIONS LIST:" page, Patrick O'Neil's homepage](#)  
...Invited paper to the IEEE Bulletin on Data Engineering, March 1994, 45-50, Vol. 17, No. 1. With **Gerhard Weikum**, "A Log-Structured History Data Access Method (LHAM)". Presented at the Fifth International...  
<http://www.cs.umb.edu/~poneil/publist.html> [Cached](#) [Annotated](#)

Thursday Oct 26, 2006 [Posts message "CIDR 2007 Accepted Papers and Demos" on Dbworld](#)  
...25-Oct-2006 conf. ann. **Gerhard Weikum** CIDR 2007 Accepted Papers and Demos 7-Jan-2007 web page...  
<http://www.cs.wisc.edu/dbworld/browse.html> [Cached](#) [Annotated](#)

Saturday Oct 21, 2006 [PC member of SMDB 2007](#)  
Page monitored for first time  
...Canada Ken Salem (University of Waterloo, Canada) Dennis Shasha (New York University, USA) **Gerhard Weikum** (Max-Planck-Institut fur Informatik, Germany)...  
<http://www.cs.wisc.edu/dbworld/messages/2006-10/1161389516.html> [Cached](#) [Annotated](#)

Wednesday Oct 18, 2006 [Vagelis Hristidis's homepage](#)  
...Demos/Posters/Short Papers 18. Nishant Kapoor, Gautam Das, Vagelis Hristidis, S. Sudarshan, **Gerhard Weikum**. STAR: A System for Tuple and Attribute Ranking of Query Answers. Demo paper, ICDE 2007 19. ...  
<http://www.cs.fiu.edu/~vagelis/index.html> [Cached](#) [Annotated](#)

137 total mentions occurring in 90 pages.  
0 new mentions found in the last 24 hours.

### Related People

- [Vagelis Hristidis](#)
- [Surajit Chaudhuri](#)
- [Gustavo Alonso](#)
- [Nick Koudas](#)

[more](#)

### Related Topics

- [xml](#)
- [optimization](#)
- [performance](#)
- [algorithms](#)

[more](#)

### Services

- [ICDE 2007 \(PC member\)](#)
- [SIGMOD 2007 \(PC member\)](#)
- [SMDB 2007 \(PC member\)](#)
- [WWW 2007 \(PC member\)](#)

[more](#)

### Publications

- [Efficient and Decentralized PageRank Approximation in a Peer-to-Peer Web Search Network](#)
- [IO-Top-k: Index-access Optimized Top-k Query Processing](#)
- [Foundations of Automated Database Tuning](#)
- [Counting at Large: Efficient Cardinality Estimation in Internet-Scale Data Networks](#)

[more \(filtered\)](#)

### Co-Authors

- [Hans-Jörg Schek](#) [05 \(2\)](#), [03 \(2\)](#), ...
- [Ralf Schenkel](#) [06 \(6\)](#), ...
- [Sebastian Michel](#) [06 \(4\)](#), [05 \(7\)](#), ...
- [Matthias Bender](#) [06 \(3\)](#), [05 \(5\)](#), ...

[more \(filtered\)](#)

### Talks

- [DBRank 2007](#)

Done



# Semantic Data Enrichment

- Find mentions of entities on a page: researchers, conferences, papers, talks, etc.
  - A mention is a span of text referring to an entity (e.g., first name, middle initial, and last name)
- Infer meta-data for each mention

**Timestamps** – When the mention first appeared

**Explanations** – Why the mention is there



- Many sophisticated techniques are known
  - Must *exploit domain knowledge* to do a better job
- We currently find about 114,400 mentions per day



# Challenges for the DBLife Project

- Structure related:
  - Extraction
    - Domain-level vs. site-level
    - Compositional approach to extraction planning
  - *Maintenance of extracted information*
    - Critical, but very little has been done
  - Exploitation
    - Search/query over extracted structures
    - Detect interesting events and changes
- Collaborative-editing related:
  - Incentives, reputation – ability to correct/add imported data
  - Handling malicious/spam users
  - Ownership model
    - My home page vs. a citation that appears on it
  - Reconciliation
    - Extracted vs. manual input
    - Conflicting input from different users
  - Fundamental challenge: Determine if two mentions refer to same entity  
“John Smith” = “J. Smith”? “Dave Jones” = “David Jones”?

- Entity identity
- Potentially inconsistent information
- Collaborative editing





# Community Systems Opportunities and Challenges

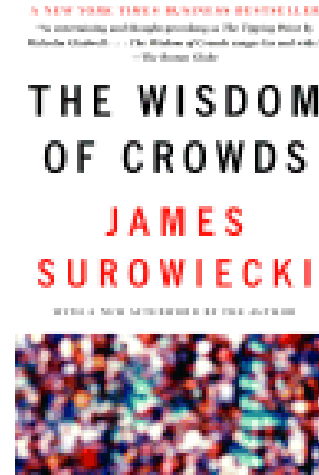
- Cull information from independently-developed sites and *map into a set of queryable structures*
- More generally: superimposition of a unifying view on content not created from shared ontologies (or any explicit ones at all)
  - See Brachman, “*Viewing Databases Through a Knowledge Representation Lens*,” *Knowledge Building and Knowledge Sharing*, 1994
  - Key sub-problem: entity matching
  - What to do in face of irreconcilable differences?
- Representation, classification of human resources to support routing of requests

