It's Y!ou—Content Optimization at Yahoo!

Raghu Ramakrishnan

Chief Scientist, Search & Cloud Computing

Yahoo! Fellow

The Message

- The Web will be increasingly personalized, but “personalization” will mostly happen within the context of
  - Content optimization
  - Semantic interpretation of web content and user intent
  - Socialization of the web

Content Optimization

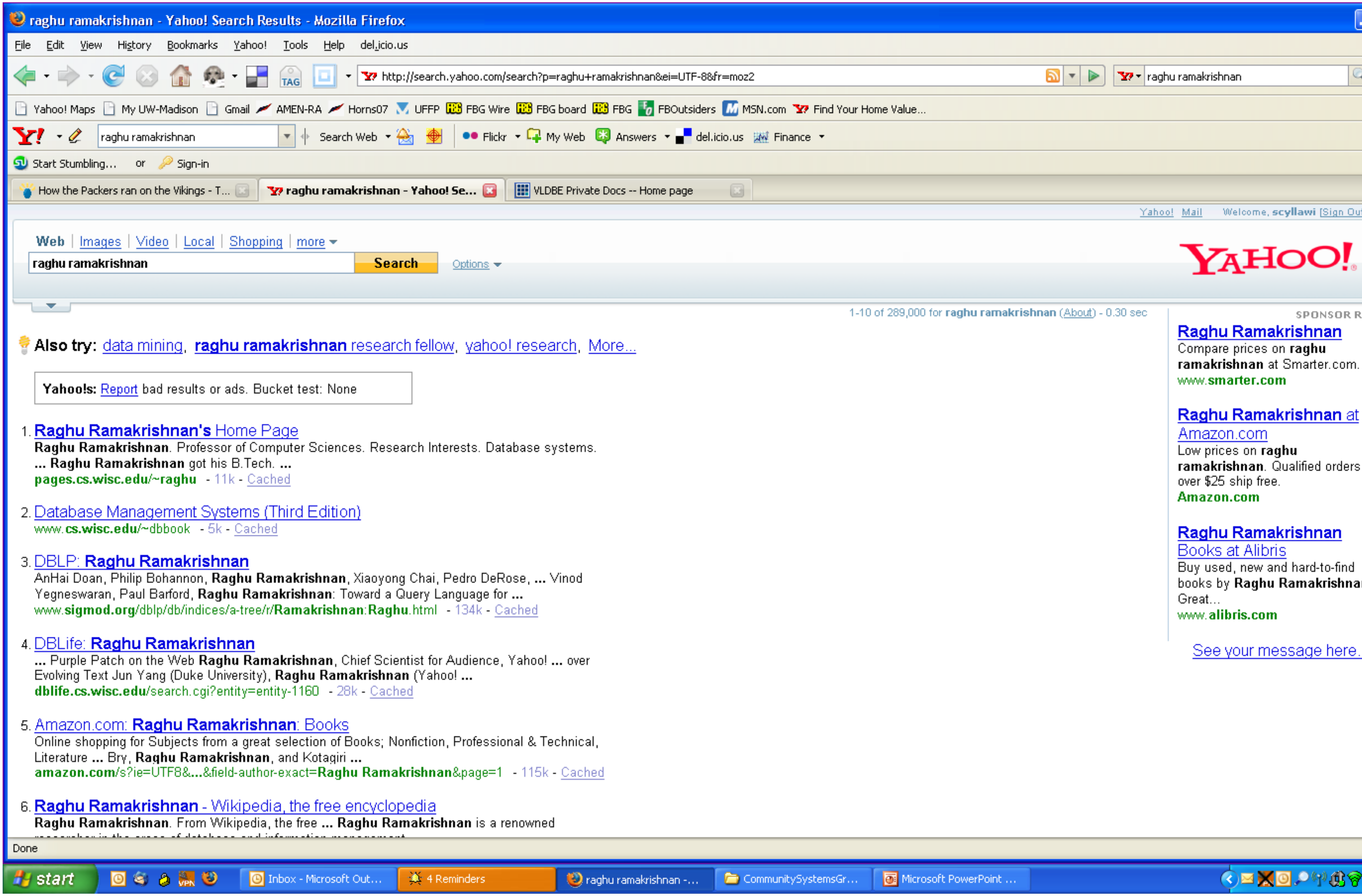
- Content
  - Portal content, Search, Social
- Optimization objective
  - Engagement, revenue, retention, `voice’
- Signals:
  - Content features
    - Topic/entity
  - Popularity
    - Links, referrals, clicks, CTRs
  - User features
    - Content consumption, social, explicit

Personalization

- Why?
  - Engagement, revenue, retention, `voice’
- Why not?
  - Discovery, privacy, search bias
- Who?
  - User profiles
- How?
  - Algorithms, infrastructure

