Emerging Sciences of the Internet: Some New Opportunities

Ron Brachman

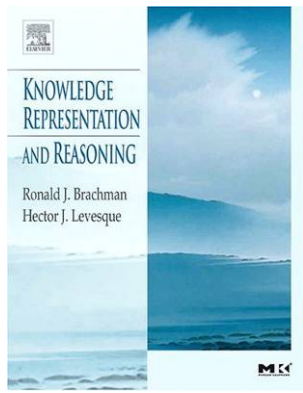
Vice President,
Worldwide Research Operations

YAHOO! RESEARCH

4th 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

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

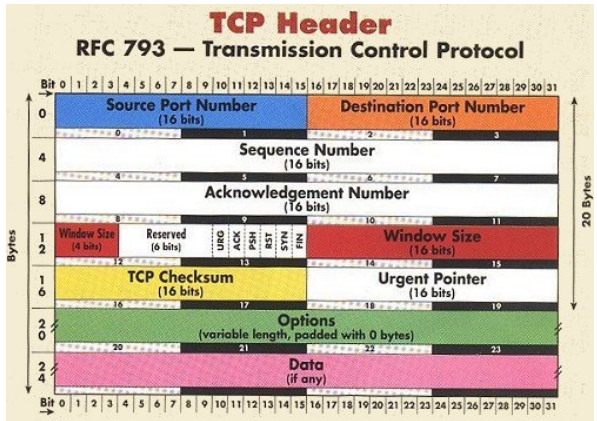
- In Yahoo! Research, we're focusing on the deep, underlying knowledge needed to move ahead – the sciences behind Web n.0
- I think there are interesting opportunities here…

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

"Stuff" = data/knowledge, help, things, entertainment

Example views:

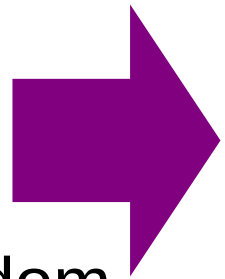
- searching/finding
- communicating (email, IM)
- socializing/sharing (groups, communities)
- buying/selling
- matchmaking/ recommending
- experiencing media

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

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

1994-1997 AV, Excite, Lycos, etc

- **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)

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
  - Downloads
  - Shop

  Innsbruck weather

  MP3 to WMA converter

  Canon SD550

  VOIP companies

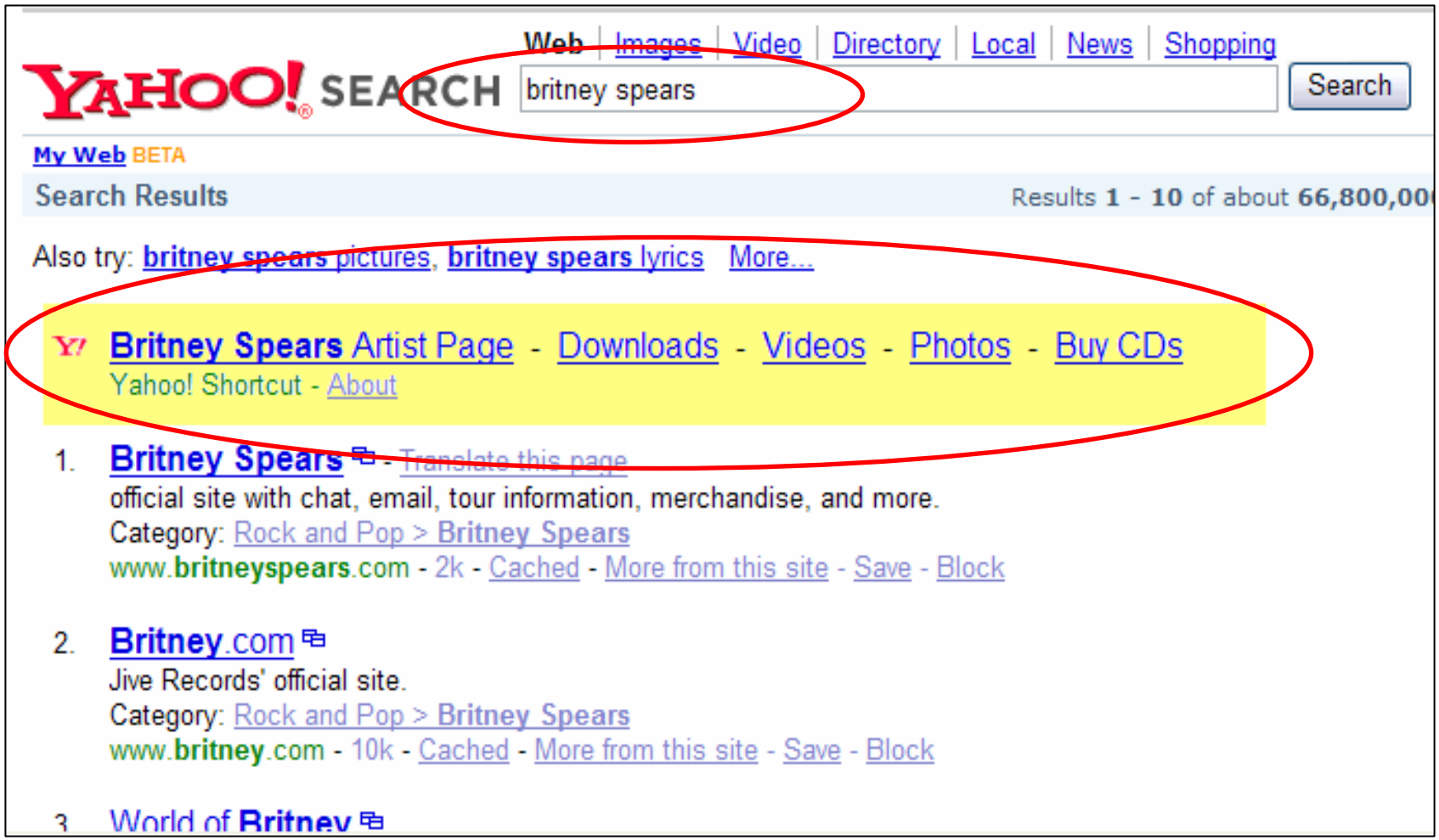
- Gray areas
  - Find a good hub
  - Exploratory search to “see what’s there”

Courtesy: Andrei Broder

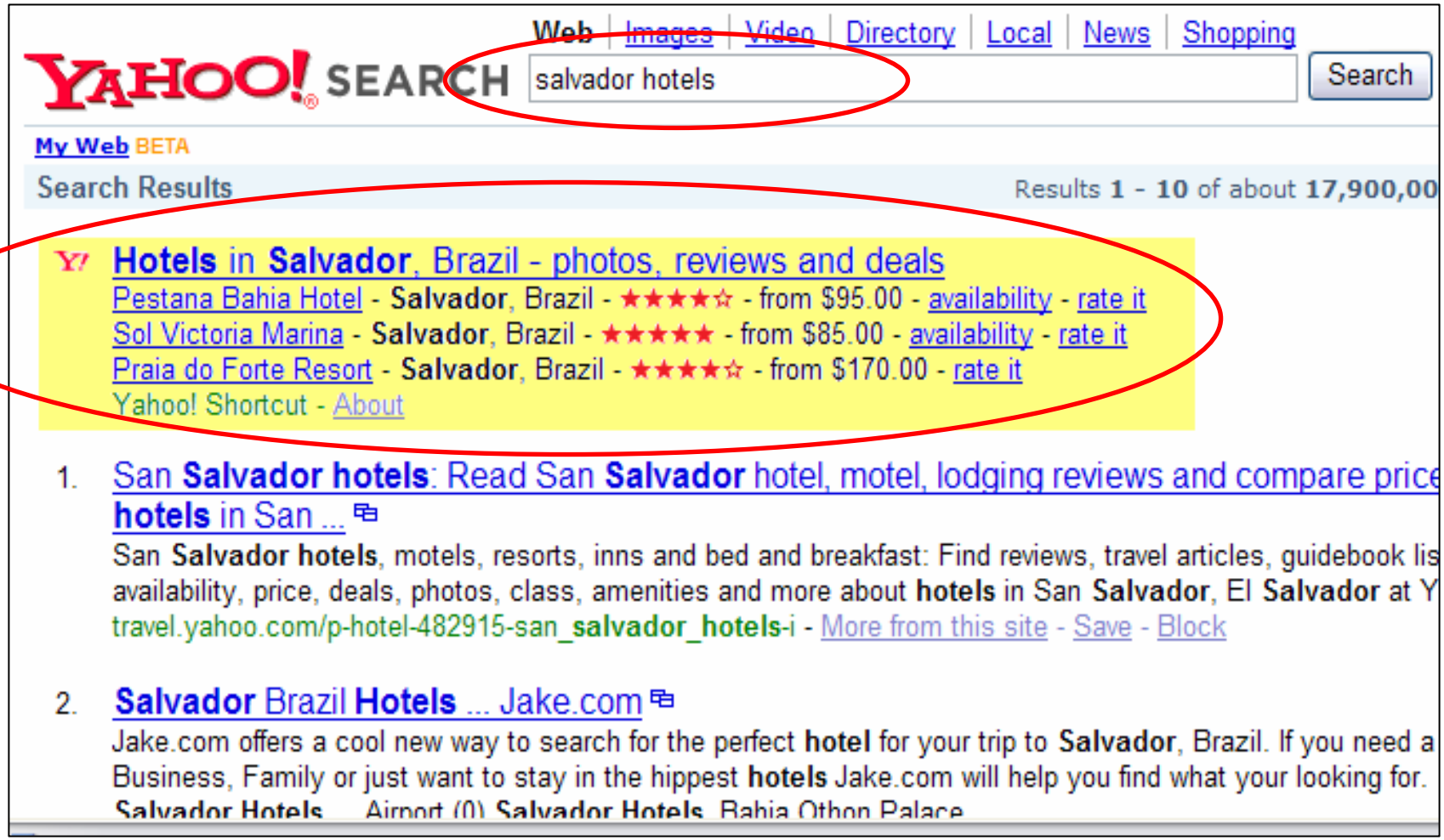
Third Generation Search

- 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

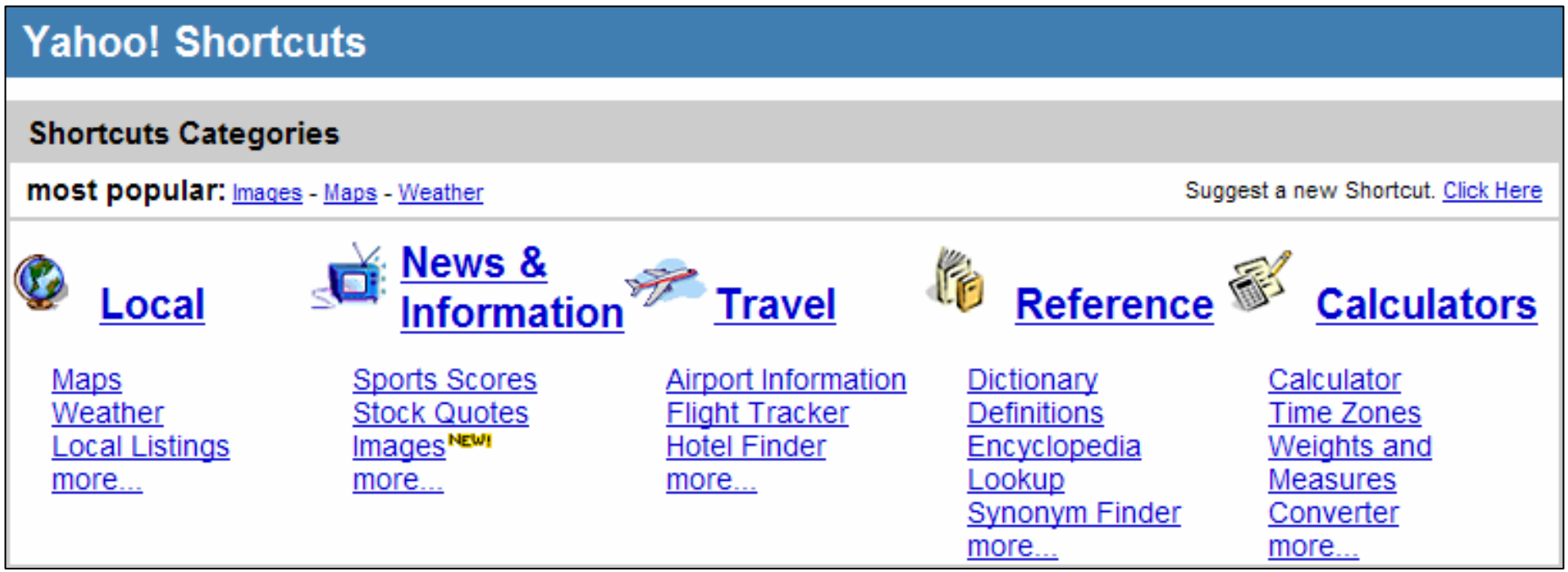


Yahoo!: salvador hotels



Yahoo Shortcuts

- Various types of queries that are “understood”