**All at Web scale!**



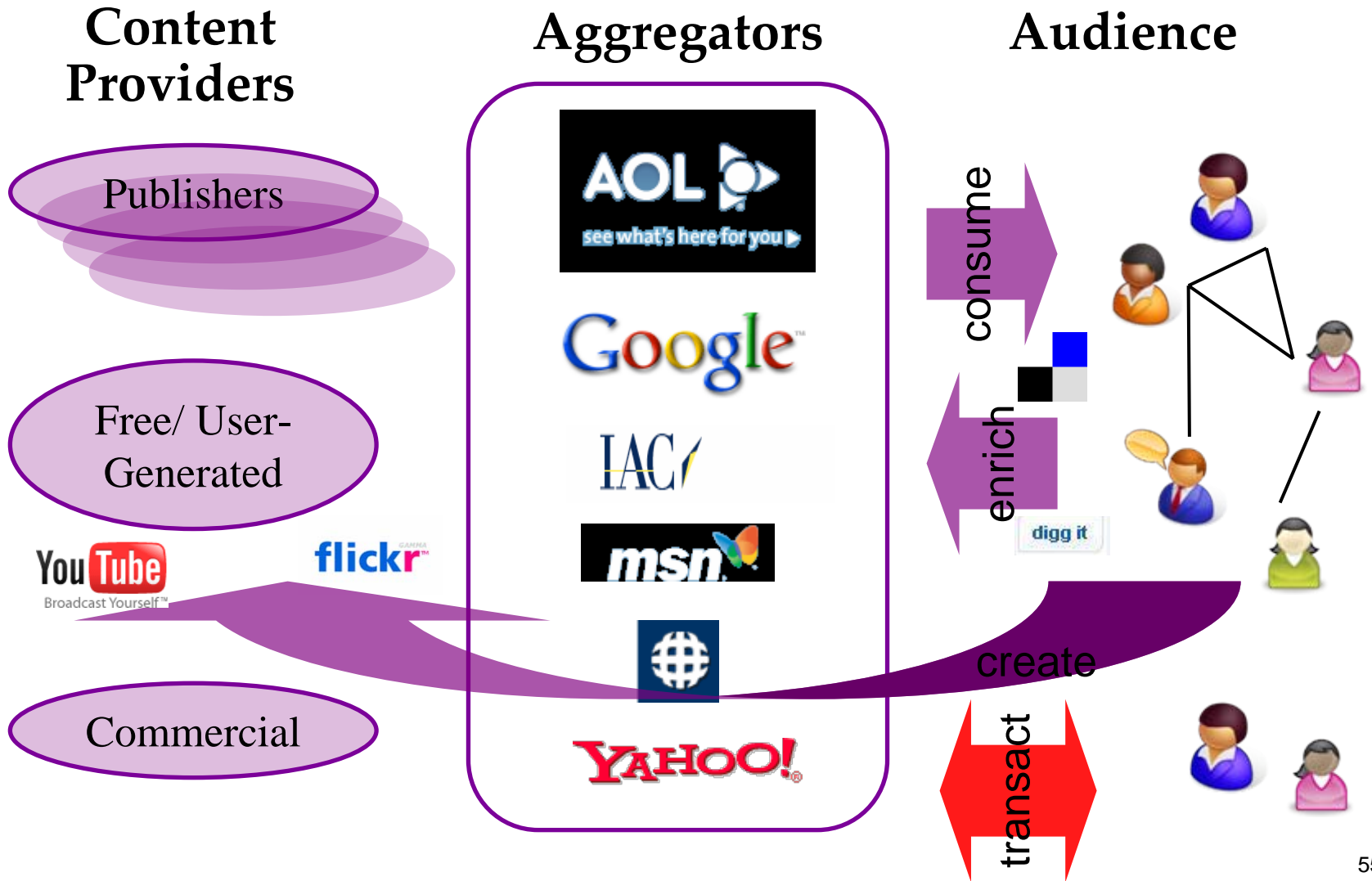
# Algorithmic Advertising/ Computational Micro-economics

- Advertising keeps us in business (paid placement)
  - Matching to context
- Ad placement
- Markets
  - “Wisdom of crowds”
- Incentive systems





# Overall Context





# Paying the Bills

- Aggregators draw content consumers
  - Search is a significant “hook”
- Consumer reveals clues about his current situation
  - The keyword(s) he types (e.g., ***miele***)
  - Keyword(s) in his email (gmail)
  - Context information (Yahoo! ...)





# Paid Placement

- Aggregator gives consumer opportunity to click through to an advertiser
  - Compensated by advertiser for click through
- Whose advertisement is displayed?
  - In the simplest form, *auction bids* for each keyword
  - Contracts:
    - “At least 20000 presentations of my advertisement to searchers typing the keyword **nfl**, on Super Bowl Sunday”.
    - “At least 100,000 impressions to searchers typing **wilson** in the Yahoo! Tennis category in August”.



# Paid Placement

- Leads to complex logistical problems: selling contracts, scheduling ads – supply chain optimization
- Interesting issues at the interface of search and paid placement:
  - If you search for *miele*, did you really want the home page of the Miele Corporation at the top?
  - If not, which appliance vendor?