User Interest Modeling

- User—Topic (entities, categories) affinity
  - From logs (clicks, views, purchases, etc.)
- Latent Factor models
  - Topic affinities based on activity of user and "similar" users
- Search history
  - Online and offline
- Responses to recommendations
  - Current session history
- Explicit user declaration
  - When and what should we ask users?


raghu ramakrishnan - Yahoo! Search Results - Mozilla Firefox
File Edit View History Bookmarks Yahoo! Tools Help del.icio.us
http://search.yahoo.com/search?p=raghu+ramakrishnan&ei=UTF-8&fr=moz2
Yahoo! Maps My UW-Madison Gmail AMEN-RA Horns07 UFFP FBG Wire FBG board FBG FBOutsiders MSN.com Find Your Home Value...
How the Packers ran on the Vikings - T...
raghu ramakrishnan - Yahoo! Se...
VLDBE Private Docs -- Home page
Yahoo! Mail Welcome, scyllawi [Sign Out]
Web | Images | Video | Local | Shopping | more
raghu ramakrishnan
Search
Options
YAHOO!
1-10 of 289,000 for raghu ramakrishnan (About) - 0.30 sec
Also try: data mining, raghu ramakrishnan research fellow, yahoo! research, More...
Yahoo!s: Report bad results or ads. Bucket test: None
1. Raghu Ramakrishnan's Home Page
Raghu Ramakrishnan. Professor of Computer Sciences. Research Interests. Database systems. ... Raghu Ramakrishnan got his B.Tech. ...
pages.cs.wisc.edu/~raghu - 11k - Cached
2. Database Management Systems (Third Edition)
www.cs.wisc.edu/~dbbook - 5k - Cached
3. DBLP: Raghu Ramakrishnan
AnHai Doan, Philip Bohannon, Raghu Ramakrishnan, Xiaoyong Chai, Pedro DeRose, ... Vinod Yegneswaran, Paul Barford, Raghu Ramakrishnan: Toward a Query Language for ...
www.sigmod.org/dblp/db/indices/a-tree/r/Ramakrishnan:Raghu.html - 134k - Cached
4. DBLife: Raghu Ramakrishnan
... Purple Patch on the Web Raghu Ramakrishnan, Chief Scientist for Audience, Yahoo! ... over Evolving Text Jun Yang (Duke University), Raghu Ramakrishnan (Yahoo! ...
dblife.cs.wisc.edu/search.cgi?entity=entity-1160 - 28k - Cached
5. Amazon.com: Raghu Ramakrishnan: Books
Online shopping for Subjects from a great selection of Books; Nonfiction, Professional & Technical, Literature ... Bry, Raghu Ramakrishnan, and Kotagiri ...
amazon.com/s?ie=UTF8&...&field-author-exact=Raghu Ramakrishnan&page=1 - 115k - Cached
6. Raghu Ramakrishnan - Wikipedia, the free encyclopedia
Raghu Ramakrishnan. From Wikipedia, the free ... Raghu Ramakrishnan is a renowned
SPONSOR RE
Raghu Ramakrishnan
Compare prices on raghu ramakrishnan at Smarter.com.
www.smarter.com
Raghu Ramakrishnan at Amazon.com
Low prices on raghu ramakrishnan. Qualified orders over $25 ship free.
Amazon.com
Raghu Ramakrishnan Books at Alibris
Buy used, new and hard-to-find books by Raghu Ramakrishnan. Great...
www.alibris.com
See your message here...
Done
start
Inbox - Microsoft Out...
4 Reminders
raghu ramakrishnan -...
CommunitySystemsGr...
Microsoft PowerPoint ...

- "Just look at our homepage, for example. Since we began pairing our **content optimization technology** with editorial expertise, we've seen **click-through rates in the Today module more than double**. And we're making additional improvements to this technology that will make the user experience ever more personally relevant."
- **Carol Bartz, Analyst Call, January 27, 2010**