Google:
andrei broder new york

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"

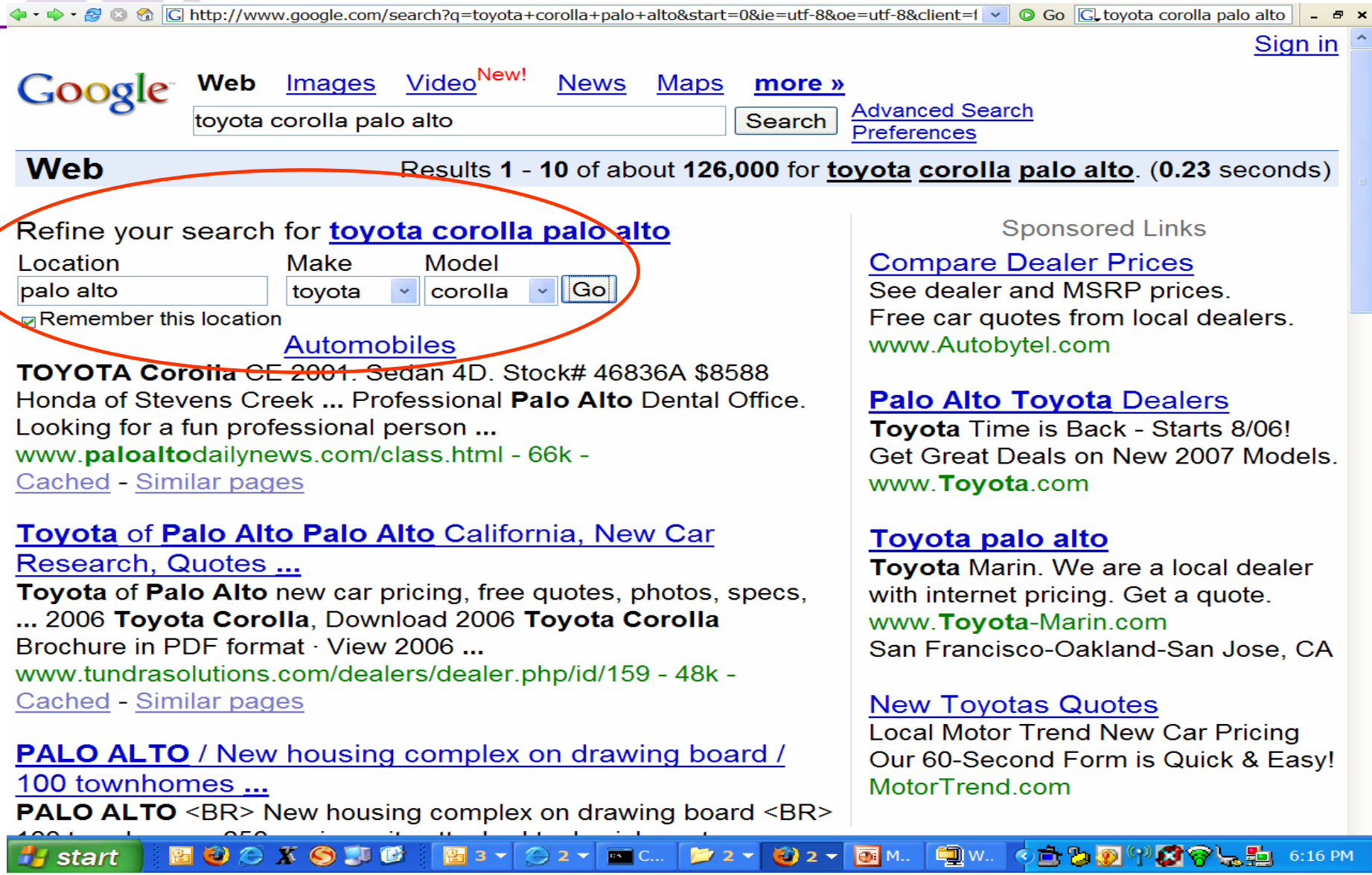
CLASSIFIERS (e.g., SVM)

Category: restaurant
Location: San Francisco

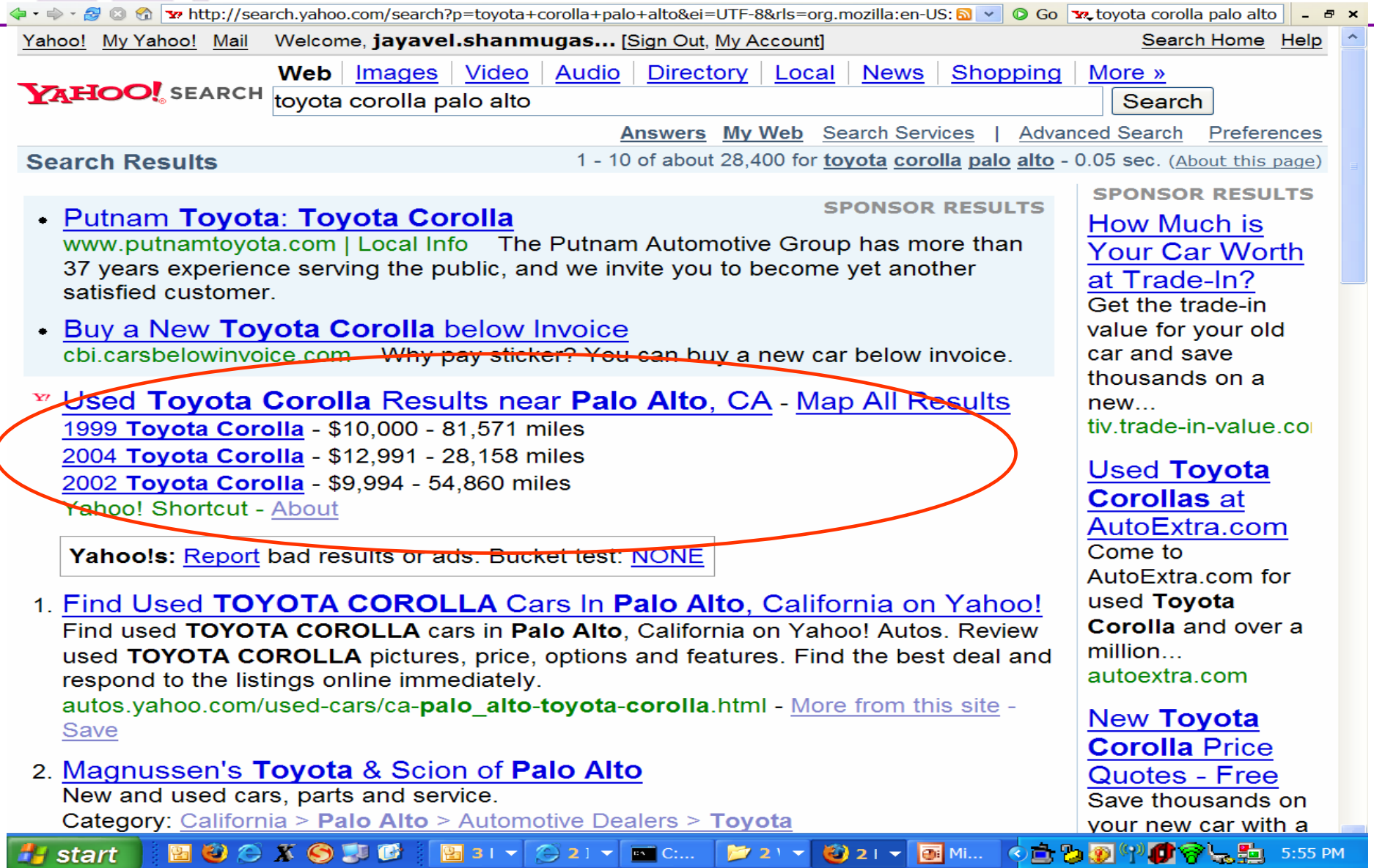
- Can apply machine learning (ML) to extract structure from user context (query, session, …), content (web pages), and ads
- Alternative: We can elicit structure from users in a variety of ways

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
Google
Web
Images
Video New!
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.paloaltodailynews.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 - 48k -
Cached - Similar pages
PALO ALTO / New housing complex on drawing board / 100 townhomes ...
PALO ALTO
 New housing complex on drawing board
Sponsored Links
Compare Dealer Prices
See dealer and MSRP prices.
Free car quotes from local dealers.
www.Autobytel.com
Palo Alto Toyota Dealers
Toyota Time is Back - Starts 8/06!
Get Great Deals on New 2007 Models.
www.Toyota.com
Toyota palo alto
Toyota Marin. We are a local dealer with internet pricing. Get a quote.
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
start
6:16 PM

Y! Shortcuts


http://search.yahoo.com/search?p=toyota+corolla+palo+alto&ei=UTF-8&rls=org.mozilla:en-US:
Go
toyota corolla palo alto
Yahoo! My Yahoo! Mail Welcome, jayavel.shanmugas... [Sign Out, My Account] Search Home Help
YAHOO! SEARCH
Web Images Video Audio Directory Local News Shopping More »
toyota corolla palo alto
Search
Answers My Web Search Services | Advanced Search Preferences
Search Results
1 - 10 of about 28,400 for toyota corolla palo alto - 0.05 sec. (About this page)
SPONSOR RESULTS
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.
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
Yahoo!s: 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
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 Corollas at AutoExtra.com
Come to AutoExtra.com for used Toyota Corolla and over a million...
autoextra.com
New Toyota Corolla Price Quotes - Free
Save thousands on your new car with a
start
5:55 PM

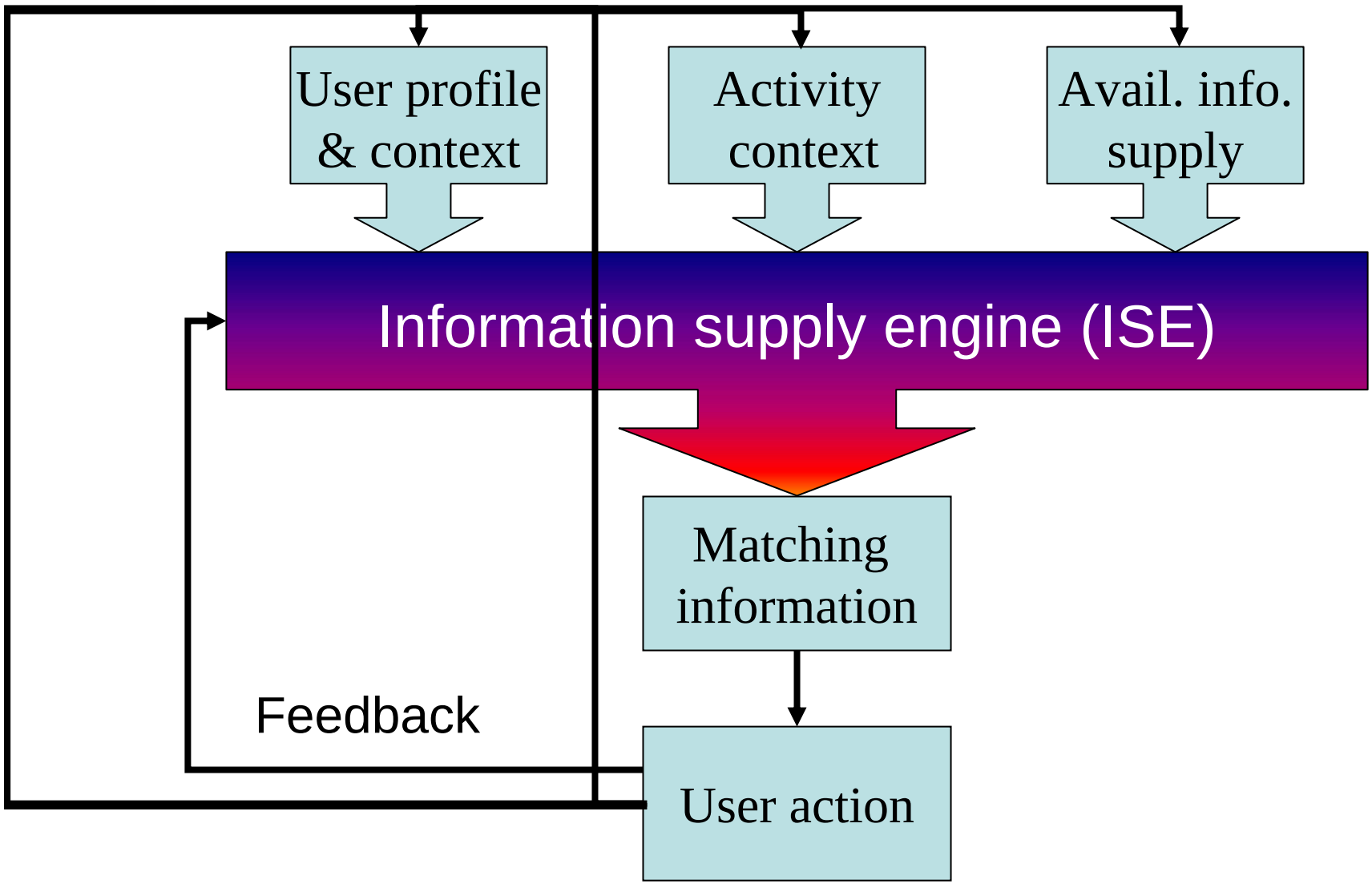
From Information Retrieval to Information Supply