Sponsored Search/  
Search Advertising

Web Results 1 - 10 of about 11,300,000 for tundra [definition]. (0.06 seconds)

The **tundra** biome  
Tundra is the coldest of all the biomes. Tundra comes from the Finnish word ... Arctic tundra is located in the northern hemisphere, encircling the north ...  
www.ucmp.berkeley.edu/exhibits/biomes/tundra.php - 14k - Cached - Similar pages

**Tundra Biome**  
All about the Alpine biome, description, global position, climate, and information on specific plants and animals of this biome.  
www.blueplanetbiomes.org/tundra.htm - 16k - Cached - Similar pages

**Tundra**  
The tundra is the simplest biome in terms of species composition and food chains ...  
Distribution: The tundra biome is restricted to the high latitudes of ...  
www.radford.edu/~swoodwar/CLASSES/GEORG235/biomes/tundra/tundra.html - 11k - Cached - Similar pages

**Tundra Semiconductor**  
Designer, developer and marketer of IC's for system interconnections.  
www.tundra.com/ - 45k - Cached - Similar pages

**Tundra 5**  
There are two types of tundra in the world, Arctic and Alpine. The arctic tundra is at the top of the world around the North Pole. ...  
ths.sns.lane.edu/biomes/tundra5/tundra5.html - 10k - Cached - Similar pages

Ads go in slots like these

Higher slots get more clicks

Sponsored Links

- [Nissan Titan](#)  
Standard V8, Fully Boxed Frame, 317HP, 385 LBS Ft. Torque.  
www.NissanUSA.com/Titan
- [The 2007 Toyota Tundra](#)  
The all-new Tundra has arrived. The truck that's changing it all.  
www.Toyota.com/Tundra
- [Top New Truck Prices](#)  
Find out our Lowest Possible Price on a new Truck Purchase or Lease!  
www.CarPriceSecrets.com
- [AutoUSA](#)  
Buy A New Car at a Low Price Available From Dealers in Your Area  
www.autousa.com/Toyota  
San Francisco-Oakland-San Jose, CA

My Yahoo! Welcome, rbrachman1 [ Sign Out, My Account ]

**MY YAHOO!** MAKE WEB SITE MAGIC. It's easy with Yahoo! Web Hosting. Now Get 25% off for a limited time. **YAHOO! SMALL BUSINESS** Get Started

Category: Web | Images | Video | Local | News | Shopping

Search:  Search

AP: Top Stories

- U.S. deaths near grim Memorial Day mark - 3 hours ago
- U.S. frees 42 al-Qaida captives in raid - 30 minutes ago
- Israel pledges more strikes on Hamas - 55 minutes ago

Weather: Westfield, NJ 64...86 F Sunnvale, CA 52...77 F

Search by Zip Code or City

Search **I want my Credit Score!**

Crave: The Gadget Blog Tech to hunger for: hot gadgets, cool games

YAHOO! HEALTH Welcome, rbrachman1 [ Sign Out, My Account ]

BETTER IN '07: SUMMER SHAPE UP

Live well. Eat well. Exercise well. It's all about having fun while improving your fitness.

Special K JAM: Get Healthy Now MAR: Spring Nutrition Guide MAY: Summer Shape Up

**Brand/"Graphical" Context-Based Advertising**

ESPN.com Philadelphia Eagles Clubhouse

Member Name:  Password:  Sign In

**Lamisil AT Gel®. Half the treatments. All the cure. Learn more now**

**Philadelphia Eagles**

Search  Download Toolbar | Shop

ESPN NFL MLB NBA NASCAR Autos NHL College Golf Soccer Tennis Page 2 Nation Insider Fantasy 360 More +

Eagles Home | Stats | Schedule | Roster | Photos | Rankings | Depth Chart | Content | Trans | Message Board | Tickets | Shop

Philadelphia Eagles - 2007 NFL DraftTracker

RD	PK(OVR)	NAME	POS	SCHOOL
2	4(36)	Kevin Kolb	QB	HOUSTON
2	25(57)	Vince Abigiani	DE	NOTRE DAME

Select Another Clubhouse

Five Game Schedule/Results

DATE	OPPONENT	RESULT
W1: Sun 9/9	@Green Bay	1:00 PM ET

SELF

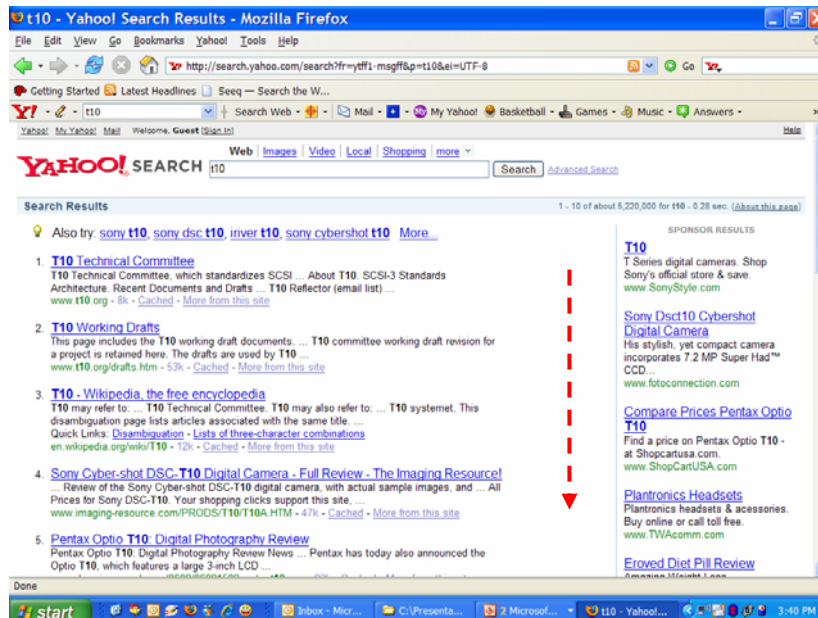
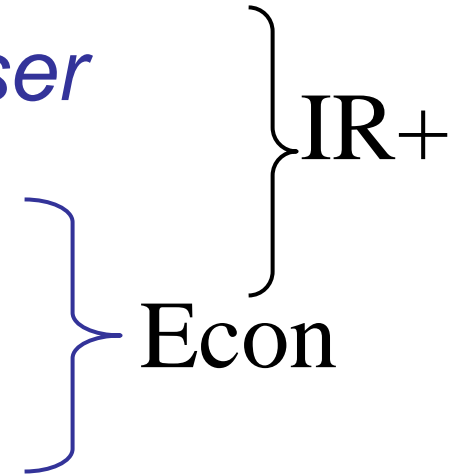
Advertisement: Click to hear why they stopped skipping. Stories based on consumer insights and nutritional support. Research shows women who eat breakfast weigh less. Exercise / healthy lifestyle are part of any diet.