CONTENT OPTIMIZATION FOR PORTALS

Team from Y! Research

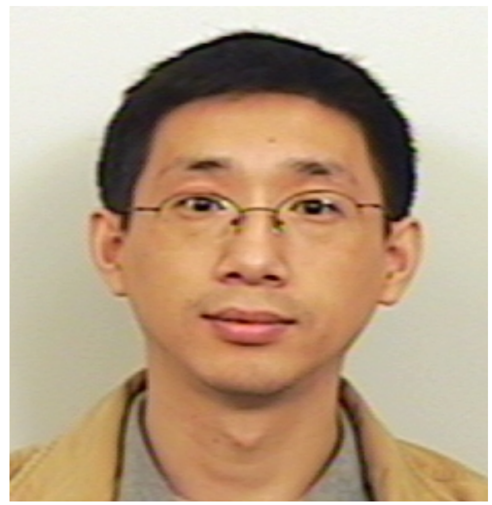
Deepak Agarwal

BeeChung Chen

Pradheep Elango

Seung-Taek Park

Raghu Ramakrishnan



Wei Chu

Team from Y! Engineering

Nitin Motgi

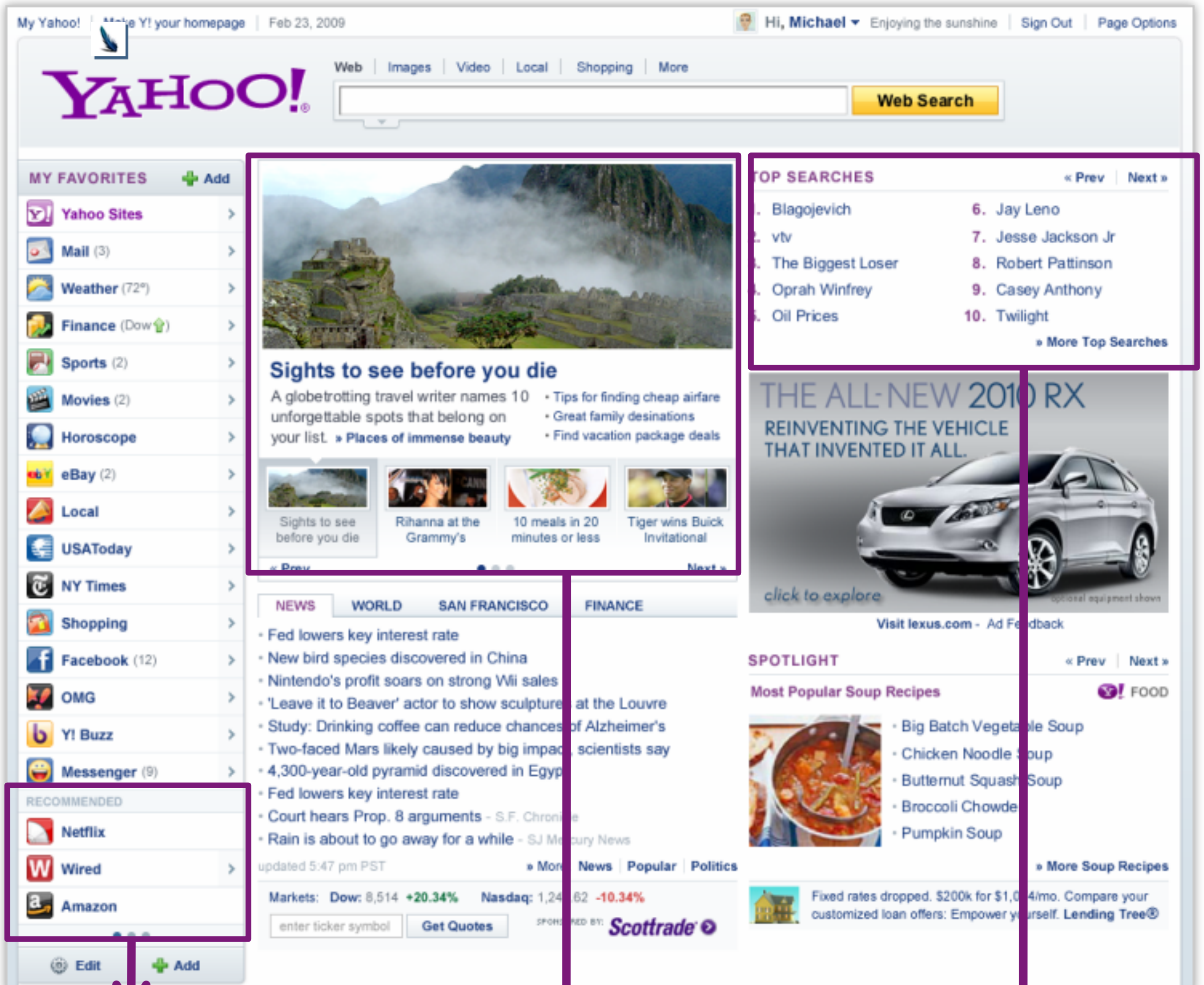
Joe Zachariah

Scott Roy

Kenneth Fox

Todd Beaupre

Content Optimization



Recommended links

+79% clicks
vs. randomly selected

News Interests

+200% clicks
vs. one size fits all

Top Searches

+43% clicks
vs. editor selected

Key Features

Package Ranker (CORE)

Ranks packages by expected CTR based on data collected every 5 minutes

Dashboard (CORE)

Provides real-time insights into performance by package, segment, and property

Mix Management (Property)

Ensures editorial voice is maintained and user gets a variety of content

Package rotation (Property)