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 Search

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

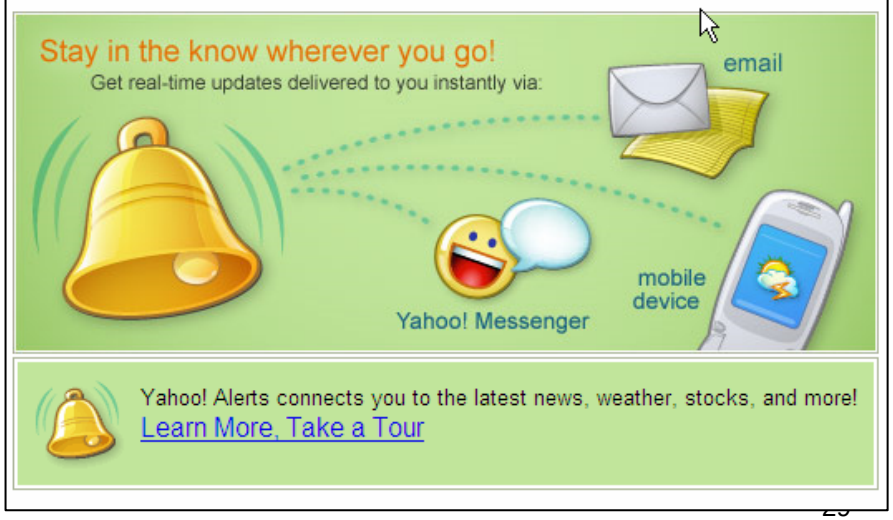
Alert me when articles match all these words: **andrei broder**

ADD TO MY YAHOO! / RSS

Add your news search for **andrei broder** to My Yahoo!: MY YAHOO!

View as RSS: XML

YAHOO! ALERTS



Annotations

"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

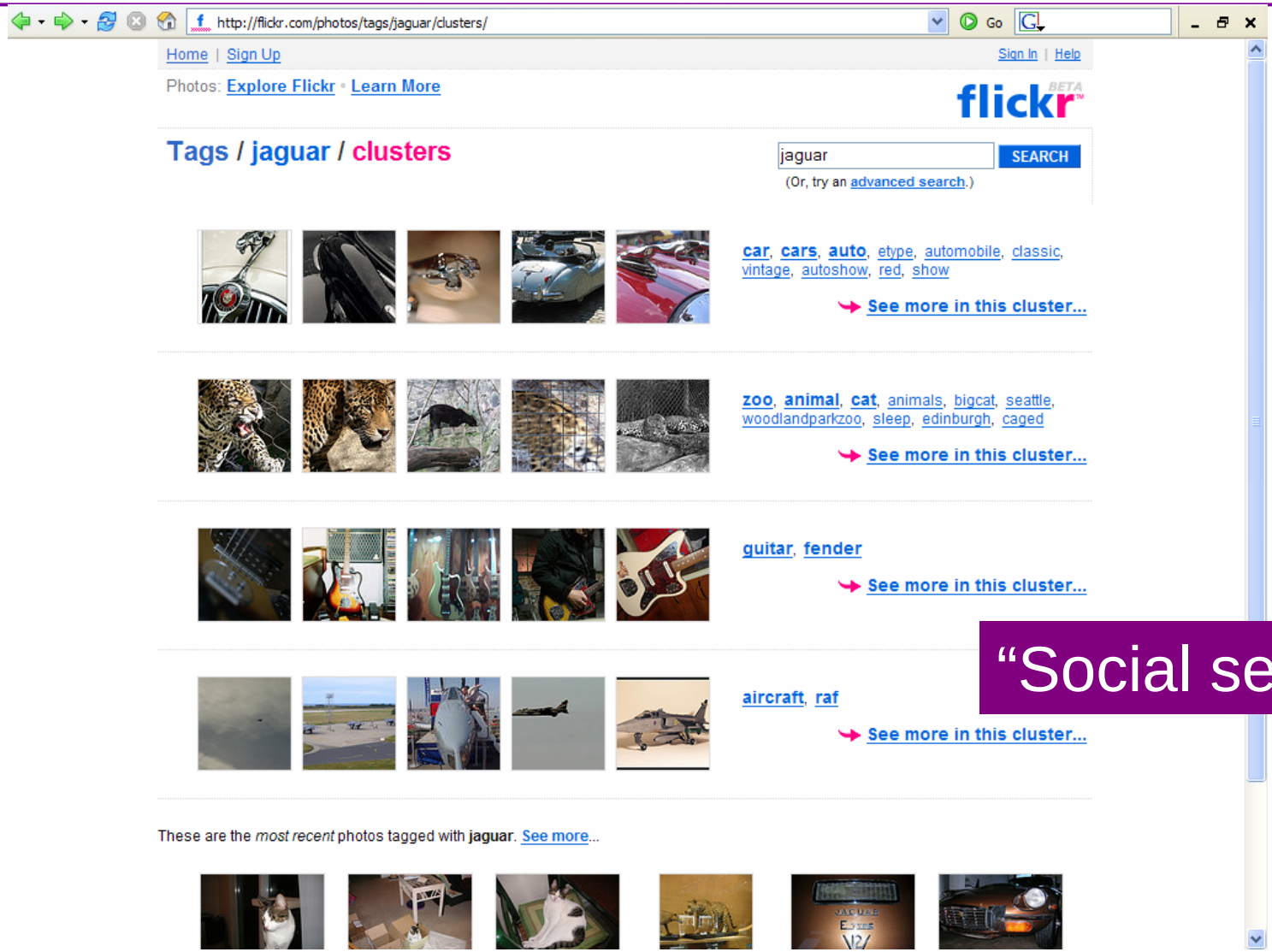
- “Interested in the following accessories for your iPod?:…”
- “Now that you’ve booked your flight, would you like to rent a car?”
- “You seem to be writing a business letter…”
- Contextual and search-driven ads
- Recommendations

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

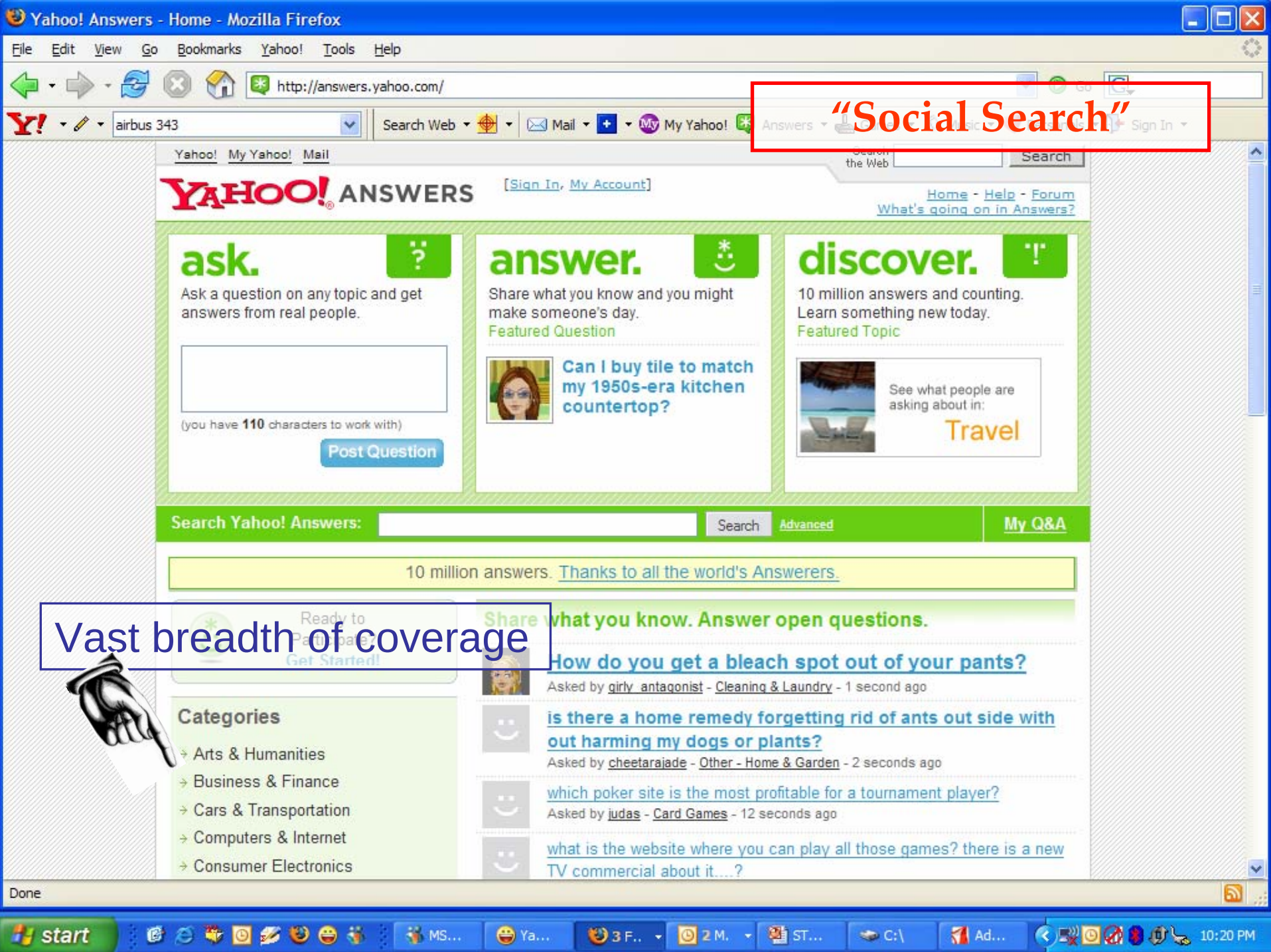


"Social search"

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


Yahoo! Answers - Home - Mozilla Firefox
File Edit View Go Bookmarks Yahoo! Tools Help
http://answers.yahoo.com/
airbus 343
Search Web
Mail
My Yahoo!
Answers
Sign In
"Social Search"
Yahoo! My Yahoo! Mail
Search the Web
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
Ready to Participate?
Get Started!
Share what you know. Answer open questions.
How do you get a bleach spot out of your pants?
Asked by girly_antagonist - 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....?
Categories
Arts & Humanities
Business & Finance
Cars & Transportation
Computers & Internet
Consumer Electronics
Done
start
10:20 PM