Share Your Success! Share your experiences, trade weight loss tips, and celebrate your success on the



# Three Subproblems

1. Match ads to query/context/user
2. Order the ads
3. Pricing on a click-through





# The Context Match Challenge

Find the “best match” between a given user in a given context and a suitable advertisement

## Examples

- Context = search results → sponsored search
- Context = publisher page – content on the page → banner ads
- Other contexts: mobile, video, newspapers, etc







# Search Page Context

The screenshot shows a Live Search results page for the query "scooter libby". The search bar at the top is circled in red. The page displays several search results, with the first one being a Wikipedia entry for Lewis Libby. On the right side, there is a "Related searches" section with "Timeline Scooter Libby Investigation" and a "SPONSORED SITES" section. One of the sponsored sites, "Compare Top Scooters Online. Research scooter libby", is circled in red. The page also includes navigation tabs for "Web", "Images", "News", "Maps", "Classifieds", and "More".

Live Search

scooter libby

Sign in

Web Images News Maps Classifieds More

scooter libby Page 1 of 249,234 results • Options • Advanced

Back to msn

**Lewis Libby** - Wikipedia, the free encyclopedia  
I. Lewis **Scooter Libby** Jr. (born August 22, 1950) is the former Chief of Staff and assistant for National Security Affairs to U.S. Vice President Dick Cheney and a former presidential advisor. His ...  
[en.wikipedia.org/wiki/Scooter\\_Libby](http://en.wikipedia.org/wiki/Scooter_Libby) • [Cached page](#)

**Libby Legal Defense Trust**  
March 6, 2007. As **Scooter's** lawyer, Ted Wells, said today, we are all disappointed in the verdict. His attorneys will seek a new trial and if that's denied, they will appeal.  
[www.scooterlibby.com](http://www.scooterlibby.com) • 5/26/2007 • [Cached page](#)

**Libby Legal Defense Trust**  
The **Libby** Legal Defense Trust was formed to help defray the legal defense costs for Lewis "**Scooter**" **Libby** and his family against the recent legal charges brought against him.  
[www.scooterlibby.com/about](http://www.scooterlibby.com/about) • 5/25/2007 • [Cached page](#)  
[+Show more results from www.scooterlibby.com](#)

**Jury convicts Libby on four charges** - Politics - MSNBC.com  
Vice President Dick Cheney's former chief of staff, I. Lewis "**Scooter**" **Libby**, was convicted Tuesday ... Jurors convict **Libby** on four of five charges Cheney's ex-aidé faces jail time in CIA leak case ...  
[www.msnbc.msn.com/id/17479718](http://www.msnbc.msn.com/id/17479718) • [Cached page](#)

Top Cheney aide **Libby** indicted, quits - Leak Investigation

**Related searches:**  
**Timeline Scooter Libby Investigation**

SPONSORED SITES

**scooter libby**  
Compare Top Scooters Online. Research scooter libby  
[www.coupscoots.com](http://www.coupscoots.com)

**Hey Republicans!**  
Do You Think Your Party Is Corrupt? Or Just A Few Bad Apples? Vote Now  
[www.popularq.com](http://www.popularq.com)

**Scooter Libby**  
Search Smarter for **Scooter Libby** and save  
[www.smarter.com](http://www.smarter.com)

**Scooter Libby**  
Learn about **Scooter Libby**.  
[www.shopica.com](http://www.shopica.com)

**Scooter Libby at Amazon**  
Low prices on **scooter libby**. Qualified orders over \$25 ship free.





# Content Page Context

The screenshot shows the top portion of a Slate article. The main header features the 'Slate' logo and a search bar. Below the header is a navigation menu with categories like 'News & Politics', 'Arts & Life', etc. The article title is 'Who Is Scooter Libby?' with a sub-headline 'THE SECRETIVE CHENEY AIDE AT THE HEART OF THE CIA LEAK CASE.' and author 'By John Dickerson'. A 'PRINT' button is visible on the left sidebar. A 'BizBox on Slate' advertisement is on the right. A small photo of Scooter Libby is partially visible at the bottom of the article content.



This section is titled 'ALSO IN SLATE' and contains four article thumbnails: 'Help! My Neighbor Is Ned Flanders.', 'Explainer: Is McCain's Brain Too Old for the Presidency?', 'The Snob Appeal of Tap Water', and 'Enjoying the Phil Spector Trial Way Too Much'. Below this is an 'Ads by Google' section with two ads circled in red: 'Free Hydrofoil Scooter' and 'Scooter'. The bottom of the page contains footer text: 'build your own slate', 'feedback | about us | help | advertise | newsletters | mobile', '© 2007 Washington Post, Newsweek Interactive Co. LLC', and 'User Agreement and Privacy Policy | All rights reserved'.



# Other Match Challenges

Web | Images | Video | Local | Shopping | more




**YAHOO! SEARCH**  Search [Advanced](#)

Search Results

- [RME](#)  
Makers of professional digital sound cards and converters, including the Hammerfall and DIG196 series.  
[www.rme-audio.com](http://www.rme-audio.com) - 3k - [Cached](#) - [More from this site](#)
- [RME Digiface](#), [RME Multiface](#), [RME Fireface 800](#), [RME Hammerfall DSP](#), [RME ...](#)  
...ations, [RME Audio](#), [Multiface](#), [RME ...](#)  
...versatile PC...