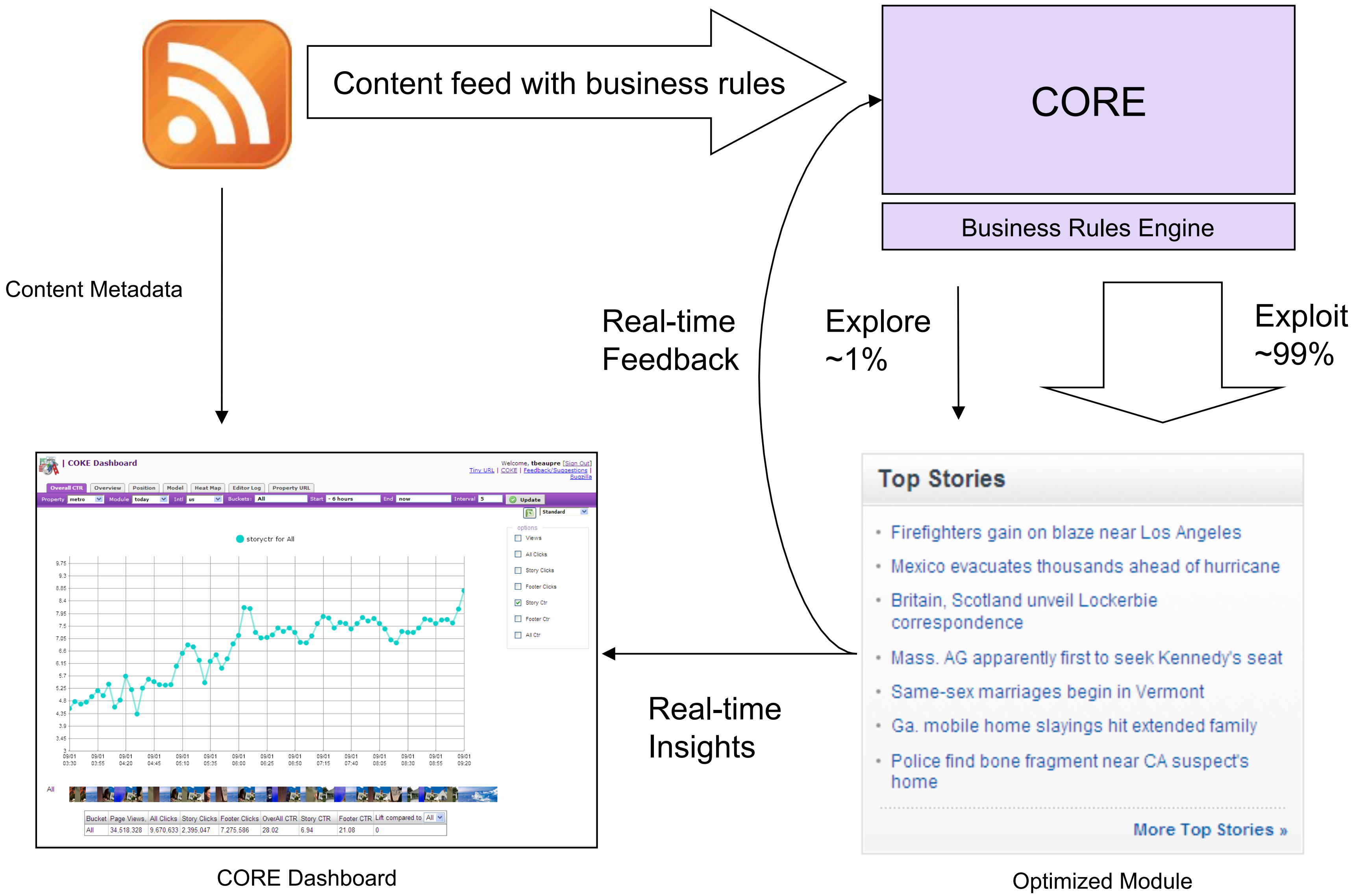
Tracks which stories a user has seen and rotates them after user has seen them for a certain period of time

Key Performance Indicators

Lifts in quantitative metrics

Editorial Voice Preserved

CORE Data Flow



Recommender Approaches

Estimate Most Popular (EMP) — *"What's most engaging overall?"*

Behavioral Affinities — *"People who did X, did Y"*

Italian
RED
94087
94089
RED
Italian

Attribute Similarities — *"Related items with similar metadata"*

Social Recommendations — *"What are my trusted connections into?"*

Business Optimization — *"What generates most business value?"*

Personalized Recommendations — *"What's most relevant to <u>me</u> based on my interests, attributes and relationships?"*

CORE Modeling Overview

Offline Modeling
- Exploratory data analysis
- Regression, feature selection, collaborative filtering (factorization)