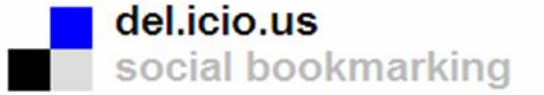
Social Search Challenges

- 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

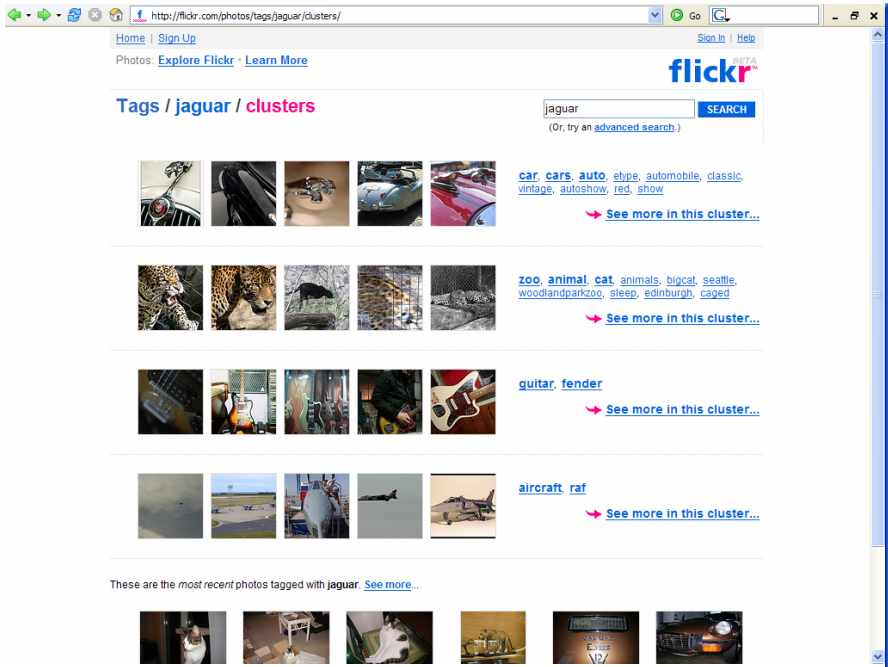
Community Science and Systems

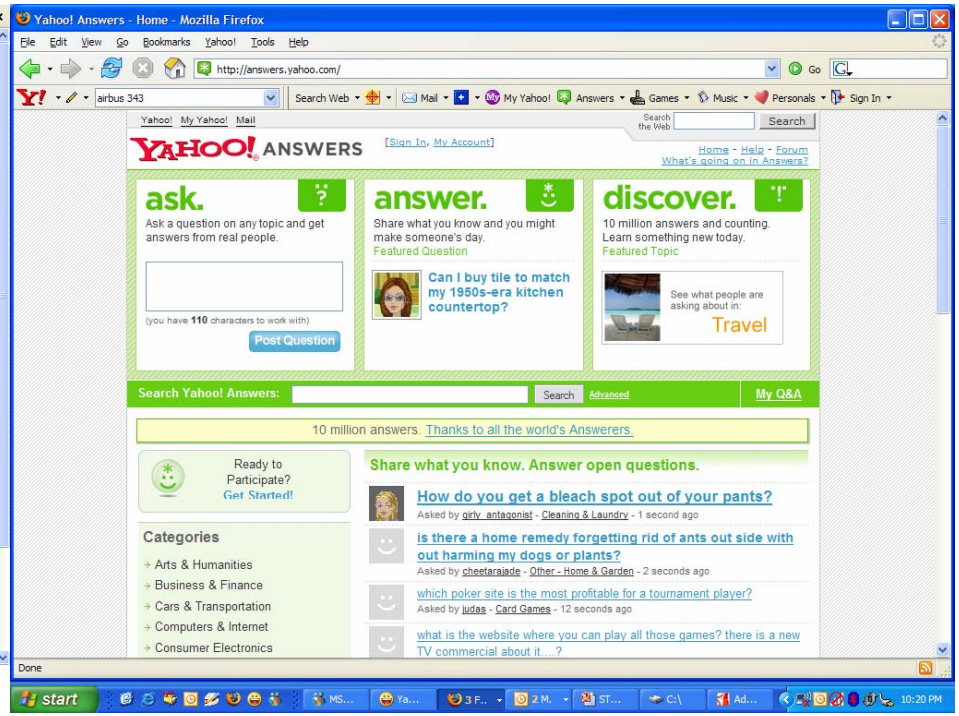
- "Social Media"
  - Different types of communities with different needs
  - Sociology, Network Theory



- Distributed/collaborative systems
  - Marriage of automatic extraction with structured augmentation
  - Imposing unifying views

The Power of Community





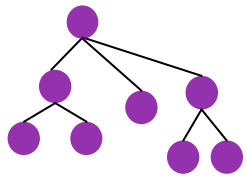
Millions of users share and tag each others' photographs

Surpassed 60 million users in twelve months

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

Understanding Online Communities

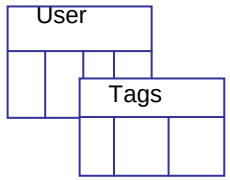
- Exploratory analysis of massive datasets of online community usage
- Challenges
  - Analyze evolving social networks of users, content and interactions to
    - Learn models of community structure and dynamics
    - Learn models of individual characteristics and preferences
    - Develop robust frameworks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



Four Types of Communities

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*



http://groups.yahoo.com/

Google search

Yahoo! My Yahoo! Mail Make Y! your home page Search: Web Search

YAHOO! GROUPS

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

Groups Home - Help

My Groups

Manage

yg-research
(No new activity)

YAHOO! GROUPS RESEARCH

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

meditation 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

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

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]

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

Community knowledge

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

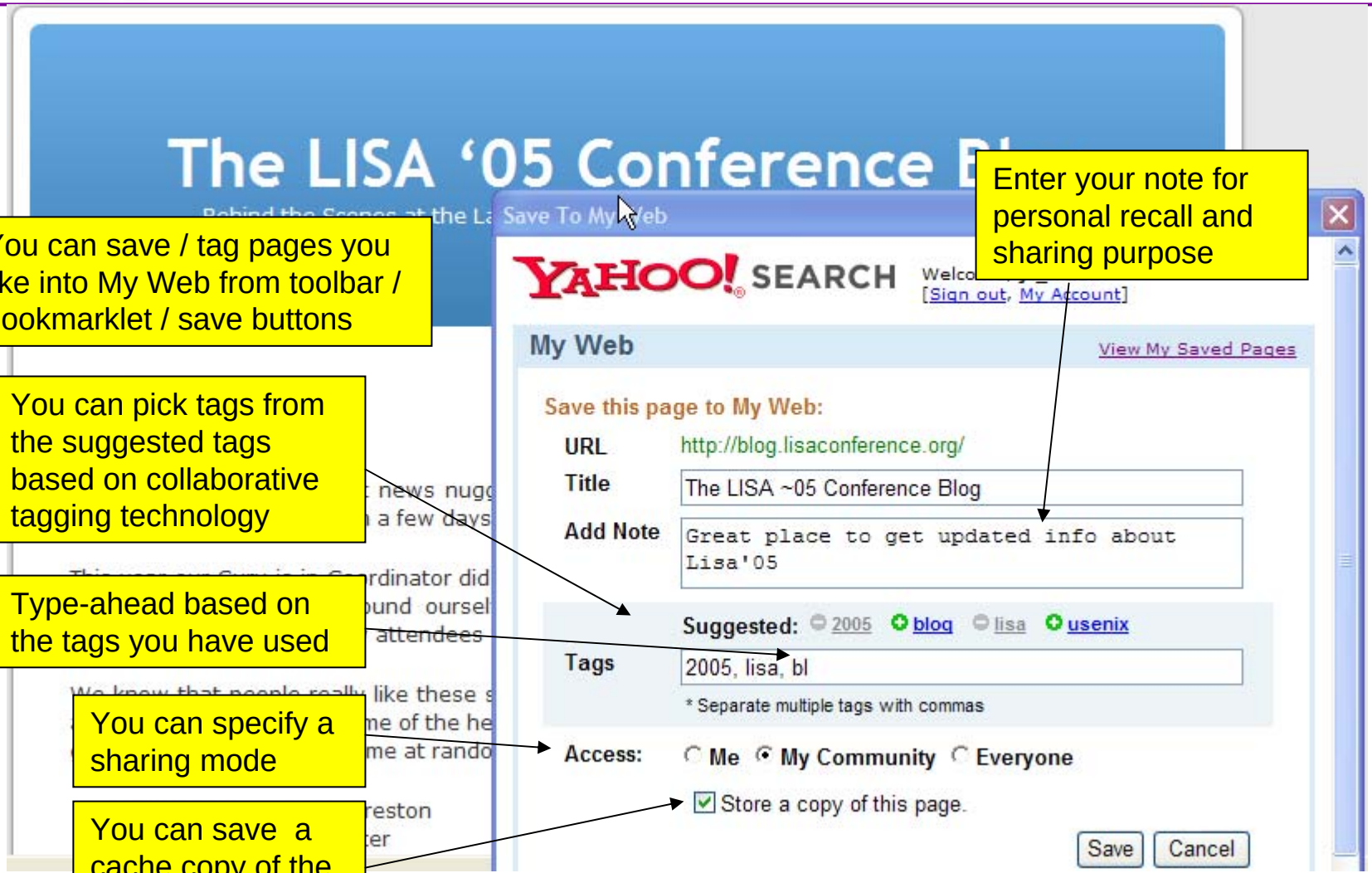
Recent Pages | Popular Pages

Tips for Working Securely from Hotspots
Shared by: hermosa1001 Tags: security, wireless 14 minutes ago - Details - Save

Public vitality
- Statistics
- Recent pages
- Popular pages

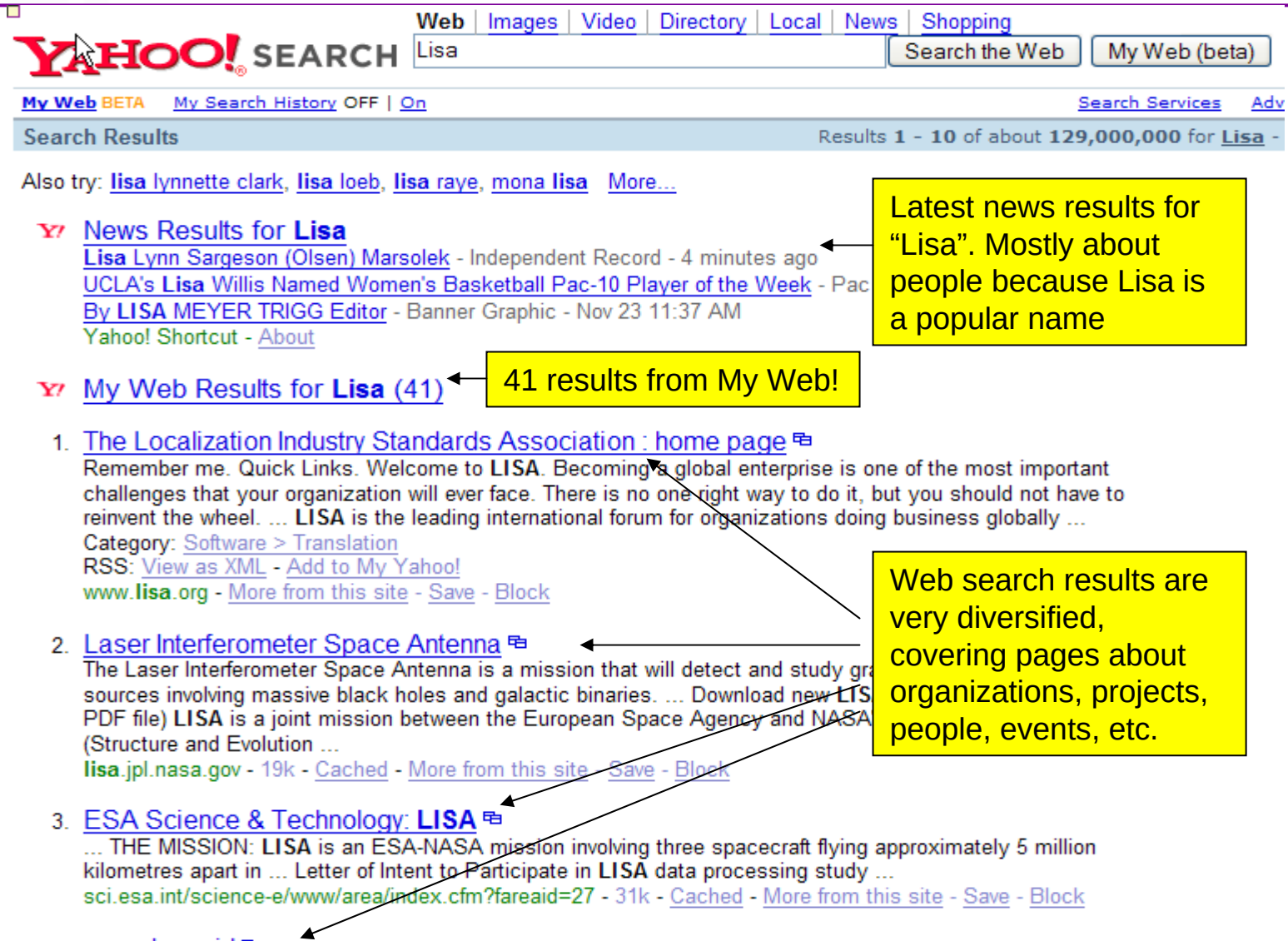
(Courtesy: Raymie Stata)

Save/Tag Pages You Like



(Courtesy: Raymie Stata)

Unaugmented Results



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

41 results from My Web!

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

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. Novembe
   (here's another burning-hot news nugget that will be posted on the official conference site wit
   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

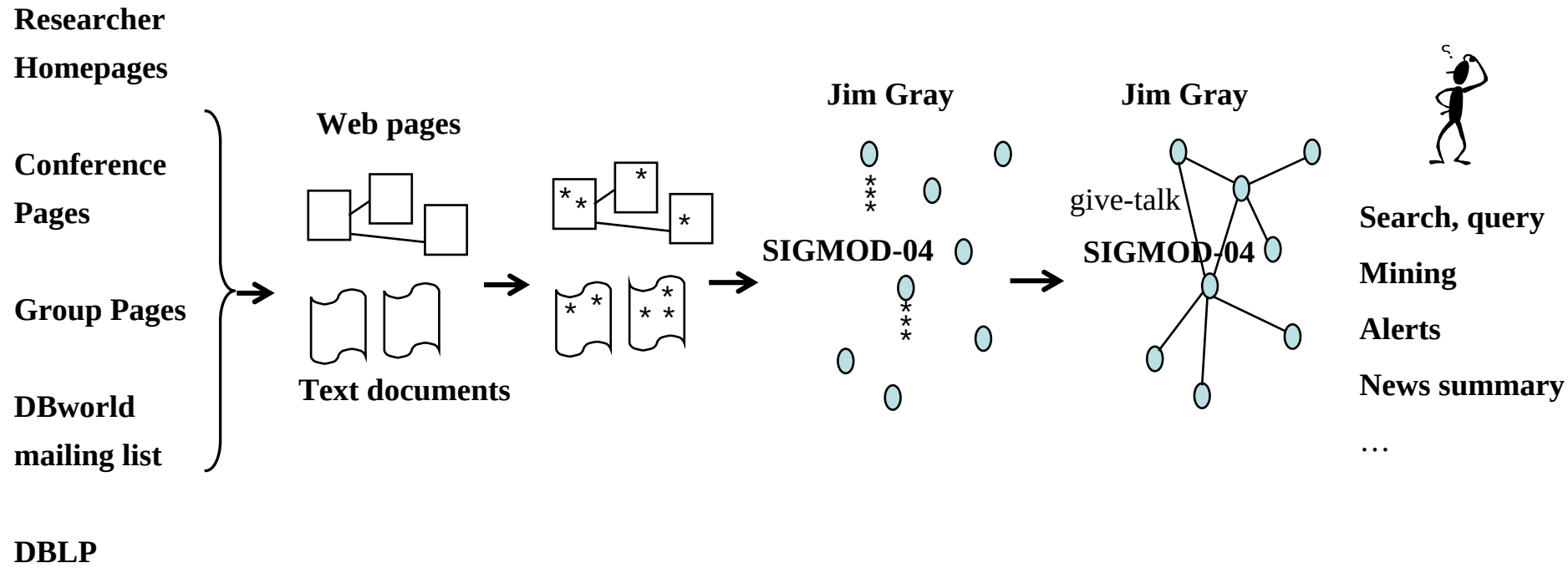
Excellent set of search results from my community because a couple of people in my community are interested in Usenix Lisa-related topics

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



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?