Shared by Yahoos

Answers 1 - 3 of 7 for rme

- [pls answe rme?](#)  
 [Ibruprofen](#) - you can get this for babies and adults. Also bonjela and frozen things...  
Best Answer - Chosen by users
- [I am seeing that belly is developing fo rme. How to arrest it ...](#)  
 Even if you dont want to change your diet, at the very least do crunches, not sit-ups. Sit-ups put too...  
Best Answer - Chosen by users
- [♥hmm....which college do u think is good fo rme?](#)  
 you don't really sound like college material.  
Best Answer - Chosen by users

[More Answers...](#) | [Ask a question...](#)



# Context Matching

## Key Challenges

- What is “best” match?
  - An ad has different utilities for publishers, advertisers, users
  - “Perceived appropriateness factor” is different
  - Might have different types of ads that are not easily compared
- Need to deal with billions of ads, contexts, users ... often, in real time
- Need to deal with constraints: budgets, business controls, incentives, etc.
- Business objectives not easy to express

• There may be very few clues

**To be successful in matching, need to**

- Understand the context
- Understand the user
- Observe the past to predict the future
- Search the available ads in real time






# Ordering the Ads


- Most generally, composite IR+Econ score ...  
but today, basically Econ
- Original GoTo/Overture scheme:
  - Order by bid


Also try: [boeing 747](#), [747 airplane](#), [747 se](#)

- My Web: [23 Results for 747](#)
- [Om Malik on Broadband](#) - Saved to
- [BBC NEWS | Business | Google s](#)
- [SeatGuru.com - Your Enlightened](#)
- [See all 23 results](#) ...

**Bid Ordering  
(Overture/former  
Yahoo!)**

- [Boeing 747-400 Family](#) 

Corporate product page from Boeing.  
Category: [Planes](#) > [Boeing 747 Family](#)  
[www.boeing.com/commercial/747family](http://www.boeing.com/commercial/747family) - 11k - [Cached](#) - [More from this site](#) - [Save](#)
- [Wikipedia: Boeing 747](#) 

The Boeing 747, commonly called the Jumbo Jet, is one of the most recognizable modern airliners and is the largest airliner currently in airline service. First flown commercially in 1970, it held the size record for more than 35 years, although ... The four-engine 747, produced by Boeing Commercial Aircraft, uses a two ...  
Quick Links: [History](#) - [Development](#) - [In service](#)  
[en.wikipedia.org/wiki/Boeing\\_747](http://en.wikipedia.org/wiki/Boeing_747) - 102k - [Cached](#) - [More from this site](#) - [Save](#)
- [Wikipedia: 747](#) 

==Events== ... This article is about the year 747. For the aircraft, see Boeing 747. For the song by Kent, see 747 (song). For other uses of 747, see 747 (number ...  
Quick Links: [Events](#) - [Births](#) - [Deaths](#)  
[en.wikipedia.org/wiki/747](http://en.wikipedia.org/wiki/747) - 16k - [Cached](#) - [More from this site](#) - [Save](#)

SPONSOR RESULTS

[747 on eBay](#)

Shop eBay for anything and everything - from specialized gifts to ...  
[www.ebay.com](http://www.ebay.com)

[Shop Boeing 747-400 Miniature Models](#)

Shop Boeing 747-400 miniature models at FlightMiniatures.com. See ...  
[www.flightminiatures.com](http://www.flightminiatures.com)

[Canon Fax L777 Toner-free Shipping](#)

Save on compatible Canon Fax L777 toner. Purchase our ...  
[www.ink-refills-ink.com](http://www.ink-refills-ink.com)

[747 Airplanes - Free Shipping](#)

Thousands of hand crafted fully assembled model plane replicas in ...



# Revenue Ordering (simplified)

- Each ad  $j$  has an expected click-through denoted  $CTR_j$
- Ad  $j$ 's bid is denoted  $b_j$
- Then, expected revenue from this ad is  $R_j = b_j \times CTR_j$
- Order ads by  $R_j$

← Estimated (how?)





Web Images Groups News Froogle Maps more »

747

Search

Advanced Search Preferences

Revenue ordering?

Results 1 - 10 of about 83,900,000 for 747. (0.10 seconds)

Web

[Boeing: Commercial Airplanes - 747 Home](#)

Boeing Announces First Engineering Releases for New 747-8 Family - Boeing Delivers Malaysia Airlines System's First 747-400 Freighter - Cargolux Expands ...  
[www.boeing.com/commercial/747family/](http://www.boeing.com/commercial/747family/) - 12k - May 3, 2006 - [Cached](#) - [Similar pages](#)

[Boeing: Commercial Airplanes - 747 - About the 747 Family](#)

The Boeing 747 Family -- The Right Choice for the Large Airplane Market. About the 747 Family. Instantly recognized by passengers around the world, ...  
[www.boeing.com/commercial/747family/background.html](http://www.boeing.com/commercial/747family/background.html) - 26k - [Cached](#) - [Similar pages](#)  
[ [More results from www.boeing.com](#) ]

[Boeing 747 - Wikipedia, the free encyclopedia](#)

The Boeing 747, commonly called the Jumbo Jet, is one of the most ... The 747SP was the longest-range airliner available until the 747-400 entered service ...  
[en.wikipedia.org/wiki/Boeing\\_747](http://en.wikipedia.org/wiki/Boeing_747) - 101k - May 3, 2006 - [Cached](#) - [Similar pages](#)

[Airliners.net: Boeing 747-400](#)