- Seed online models & explore/exploit methods at good initial points
- Reduce the set of candidate items

Large amount of historical data (user event streams)

Hadoop

Online Learning
- Online regression models, time-series models

- Model the temporal dynamics
- Provide fast learning for per-item models

Hadoop pipelines

PNUTS

Near real-time user feedback

Explore/Exploit
- Multi-armed bandits

- Find the best way of collecting real-time user feedback (for new items)

Ranking in CORE

- Pure feature based (did not work well):
  - Article: URL, keywords, categories
  - Build offline models to predict CTR when article shown to users
  - Models considered
    - Logistic Regression with feature selection
    - Decision Trees, Feature segments through clustering
- Track CTR per article in user segments through online models
  - **This worked well; the approach we took eventually**

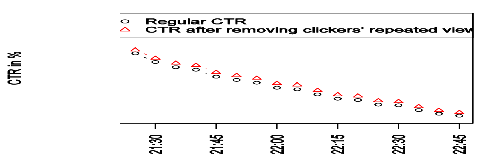
Explore/Exploit

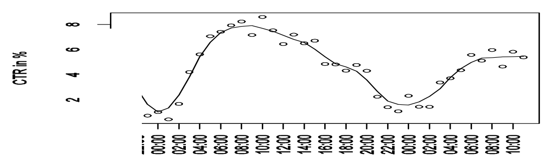
- What is the best strategy for new articles?
  - If we show it and it's bad: lose clicks
  - If we delay and it's good: lose clicks
- Solution: Show it while we don't have much data if it looks promising
  - Classical **multi-armed bandit** type problem
  - Our setup is different than the ones studied in the literature; new ML problem

(Online Models for Content Optimization, NIPS 2008: D. Agarwal, B. Chen, P. Elango, N. Motgi, S. Park, R. Ramakrishnan, S. Roy, J. Zachariah)

Challenges in Our Setting

- Dynamic item pools
- Article CTRs decay over time
- Time-of-day /day-of-week effects