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

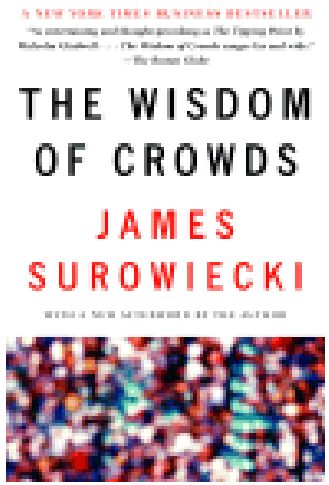
Community Systems Opportunities and Challenges

- 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

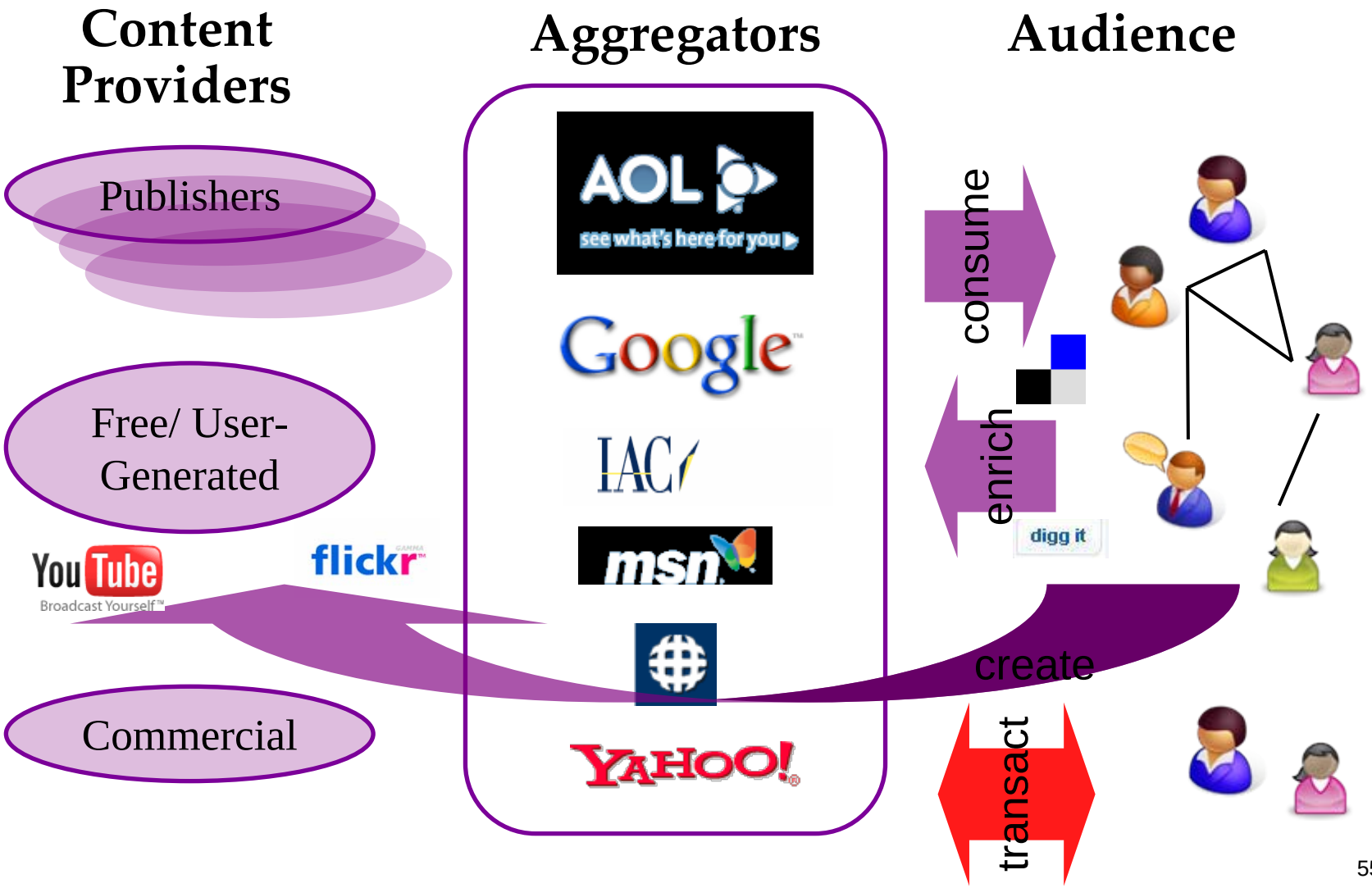
All at Web scale!

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

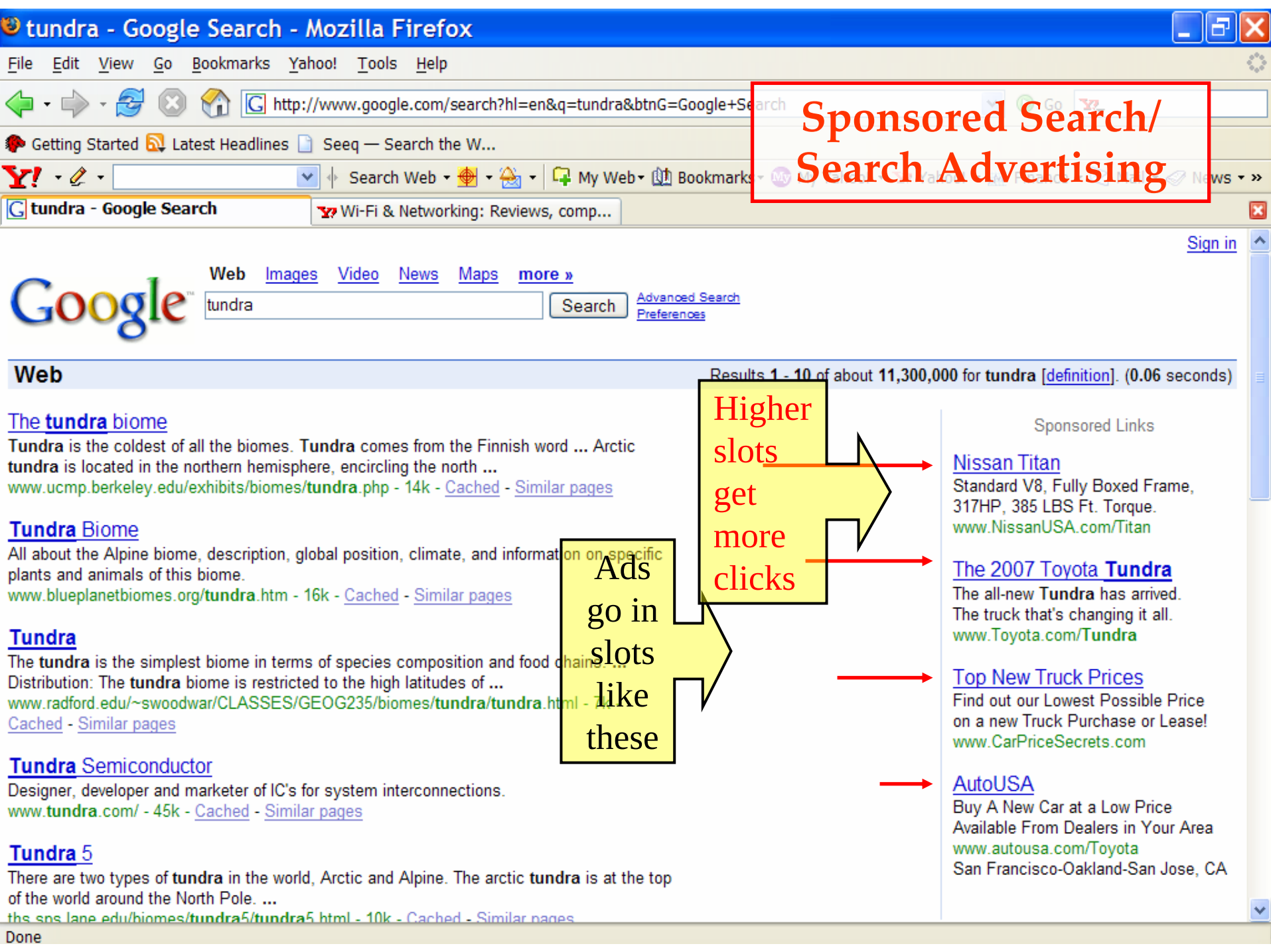
- Aggregators draw content consumers
  - Search is a significant “hook”
- Consumer reveals clues about his current situation
  - The keyword(s) he types (e.g., ***miele***)
  - Keyword(s) in his email (gmail)
  - Context information (Yahoo! …)

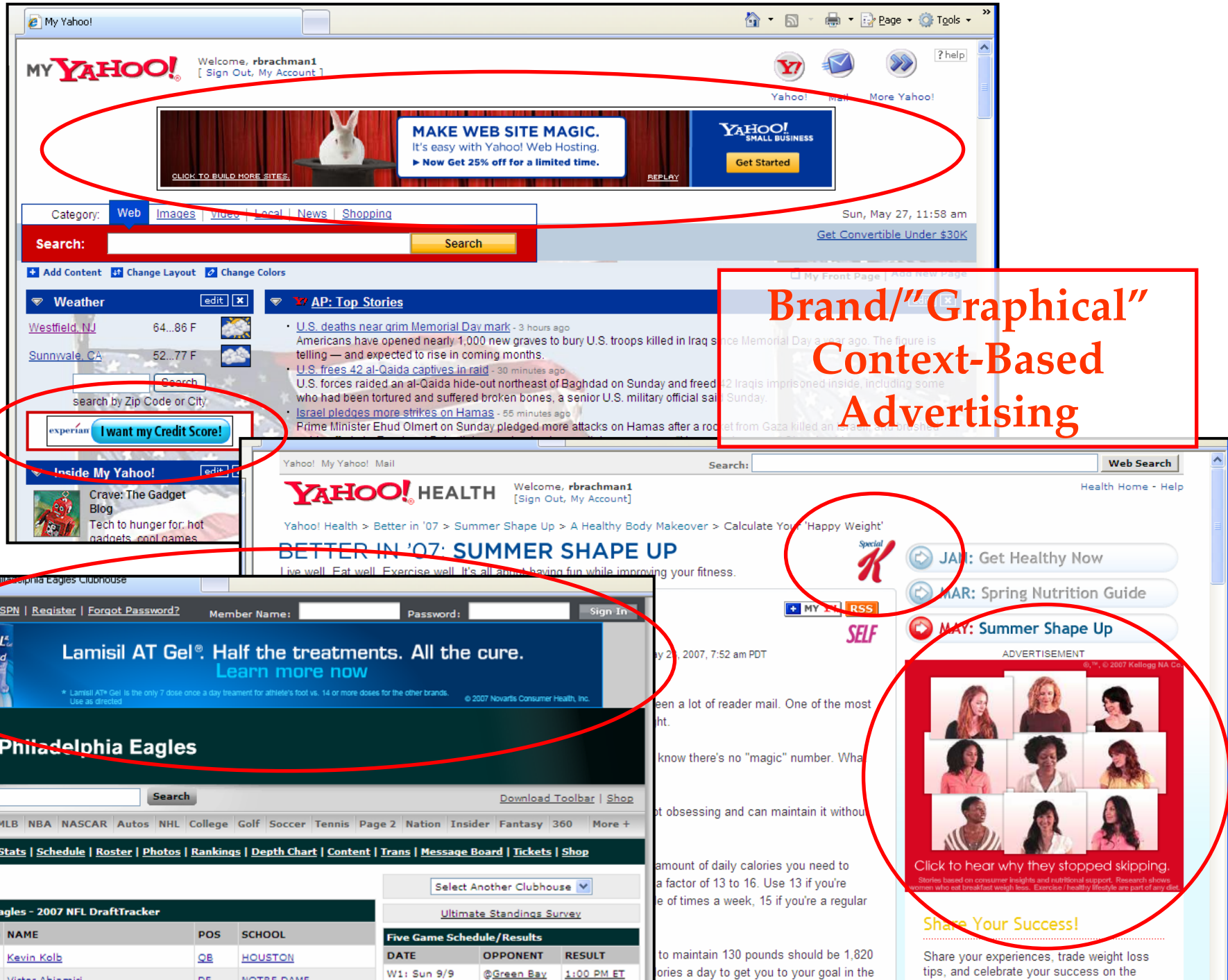
Paid Placement

- 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?