The 747-400 is the latest, longest ranging and best selling model of the 747 family. Boeing launched the 747-400 in October 1985 and the first development ...  
[www.airliners.net/info/stats.main?id=100](http://www.airliners.net/info/stats.main?id=100) - 27k - May 3, 2006 - [Cached](#) - [Similar pages](#)

[Boeing 747SP Website](#)

Boeing 747SP Production List, Photo Gallery, Incident Reports, Current Status, News and much more.  
[www.747sp.com/](http://www.747sp.com/) - 21k - May 3, 2006 - [Cached](#) - [Similar pages](#)

Sponsored Links

[747 On Sale](#)

Chairs, Desks & Cabinets  
Everyday Low Prices.  
[OfficeFurniture.BettyMills.com](http://OfficeFurniture.BettyMills.com)

[Boeing™ 747-400 Model](#)

Accurately detailed Boeing™ 747-400 desktop display models  
[www.FlightMiniatures.com](http://www.FlightMiniatures.com)

[Boeing's Photo Collection](#)

The world's leading source of still and motion aerospace imagery!  
[www.boeingphotostore.com](http://www.boeingphotostore.com)

[747](#)

Everything to do with 747 items.  
[eBay.com](http://eBay.com)





Updated Yahoo! ordering

Search Results

Also try: 747 airplane, area code 747, 747 jumbo jet, 747 pictures

1. Boeing 747 Family

Corporate product page from Boeing for the 747 series of commercial aircraft. Category: Planes > Boeing 747 Family www.boeing.com/commercial/747family - 11k - Cached - More from this site - Save

2. Boeing 747 - Wikipedia, the free encyclopedia

The Boeing 747, commonly called a Jumbo Jet, is one of the most recognizable of ... The prototype 747, City of Everett, at the Museum of Flight in Seattle, ... Quick Links: History - Development - In service | Saved by 5 people en.wikipedia.org/wiki/Boeing\_747 - 125k - Cached - More from this site - Save

3. Airliners.net: Boeing 747-100 & 200

Largest aviation ... 747-200B - Max speed 981km/h (530kt) (with RR engines) ... 747-200F - Range with 90,270kg (200,000lb) payload 9075km (4900nm) ... www.airliners.net/info/stats.main?id=97 - 26k - Cached - More from this site - Save

4. Boeing: Commercial Airplanes - 747 - About the 747 Family

... by passengers around the world, the Boeing 747 is in a class by itself. ... Along with earlier versions, 747 Freighters -- about 225 in all -- carry half of ... www.boeing.com/commercial/747family/background.html - 26k - Cached - More from this site - Save

5. Airliners.net: Boeing 747-400

1 - 10 of about 49,900,000 for 747 - 0.11 sec. (About this page)

SPONSOR RESULTS

This Part and Millions Others Available

747 - This part and millions others available. Supplying... www.cpluselectronics.com

747 Jumbo Jet

Enjoy Savings & Selection On 747 jumbo jet. www.Shopzilla.com

Reverse Lookup Any 747 #

Find Name & Address for any Number. Also works w/ Cell Phone Numbers. ReversePhoneDetective.com

747 at MSN Shopping

Save on brand name office products at MSN Shopping - shop by price. shopping.msn.com

747 Pictures

Save on Clothing. Compare products, prices & stores



# Economic Ordering

- Bid and revenue ordering: two forms of ordering by “Econ” score
- Does revenue ordering maximize revenue?
- Not necessarily – advertisers react to ordering scheme, by changing their bid behavior!
  - Lahaie+Pennock ACM EC 2007
    - Family of schemes bridging Bid and Revenue ordering



# A Sea of Challenges

- How do you design marketplaces for search advertising?
  - What tools do you give advertisers?
  - What do they get to bid on?
    - *Laptop* vs *Ultralight laptop*
  - How do you price clicks?
- Reputation and incentive mechanisms
- Matching marketplaces
  - Jobs, dates, ...
  - Online matching everywhere
    - Hardest part is estimating the payoffs, not the matching algorithm



# A New Convergence

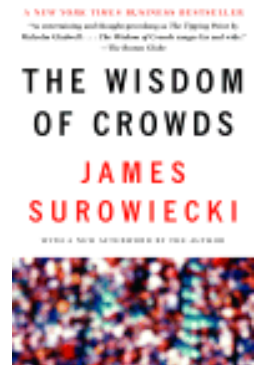
- Monetization and economic value are an intrinsic part of system design
  - Not an afterthought
  - Mistakes are costly!
- Computing meets humanities like never before
  - sociology, economics, anthropology ...





# Prediction Markets

- Futures market designed to elicit a forecast about some future event
- Leverages “wisdom of crowds”, extracting and combining information from distributed sources
- Have been used successfully to
  - Predict election outcomes [IEM, 1988-]
  - Predict corporate metrics (sales, product release times, ...) [HP, 2001-] [MSFT, Eli Lilly, Intel, Siemens] [GOOG, 2005-]
  - Predict movie box office returns [HSX], news [NewsFutures], scientific conjectures [FX], sporting events, judicial nominations, economic numbers, ... [inTrade], real estate [HedgeStreet], Mideast economic variables (proposed) [PAM], many others





# Wise Crowds vs. Irrational Ones

Not all crowds are wise – certain key features need to be present to achieve “wisdom” [Surowiecki, 2005]:

- Diversity of opinion
  - Each person should have private information even if it's just an eccentric interpretation of the known facts
- Independence
  - People's opinions aren't determined by the opinions of those around them
- Decentralization
  - People are able to specialize and draw on local knowledge
- Aggregation
  - Some mechanism exists for turning private judgments into a collective decision



# Burgeoning Research Area

- Understanding auctions, incentive mechanisms, pricing, optimization – a huge deal
  - Already multi-billion dollar business, growing fast
  - Interface of microeconomics and CS
- Advertiser's perspective – spending budget optimally
  - Many open problems, a few papers, some of them quite realistic
  - Marketplace design





# Media Experience

Creating new experiences around media

- Beyond CHI: engagement models
- New/integrated media
  - Mashup or link up?