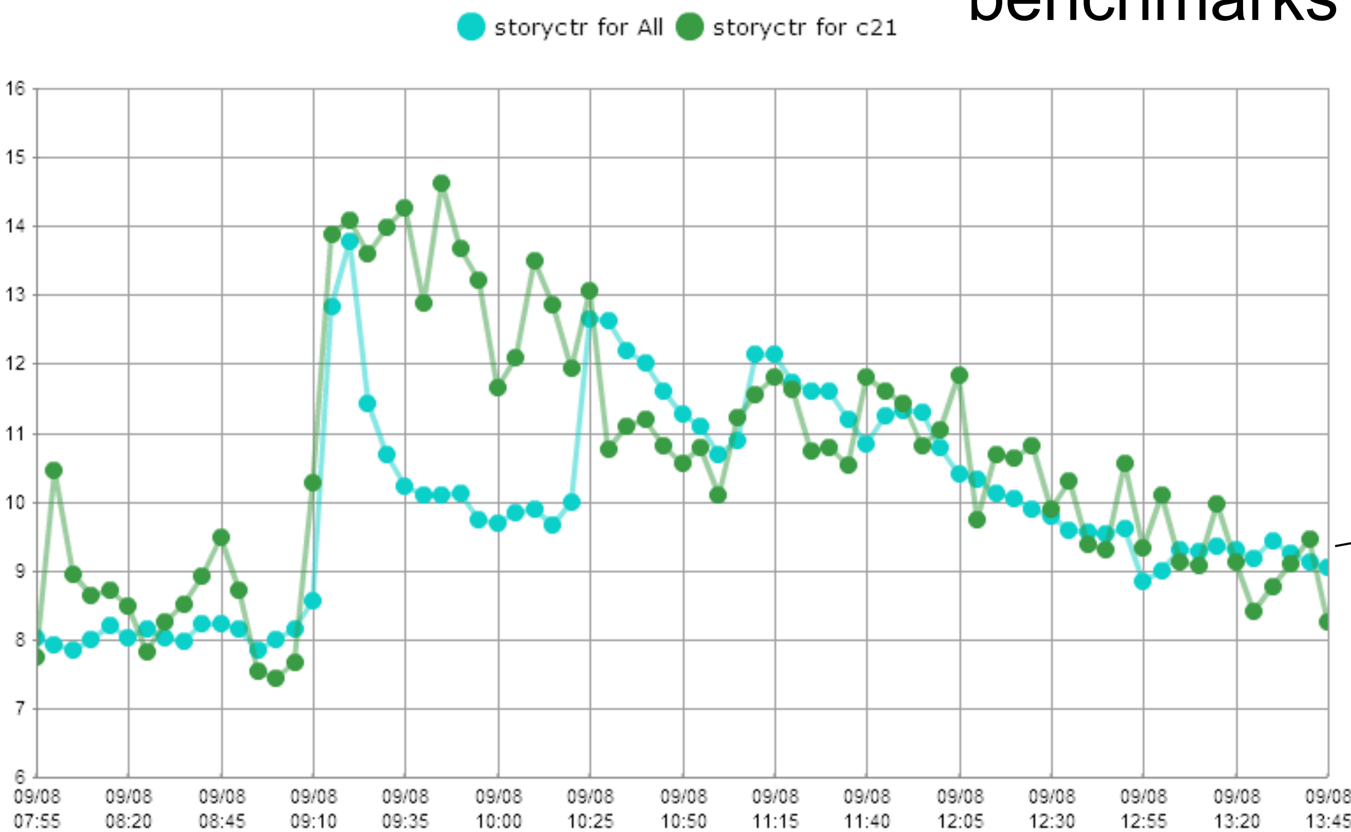


Some Other Complications

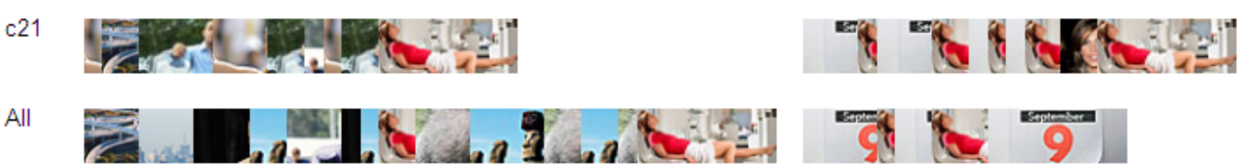
- We run multiple experiments (possibly correlated) simultaneously; effective sample size calculation a challenge
- Serving Bias: Incorrect to learn from data for serving scheme A and apply to serving scheme B
  - Need unbiased quality score
  - Bias sources: positional effects, time effect, set of articles shown together
- Incorporating feature-based techniques
  - e.g., logistic regression; tree-based (hierarchical bandit)

CORE Dashboard: Overall CTR

Compare performance of models and historical benchmarks



Compare buckets and models over time

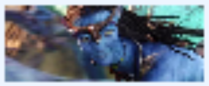
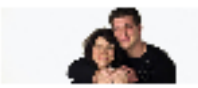
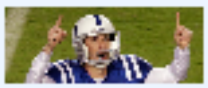
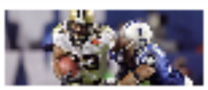
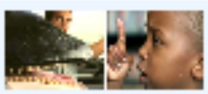
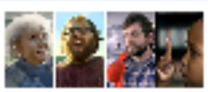
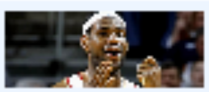
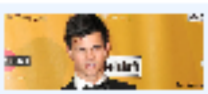


See which content was promoted most across time

Bucket	Page Views,	All Clicks	Story Clicks	Footer Clicks	OverAll CTR	Story CTR	Footer CTR	Lift compared to All
All	43,005,783	13,061,688	4,286,407	8,775,281	30.37	9.97	20.4	0
c21	281,821	84,853	29,777	55,076	30.11	10.57	19.54	6.02