tundra - Google Search - Mozilla Firefox
File Edit View Go Bookmarks Yahoo! Tools Help
http://www.google.com/search?hl=en&q=tundra&btnG=Google+Search
Getting Started Latest Headlines Seeq — Search the W...
Search Web My Web Bookmarks
tundra - Google Search
Wi-Fi & Networking: Reviews, comp...
Sponsored Search/ Search Advertising
Sign in
Google
Web Images Video News Maps more »
tundra
Search
Advanced Search
Preferences
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/GEOG235/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.sps.lane.edu/biomes/tundra5/tundra5.html - 10k - Cached - Similar pages
Higher slots get more clicks
Ads go in slots like these
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
Done

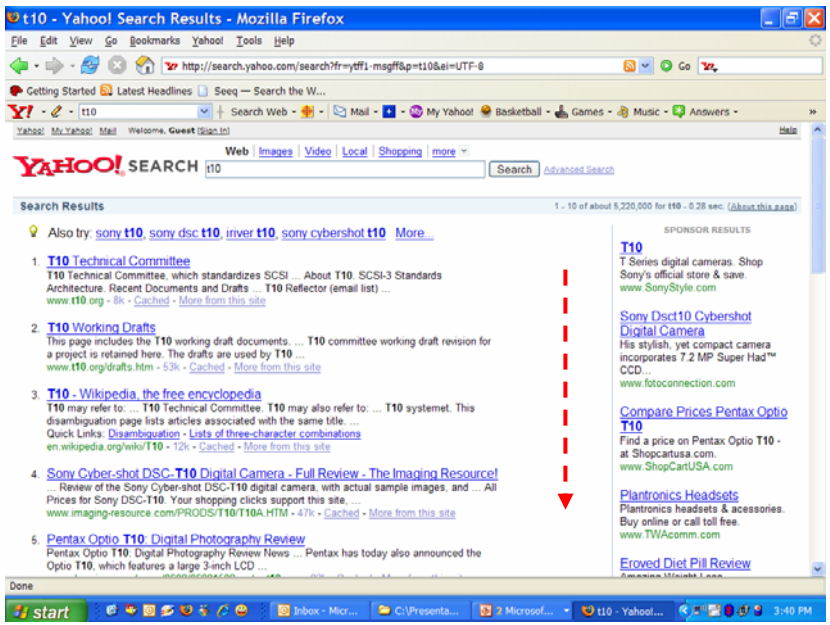

My Yahoo!
MY YAHOO!
Welcome, rbrachman1
[Sign Out, My Account]
MAKE WEB SITE MAGIC.
It's easy with Yahoo! Web Hosting.
▶ Now Get 25% off for a limited time.
CLICK TO BUILD MORE SITES.
REPLAY
YAHOO! SMALL BUSINESS
Get Started
Category: Web | Images | Video | Local | News | Shopping
Search:
Search
Sun, May 27, 11:58 am
Get Convertible Under $30K
Add Content
Change Layout
Change Colors
Weather
Westfield, NJ 64...86 F
Sunnyvale, CA 52...77 F
search by Zip Code or City
experian I want my Credit Score!
Inside My Yahoo!
Crave: The Gadget Blog
Tech to hunger for: hot gadgets, cool games
AP: Top Stories
U.S. deaths near grim Memorial Day mark - 3 hours ago
Americans have opened nearly 1,000 new graves to bury U.S. troops killed in Iraq
telling — and expected to rise in coming months.
U.S. frees 42 al-Qaida captives in raid - 30 minutes ago
U.S. forces raided an al-Qaida hide-out northeast of Baghdad on Sunday and freed
who had been tortured and suffered broken bones, a senior U.S. military official sai
Israel pledges more strikes on Hamas - 55 minutes ago
Prime Minister Ehud Olmert on Sunday pledged more attacks on Hamas after a roc
Brand/"Graphical" Context-Based Advertising
Yahoo! My Yahoo! Mail
Search:
Web Search
YAHOO! HEALTH
Welcome, rbrachman1
[Sign Out, My Account]
Health Home - Help
Yahoo! Health > Better in '07 > Summer Shape Up > A Healthy Body Makeover > Calculate Your 'Happy Weight'
BETTER IN '07: SUMMER SHAPE UP
Live well. Eat well. Exercise well. It's all about having fun while improving your fitness.
Special K
SELF
JAN: Get Healthy Now
MAR: Spring Nutrition Guide
MAY: Summer Shape Up
ADVERTISEMENT
Click to hear why they stopped skipping.
Share Your Success!
Share your experiences, trade weight loss tips, and celebrate your success on the
ESPN.com: Philadelphia Eagles Clubhouse
ESPN.com | MyESPN | Register | Forgot Password?
Member Name:
Password:
Sign In
NEW
LAMISIL Advanced
Lamisil AT Gel®. Half the treatments. All the cure.
Learn more now
* Lamisil AT® Gel is the only 7 dose once a day treament for athlete's foot vs. 14 or more doses for the other brands. Use as directed
© 2007 Novartis Consumer Health, Inc.
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
Select Another Clubhouse
Ultimate Standings Survey
Philadelphia Eagles - 2007 NFL DraftTracker
RD PK(OVR) NAME POS SCHOOL
2 4(36) Kevin Kolb QB HOUSTON
Five Game Schedule/Results
DATE OPPONENT RESULT
W1: Sun 9/9 @Green Bay 1:00 PM ET

Three Subproblems

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

IR+

Econ



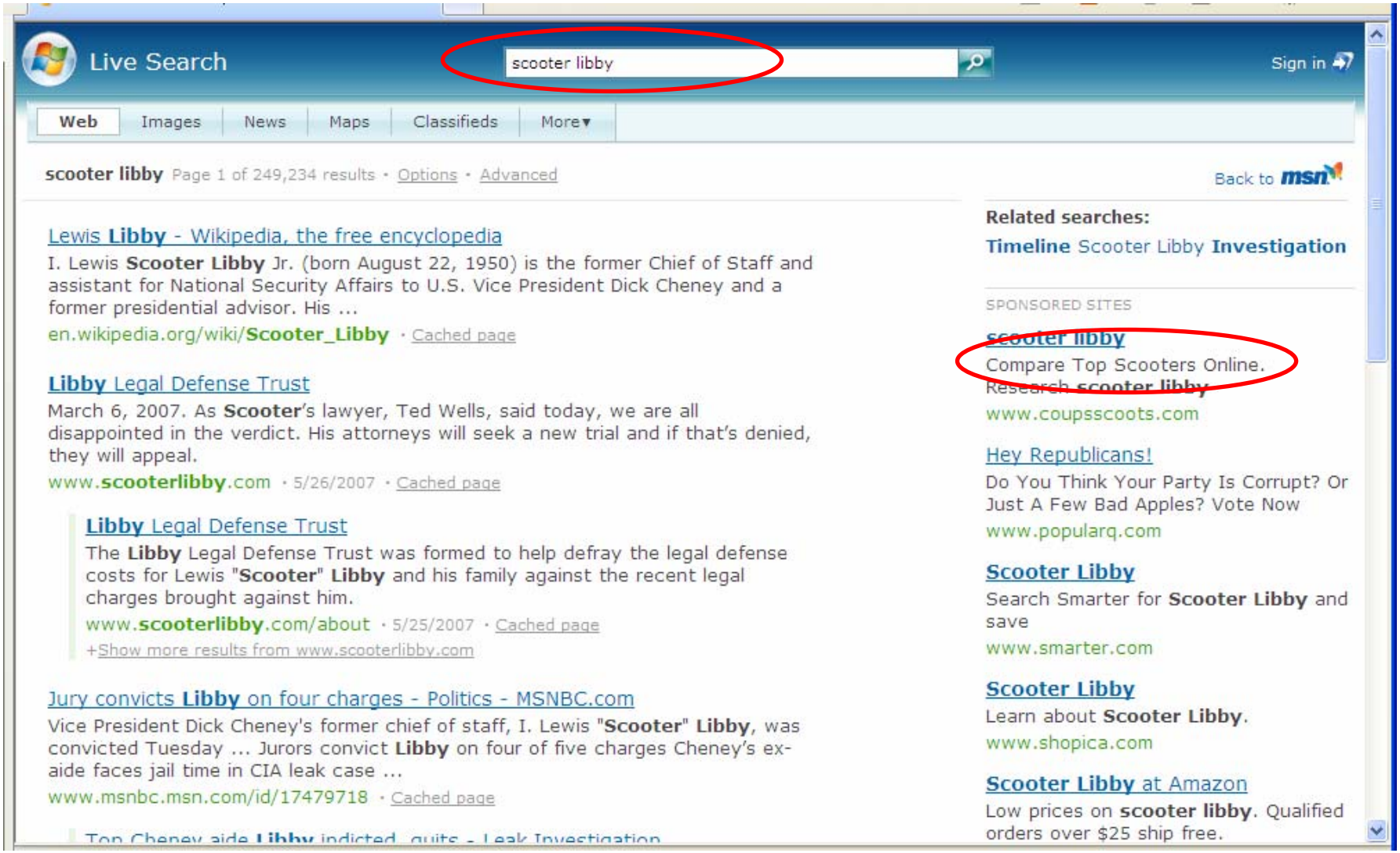
The Context Match Challenge

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



Content Page Context

Slate

SEARCH

Home ▷ Politics

Ask Web search SEARCH

News & Politics
Arts & Life
Business & Tech
Health & Science
Style & Shopping
Travel & Food
Sports
Slate on NPR

PRINT
DISCUSS
E-MAIL
RECOMMEND
NEWSLETTERS
PODCASTS
RSS

Together

POLITICS: WHO'S WINNING, WHO'S LOSING, AND WHY.

Who Is Scooter Libby?

THE SECRETIVE CHENEY AIDE AT THE HEART OF THE CIA LEAK CASE.

By John Dickerson

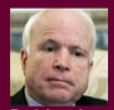
Posted Friday, Oct. 21, 2005, at 6:57 PM ET

Who is I. Lewis Libby? The not-Karl-Rove character at the center of the CIA leak investigation is so mysterious he hides his first name. Rove we know: He's Bush's political id—a self-taught master of political hardball, a brash Texan who has plotted the president's advance for 25 years.

• • •

ALSO IN SLATE

Help! My Neighbor Is Ned Flanders.



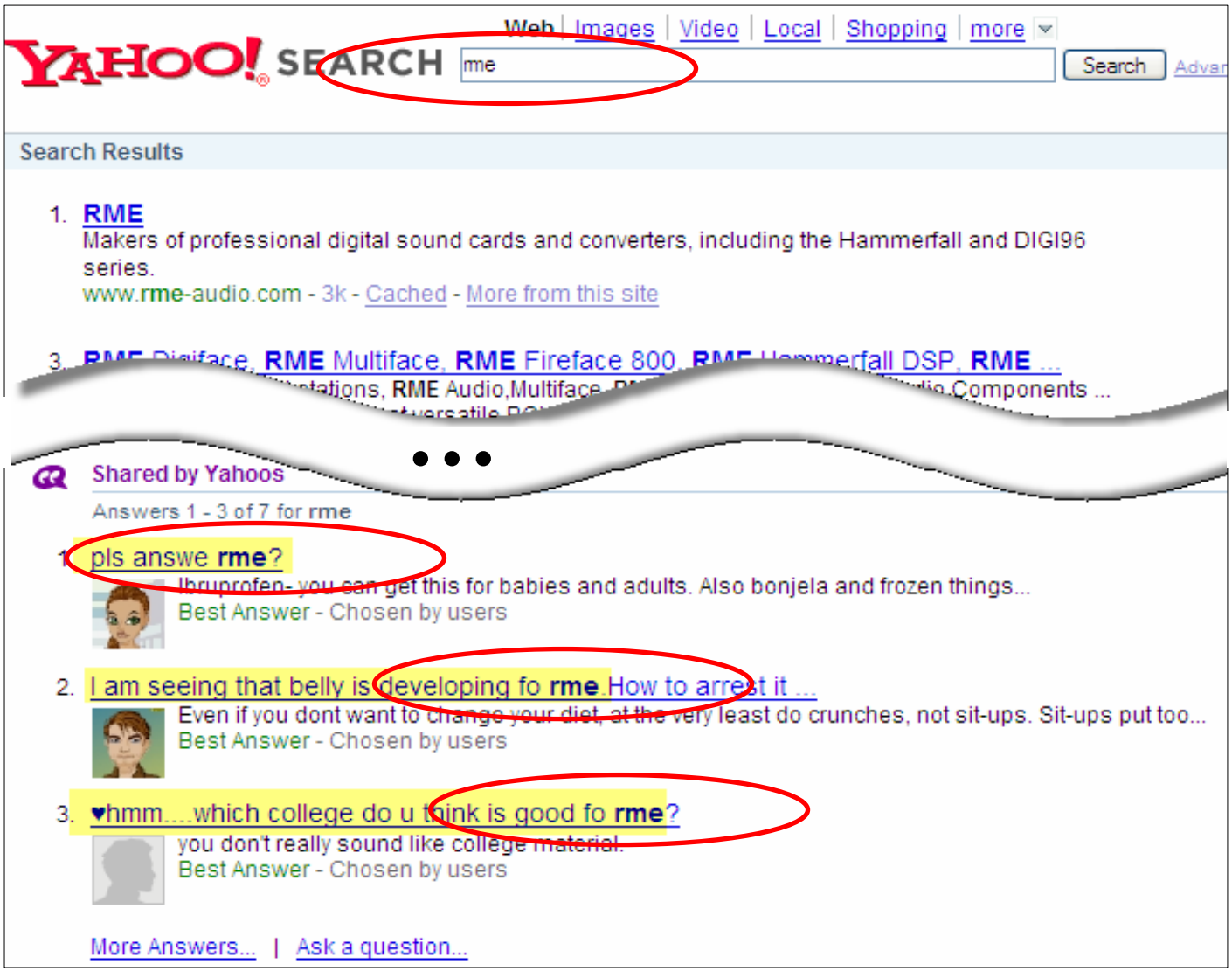
Explainer: Is McCain's Brain Too Old for the Presidency?