# Beyond HCI

- 30 years ago, a collection of sociologists, cognitive psychologists and computer scientists gave us
  - The personal computer
  - The field of HCI/CHI
- Now we need to consider a science of online audience engagement
  - Not just about people interacting with computers or the web
  - But about people interacting with other people, with computers/the web as the medium
- An intrinsically data-driven science





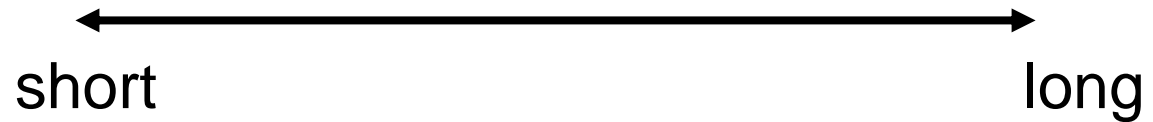
# Some Questions

- Why do people choose to lurk or participate?
- Why do people create new online personas?
- Why are YouTube, Myspace and Flickr successful? (and many others, not)
- What new genres are emerging - and what can we provoke?
  - For content creation
  - Enrichment
  - Participation

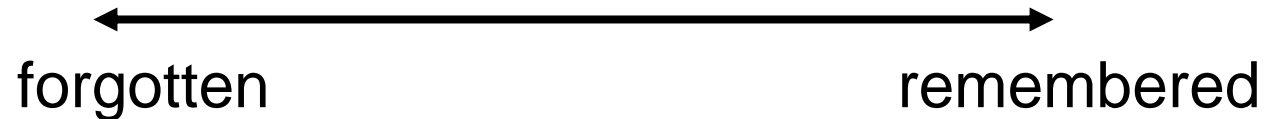


# Some Dimensions of Experience

**Duration**



**Memorability**



**Social context**





Future Galaxy Servers of Players

### Hospital

TEST IMAGE, DO NOT USE

**NiteHawk**

Level:	1
Strength:	14 4
Dexterity:	15 5
Vitality:	15 5
Intelligence:	8 -2
Perception:	8 -2
Charisma:	8 -2
Attack:	1
Armor:	0
Money:	100
Hitpoints:	40/40
Status:	Ready

Disconnect Connected Freeze

<New York Streets 2> Room 2. The street continues north and south.  
<New York Streets 3> Room 3 East. The room continues south and east.  
You cannot go that way!  
<Hospital> The Hospital. You can head out to the west.  
<New York Streets 3> Room 3 East. The room continues south and east.  
<New York Streets 2> Room 2. The street continues north and south.  
<New York Streets> You are in the New York streets. There is a smelly man with pissed pants in a box to the left of you. The street continues to the north and south.  
<New York Streets 2> Room 2. The street continues north and south.  
<New York Streets 3> Room 3 East. The room continues south and east.  
<Hospital> The Hospital. You can head out to the west.  
<New York Streets 3> Room 3 East. The room continues south and east.  
<New York Streets 2> Room 2. The street continues north and south.  
<New York Streets> You are in the New York streets. There is a smelly man with pissed pants in a box to the left of you. The street continues to the north and south.  
<New York Streets 2> Room 2. The street continues north and south.  
<New York Streets 3> Room 3 East. The room continues south and east.  
<Hospital> The Hospital. You can head out to the west.

Second life  
Future Galaxy  
WOW







mestic Try Meetic for free I am searching: A man Age: 30 years FREE SUBSCRIPTION

Property Jobs Read before you move In Sweden, in English

24th April 2007 NEWS | BLOG | JOBS | DATING | TRAVEL | DISCUSS | NOTICEBOARD | WHO'S WHO | WHAT'S WHAT | WEATHER |

National Business Politics Sport Swedish Life Society Science & Technology Analysis & Opinion

- Home SIGN IN REGISTER Why Register? Contact Us Advertise FAQ XML Newsfeeds Privacy Policy Unsubscribe

# Sweden to set up embassy in Second Life

Published: 26th January 2007 18:02 CET

Sweden is to become the first country to establish diplomatic representation in the virtual reality world of Second Life, officials said on Friday.

"We are planning to establish a Swedish embassy in Second Life primarily as an information portal for Sweden," Swedish Institute (SI) director Olle Wästberg told AFP.

The embassy would not provide passports or visas but would instruct visitors how to obtain such documents in the real world and act as a link to web-based information about the Scandinavian country.

"Second Life allows us to inform people about Sweden and broaden the opportunity for contact with Sweden easily and cheaply," Wästberg said.

The Swedish Institute is an agency of the

### Related Articles

- New Miss Sweden risks disqualification 24th April 2007
Four injured by skating knifeman 24th April 2007
Love causes man to steal from 90-year-old mother 23rd April 2007

### Article Options

### Events in Second Life

Experts in Virtual World Events and Real World Tie-Ins - TheSLAgency www.theslagency.com

### Free Second Life Scripts

download free scripts and tutorials for second life www.2lifeblog.com

### Swedish Vacation & Volvo

2 Complimentary Tickets to Europe & More - Buy Your Next Volvo Overseas www.volvocars.us/mybagsarepacked

Ads by Google

### Today's most popular articles

- 1. Swedes fake illness to skip work
2. New Miss Sweden risks disqualification
3. Introducing... Marcus Schenkenberg
4. Swedish security police "harasses Muslims"
5. Sexual harassment rife in Swedish parliament - MPs

### Highlights

### Buy train ticket

Departing from...