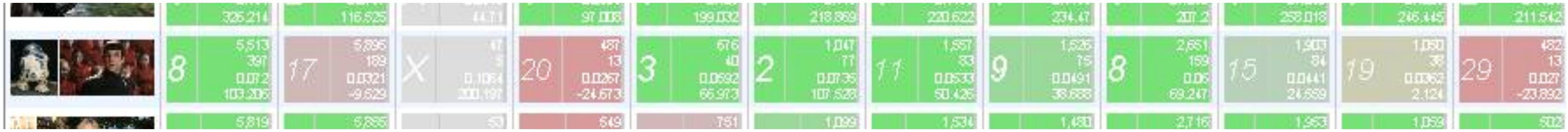
Compare bucket metrics

CORE Dashboard:
Segment Heat Map

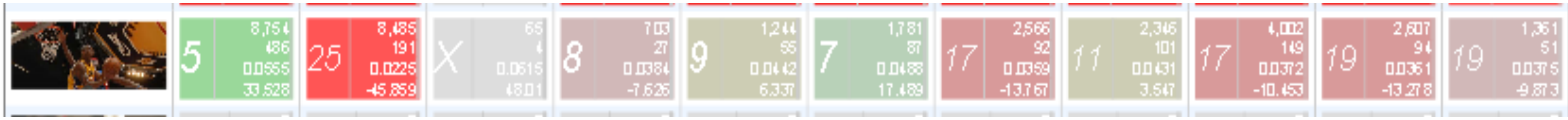
Package	male	female	OMG	BUAuto	BUEnt	BU Fin	Health	BUSport▲	NBA	BUTrav	ALL
	408,260 18,440 0.0452 8.477	390,404 14,449 0.037 -11.113	270,039 16,940 0.0627 50.661	121,080 7,389 0.061 46.564	270,038 16,940 0.0627 50.661	325,873 20,012 0.0614 47.488	195,796 12,763 0.0652 56.553	350,152 21,454 0.0613 47.152	132,916 9,457 0.0712 70.879	123,388 7,896 0.064 53.691	923,611 38,457 0.0416 0
	1 8,067 852 0.1056 153.654	1 7,657 674 0.088 111.405	1 5,125 720 0.1405 237.406	1 2,382 286 0.1201 188.362	1 5,125 720 0.1405 237.406	1 6,415 858 0.1337 221.221	1 3,769 532 0.1412 239	1 6,750 917 0.1359 226.272	1 2,585 385 0.1489 257.696	1 2,490 330 0.1325 218.294	1 18,137 1,738 0.0958 130.143
	5 9,968 644 0.0646 55.164	3 12,847 777 0.0605 45.256	2 8,569 885 0.1033 148.043	4 3,529 326 0.0924 121.86	2 8,569 885 0.1033 148.043	3 9,744 922 0.0946 127.252	3 6,067 643 0.106 154.537	2 10,187 1,004 0.0986 136.702	5 3,820 420 0.1099 164.058	2 4,037 433 0.1073 157.598	4 25,744 1,595 0.062 48.798
	2 3,326 249 0.0749 79.8	5 3,954 212 0.0536 28.769	5 2,521 231 0.0916 120.066	2 1,004 102 0.1016 143.995	5 2,521 231 0.0916 120.066	5 3,016 276 0.0915 119.782	5 1,860 186 0.1 140.167	3 3,291 310 0.0942 126.229	3 1,141 136 0.1192 186.264	3 1,039 100 0.0962 131.152	3 8,500 541 0.0636 52.859
	11 2,562 133 0.0519 24.677	13 2,004 81 0.0404 -2.926	3 1,250 122 0.0976 134.403	6 629 51 0.0811 94.73	3 1,250 122 0.0976 134.403	4 1,608 151 0.0939 125.53	2 919 103 0.1121 169.175	4 1,669 154 0.0923 121.604	4 655 74 0.113 171.334	4 591 55 0.0931 123.506	10 5,342 252 0.0472 13.295
	3 2,881 206 0.0715 71.727	2 3,242 230 0.0709 70.384	4 2,071 196 0.0946 127.295	3 949 95 0.1001 140.42	4 2,071 196 0.0946 127.295	2 2,614 254 0.0972 133.368	4 1,605 165 0.1028 146.901	5 2,740 239 0.0872 109.489	10 1,036 94 0.0907 117.912	9 958 78 0.0814 95.543	2 7,043 493 0.07 68.114
	6 10,785 649 0.0602 44.523	4 12,768 742 0.0581 39.571	7 8,580 694 0.0809 94.261	7 3,511 283 0.0806 93.584	7 8,580 694 0.0809 94.261	6 9,725 795 0.0817 96.332	6 6,138 550 0.0896 115.204	6 10,670 866 0.0812 94.925	11 3,669 321 0.0875 110.122	5 3,785 339 0.0896 115.104	5 27,331 1,641 0.06 44.2
	10 22,202 1,212 0.0546 31.106	7 23,328 1,200 0.0514 23.543	6 15,593 1,289 0.0827 98.535	5 6,552 533 0.0813 95.374	6 15,593 1,289 0.0827 98.535	7 17,652 1,376 0.078 87.214	8 10,797 915 0.0847 103.532	7 19,050 1,522 0.0799 91.882	9 6,639 604 0.091 118.498	7 6,435 552 0.0858 106.018	6 52,978 2,786 0.0526 26.299
	22 26,685 1,160 0.0435 4.401	10 35,405 1,530 0.0432 3.786	8 19,832 1,572 0.0793 90.371	9 7,844 552 0.0704 69.011	8 19,832 1,572 0.0793 90.371	8 21,743 1,641 0.0755 81.26	7 13,721 1,167 0.0851 104.267	8 22,168 1,743 0.0786 88.836	8 8,249 788 0.0955 129.424	8 8,327 689 0.0827 98.721	18 74,559 3,167 0.0425 2.014
	4 7,745 518 0.0669 60.628	26 7,202 185 0.0257 -38.308	13 4,898 322 0.0657 57.889	15 2,308 148 0.0641 54.007	13 4,898 322 0.0657 57.889	11 6,051 423 0.0699 67.891	19 3,652 235 0.0643 54.544	9 6,436 506 0.0786 88.82	2 2,562 308 0.1202 188.726	12 2,359 169 0.0716 72.057	7 17,235 834 0.0484 16.217
	7 7,699 460 0.0597 43.495	29 7,201 169 0.0235 -43.635	11 4,809 340 0.0707 69.8	10 2,269 158 0.0696 67.239	11 4,809 340 0.0707 69.8	9 6,004 433 0.0721 73.205	14 3,544 243 0.0686 64.674	10 6,247 475 0.076 82.615	6 2,482 257 0.1035 148.682	11 2,329 167 0.0717 72.211	12 17,169 783 0.0456 9.529
	12 7,688 393 0.0511 22.77	9 7,229 336 0.0465 11.628	9 4,785 363 0.0759 82.196	17 2,280 139 0.061 46.418	9 4,785 363 0.0759 82.196	12 6,037 403 0.0668 60.324	13 3,501 245 0.07 68.069	11 6,319 430 0.068 63.431	15 2,397 182 0.0759 82.355	15 2,312 152 0.0657 57.895	8 17,275 833 0.0482 15.809