The Snob Appeal of Tap Water

Enjoying the Phil Spector Trial Way Too Much

build your own slate
feedback | about us | help | advertise | newsletters | mobile

Other Match Challenges



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

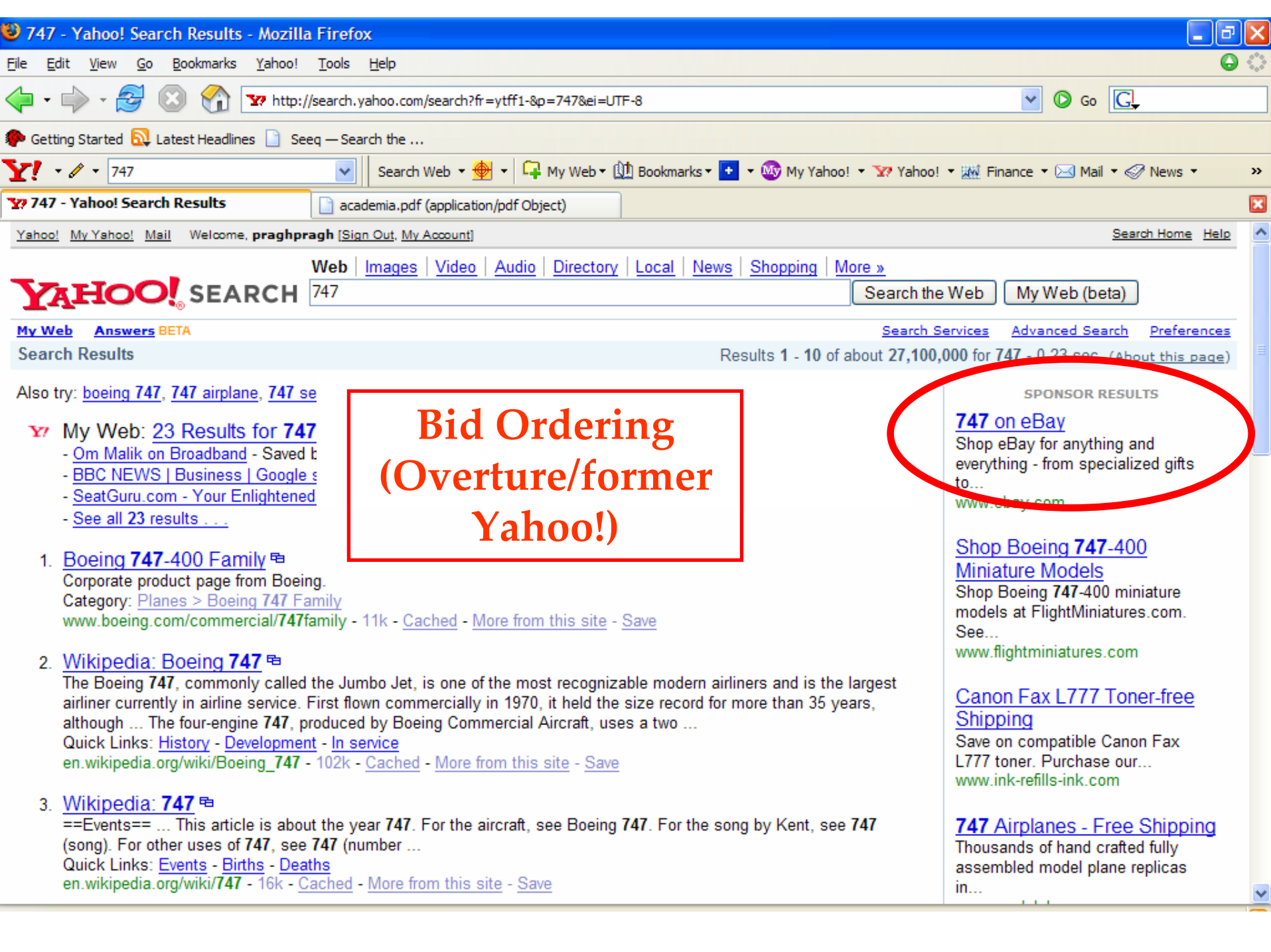
- There may be very few clues

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

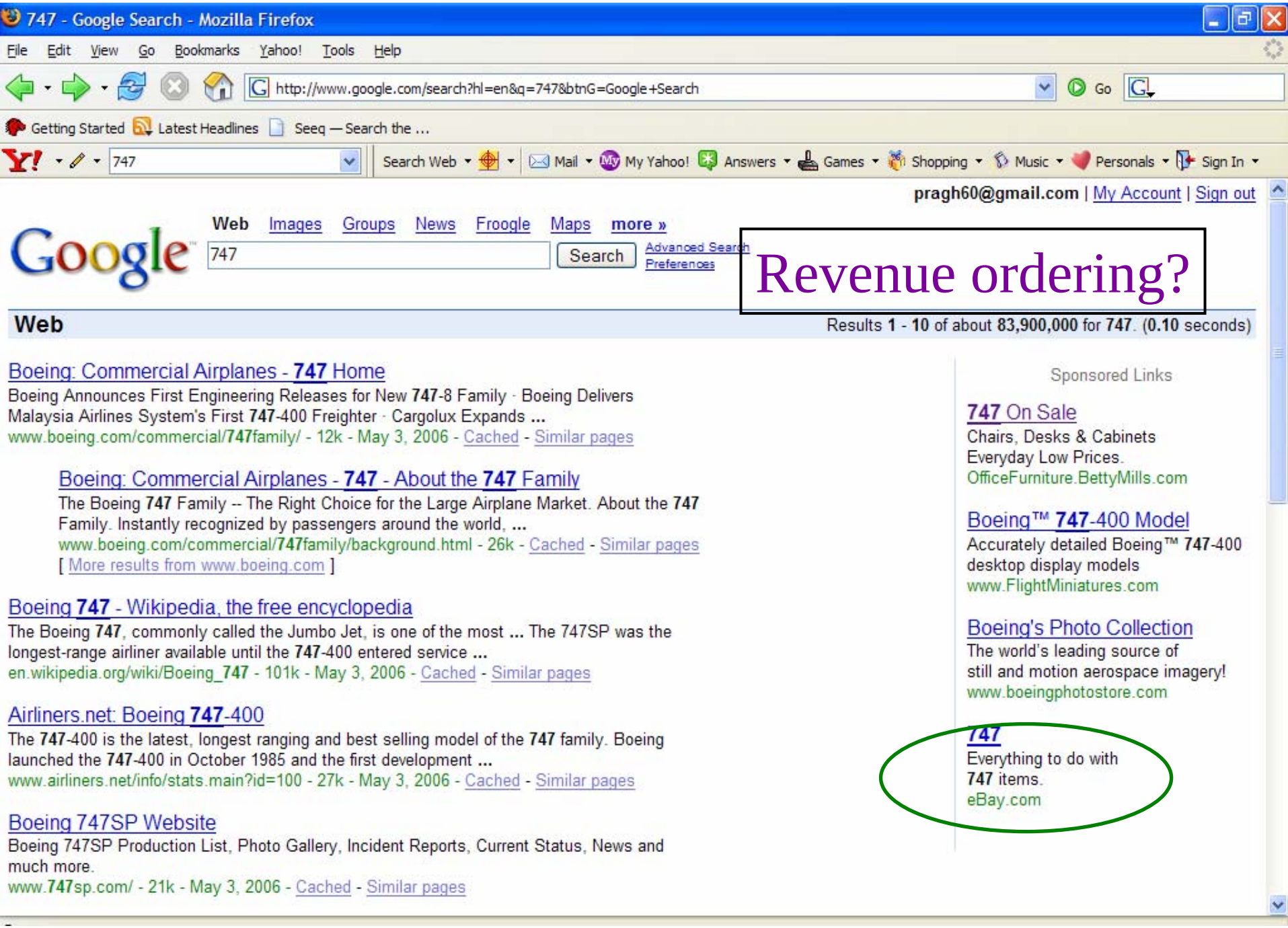
- Most generally, composite IR+Econ score … but today, basically Econ
- Original GoTo/Overture scheme:
  - Order by bid