Going to...

NEXT >>



### Noticeboard

- 24/04 Accommodation rent
24/04 Employment for an extremely hardworking American
24/04 Employment for
24/04 Looking a job
24/04 talkative
23/04 English / french lessons

SHOW ALL NOTICES >>

### Discuss

23:05 International diplomas?? /



# Proxies for engagement

Increasing engagement



- Clicks, click-throughs and click chains
- Page views, repeat visits
- Time elapsed within page and between pages
- Creation of user IDs
- Creation of user profiles
- Content generation
- Content transport through networks
- Downloads, purchases
- Subscriptions

## Challenges:

- Devise and standardize defensible metrics of online engagement
- Use these to predictively devise online experiences





# Structure Matters in Media

- And for large-scale providers there's a better chance it will be provided
- RDF lives even on Yahoo!

The screenshot shows the Yahoo! Food website interface. At the top, there's a navigation bar with 'Food', 'Web', 'Images', 'Video', and 'Audio' tabs. Below that is the 'YAHOO! FOOD' logo and a search bar. The main content area is divided into several sections:

- FOOD HOME**, **EVERYDAY**, **SPECIAL DAYS**, **RECIPES**, **RESTAURANTS & COOKS**
- BUZZ:** brunch, cake, salad, indian, meatloaf, macaroni salad, sangria, pasta, salmon, cookies
- Food & Wine's Grow for Good:** Helping to support local farms and encourage sustainable agriculture. Super Salad Specialists Visit the online auction.
- Quick Recipe Search:** A search bar with a 'SEARCH' button.
- what's for dinner?:** Quick Dinners from *Everyday with Rachael Ray*. Features a 'Big Bird Burger' and a 'Grilling Tips: Heat, Cook, and Glass' section.
- ADVERTISEMENT:** Introducing NEW! Rice Krispies® with Real Strawberries. Includes a 'CLICK HERE TO GET A TASTE' button.
- YAHOO! FOOD VIDEOS:** MOST WATCHED FOOD VIDEOS. Features a video thumbnail for 'Cheesecake Brownies - Kraft'.





# ZoneTag

## ZONETAG™ PHOTOS

[Install](#) [Instructions](#) [FAQ](#) [My ZoneTag](#) [About Us](#)

Take a picture — Instant upload to Flickr with location tags



[Try it now!](#)

[Explore ZoneTag](#)

### ZoneTag Highlights

#### Tagging made easy

ZoneTag suggests likely tags for each photo, making it easy to add tags from your phone and even easier for you to find the photos later. The suggestions are based on tags made by you and your Flickr contacts in similar context, e.g. in the same location.

#### Two-click upload to Flickr

Get your photos uploaded to Flickr in merely two clicks. No more searching for your photo to MMS to your friends or copying them to your computer first.

#### Find your photos by location

ZoneTag can automatically tag your photos with the location, based on the cell tower, in which they were taken. The location of the cell tower is not known until our user community updates the city, state, or zip code of their photos on Flickr.

#### Real-time privacy controls

You can specify on your phone who sees each photo you upload. You can also choose if you want location tags added to each photo.

#### Easy setup

Your first photo will be on Flickr in five minutes. We'll take you through our quick and easy setup step-by-step.



# Video Remix

REMIX GALLERY  
REMIXER TOOL

REMIX HOME  
INFO ABOUT REMIX

## International REMIX

Remix clips of your favorite festival flicks in just a few minutes and add it to the International Remix gallery instantly.

For the latest Remixer news and discussions with other remixers, join our group!

TRY IT NOW



HOT PICK:  
Not Available

NEWEST REMIX:



MORE NEW REMIXES:



INTERNATIONAL REMIX | HOME | GALLERY | INFO | REMIXER | FEEDBACK

powered by Yahoo! Research Berkeley

Copyright © 2006 Yahoo! Inc. All rights reserved.



# Towards a Science of Data

- Today:
  - Yahoo! adds ~25TB per *day*  
(some agencies do >>25TB)
  - And then ... what?
- What are the useful attributes?
- How do we efficiently cull them?

March 1997:

“Walmart currently operates a 7TB system, the largest in the world.”



# The Tyranny of Moore's Law

- Storage gets cheaper faster than communications gets cheaper faster than processing
- Today: Statistical Machine Learning plus lots of heuristics
  - struggle to keep pace
- Consumer Privacy
  - AOL mishap



# Conclusions

- For large open-Web companies, things continue to move strongly towards “Web 2.0” applications
  - Relationships are fundamental driver
  - Social element is very significant
- Huge and growing amount of content, metadata is user-produced and uncontrollable
  - *Opportunities to meet halfway*



- Don't forget advertising: it fuels the machine
  - *Huge matching and inference problems*
    - *Inferring context, intent, relevant history...*
    - *Background knowledge could become important*
- New sciences are evolving
  - Science of finding things
  - Community science (and systems)
  - Algorithmic advertising
  - Computational microeconomics
  - Media experiences
  - Data science





- Towards a science of aggregated, augmented organic knowledge
  - Extraction -> integration/rationalization
  - Wrapper induction
  - Reuse/adaptation of ontologies
  - Meaning in an inconsistent world
    - Insights from belief revision
    - Collective intelligence vs. collective confusion
  - Learn from projects like Cyc
  - Role of background knowledge
    - How to avoid the “AI-complete” problem?

Let's skate to where the puck is going to be!



**Thank you!**