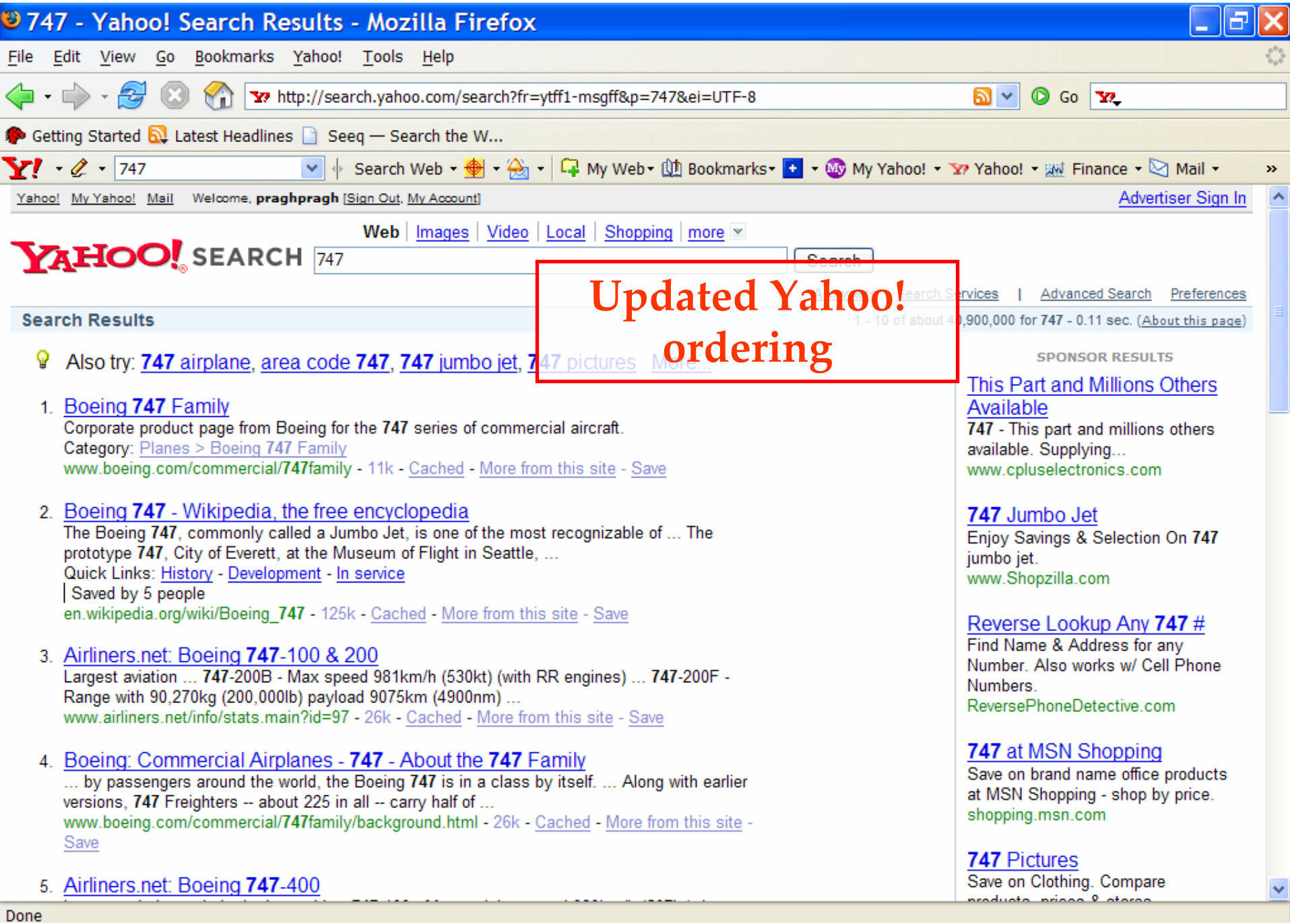

747 - Yahoo! Search Results - Mozilla Firefox
File Edit View Go Bookmarks Yahoo! Tools Help
http://search.yahoo.com/search?fr=ytff1-&p=747&ei=UTF-8
Go
Getting Started Latest Headlines Seeq — Search the ...
747
Search Web My Web Bookmarks My Yahoo! Yahoo! Finance Mail News
747 - Yahoo! Search Results
academia.pdf (application/pdf Object)
Yahoo! My Yahoo! Mail Welcome, praghpragh [Sign Out, My Account]
Search Home Help
Web Images Video Audio Directory Local News Shopping More »
YAHOO! SEARCH
747
Search the Web
My Web (beta)
My Web Answers BETA
Search Services Advanced Search Preferences
Search Results
Results 1 - 10 of about 27,100,000 for 747 - 0.23 sec (About this page)
Also try: boeing 747, 747 airplane, 747 se
My Web: 23 Results for 747
- Om Malik on Broadband - Saved b
- BBC NEWS | Business | Google s
- SeatGuru.com - Your Enlightened
- See all 23 results . . .
Bid Ordering (Overture/former Yahoo!)
1. Boeing 747-400 Family
Corporate product page from Boeing.
Category: Planes > Boeing 747 Family
www.boeing.com/commercial/747family - 11k - Cached - More from this site - Save
2. 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 - 102k - Cached - More from this site - Save
3. 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 - 16k - Cached - More from this site - Save
SPONSOR RESULTS
747 on eBay
Shop eBay for anything and everything - from specialized gifts to...
Shop Boeing 747-400 Miniature Models
Shop Boeing 747-400 miniature models at FlightMiniatures.com. See...
www.flightminiatures.com
Canon Fax L777 Toner-free Shipping
Save on compatible Canon Fax L777 toner. Purchase our...
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?)


747 - Google Search - Mozilla Firefox
File Edit View Go Bookmarks Yahoo! Tools Help
http://www.google.com/search?hl=en&q=747&btnG=Google+Search
Go
Getting Started
Latest Headlines
Seeq — Search the ...
747
Search Web
Mail
My Yahoo!
Answers
Games
Shopping
Music
Personals
Sign In
pragh60@gmail.com | My Account | Sign out
Google
Web Images Groups News Froogle Maps more »
747
Search
Advanced Search
Preferences
Revenue ordering?
Web
Results 1 - 10 of about 83,900,000 for 747. (0.10 seconds)
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/ - 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 - 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 - 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 - 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/ - 21k - May 3, 2006 - Cached - Similar pages
Sponsored Links
747 On Sale
Chairs, Desks & Cabinets
Everyday Low Prices.
OfficeFurniture.BettyMills.com
Boeing™ 747-400 Model
Accurately detailed Boeing™ 747-400 desktop display models
www.FlightMiniatures.com
Boeing's Photo Collection
The world's leading source of still and motion aerospace imagery!
www.boeingphotostore.com
747
Everything to do with 747 items.
eBay.com


747 - Yahoo! Search Results - Mozilla Firefox
File Edit View Go Bookmarks Yahoo! Tools Help
http://search.yahoo.com/search?fr=ytff1-msgff&p=747&ei=UTF-8
Go
Getting Started Latest Headlines Seeq — Search the W...
747
Search Web
My Web
Bookmarks
My Yahoo!
Yahoo!
Finance
Mail
Yahoo! My Yahoo! Mail Welcome, praghpragh [Sign Out, My Account]
Advertiser Sign In
Web Images Video Local Shopping more
YAHOO! SEARCH
747
Search
Search Services | Advanced Search Preferences
Search Results
1 - 10 of about 40,900,000 for 747 - 0.11 sec. (About this page)
Updated Yahoo! ordering
Also try: 747 airplane, area code 747, 747 jumbo jet, 747 pictures More...
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
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
Done

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

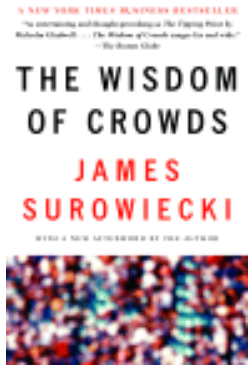
- How do you design marketplaces for search advertising?
  - What tools do you give advertisers?
  - What do they get to bid on?
    - ***Laptop*** vs ***Ultralight laptop***
  - How do you price clicks?
- Reputation and incentive mechanisms
- Matching marketplaces
  - Jobs, dates, …
  - Online matching 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 Lily, 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

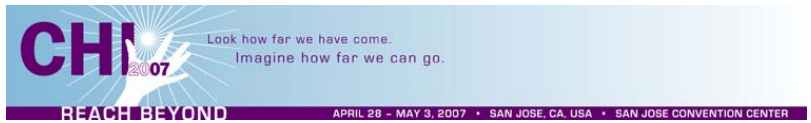
Media Experience

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

short — long

Memorability

forgotten — remembered

Social context

alone — with others

•
•
•
•

Second life
Future Galaxy
WOW

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

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

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

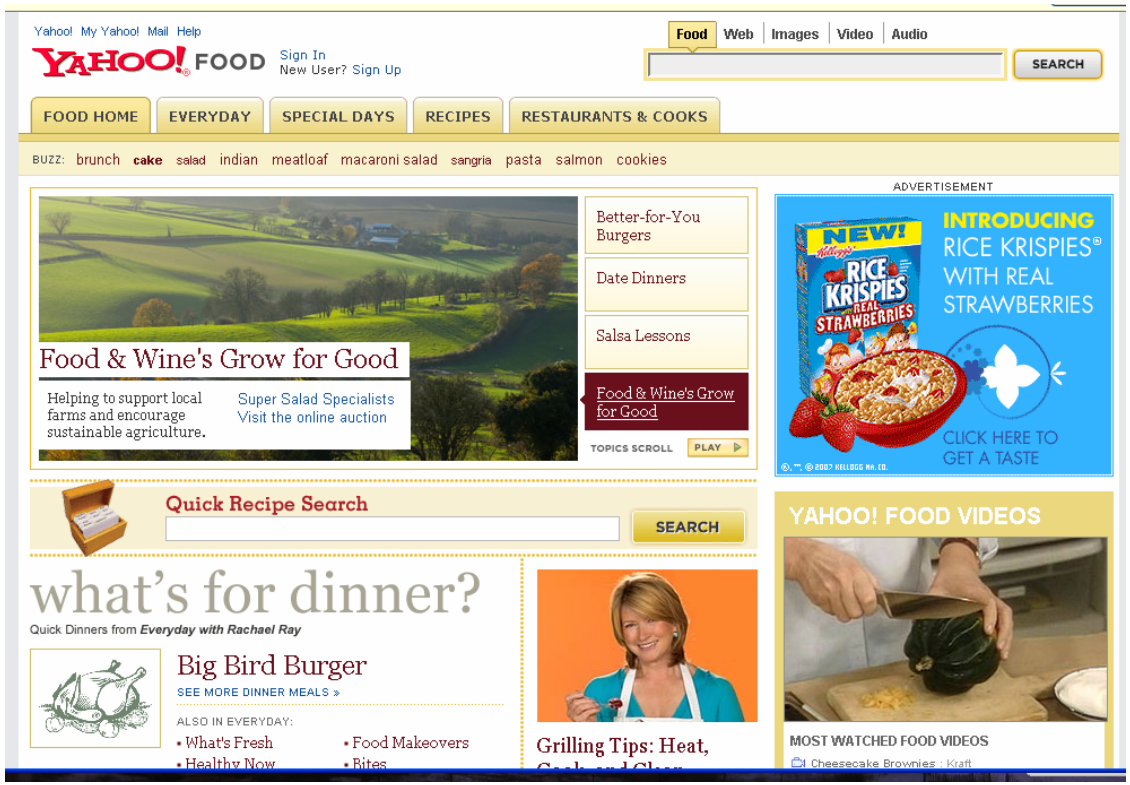
- Clicks, click-throughs and click chains
- Page views, repeat visits
- Time elapsed within page and between pages
- Creation of user IDs
- Creation of user profiles
- Content generation
- Content transport through networks
- Downloads, purchases
- Subscriptions

Challenges:

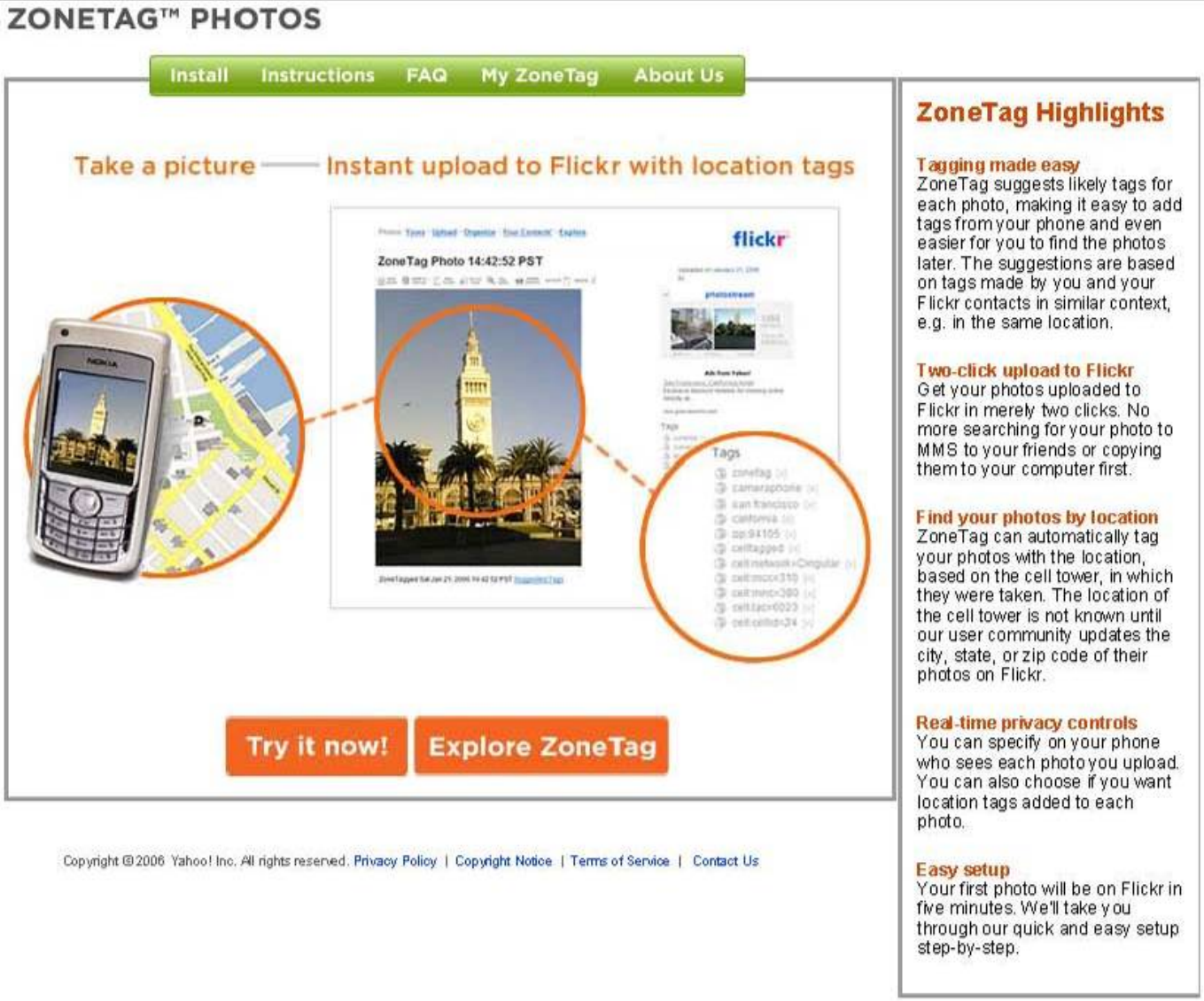
- Devise and standardize defensible metrics of online engagement
- Use these to predictively devise online experiences

Structure Matters in Media

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



ZoneTag


ZONETAG™ PHOTOS
Install
Instructions
FAQ
My ZoneTag
About Us
Take a picture —— Instant upload to Flickr with location tags
flickr
ZoneTag Photo 14:42:52 PST
Tags
Try it now!
Explore ZoneTag
Copyright © 2006 Yahoo! Inc. All rights reserved. Privacy Policy | Copyright Notice | Terms of Service | Contact Us
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 | REMIX HOME
REMIXER TOOL | 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:

A short history of ME.
shamma

MORE NEW REMIXES:


david82


pippo


peter

INTERNATIONAL REMIX HOME | GALLERY | INFO | REMIXER | FEEDBACK

powered by Yahoo! Research Berkeley

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!