Examples

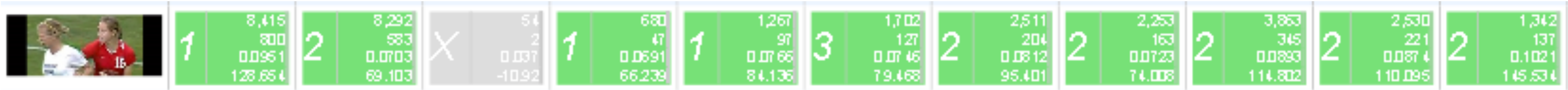
– **ACQUISITION:** A “Star Trek” package was #3 with 18-20 demo, #2 with 21-24 demo, but #9 overall. We can acquire younger audiences with targeted content like this.



– **ENGAGEMENT:** “Kobe’s astonishing shot” was #25 with women, but #5 with men. We can better engage men (or sports fans) by showing more like this, women by showing less.



– **REACH:** A package about a hair-pulling soccer player was just plain interesting to everyone (#1-3). We can maintain reach by programming content for the mass audience.



Degrees of Personalization

Most Popular
Most engaging overall based on objective metrics

Most Popular + Per User History
Engaging overall, and aware of what I've already seen

Light Personalization
More relevant to me based on my age, gender and property usage

Deep Personalization
Most relevant to me based on my deep interests and relationships

---

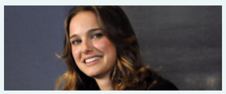
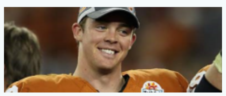
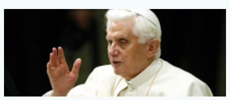
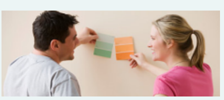
Voice and Business Rules

Real-time Dashboard

Business Optimization

Matching Users to Content

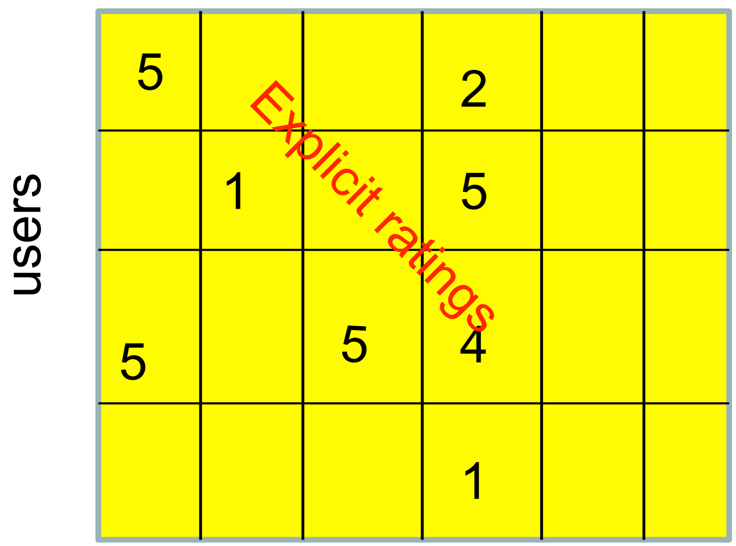
We learn how user attributes correlate with engagement for each item

	Default	Male	Female	18-24	25-34	Heavy Sports
	7.1	-0.4	+0.4	+0.3	+0.1	-0.5
	6.8	+1.0	-1.0	+0.2	+0.3	+2.1
	6.5	-0.6	+0.6	+0.5	+0.3	-0.8
	6.2	0	0	-0.7	-0.5	-0.3
	5.9	-1.1	+1.1	-0.5	-0.2	-0.2
	✓	✓		✓		✓

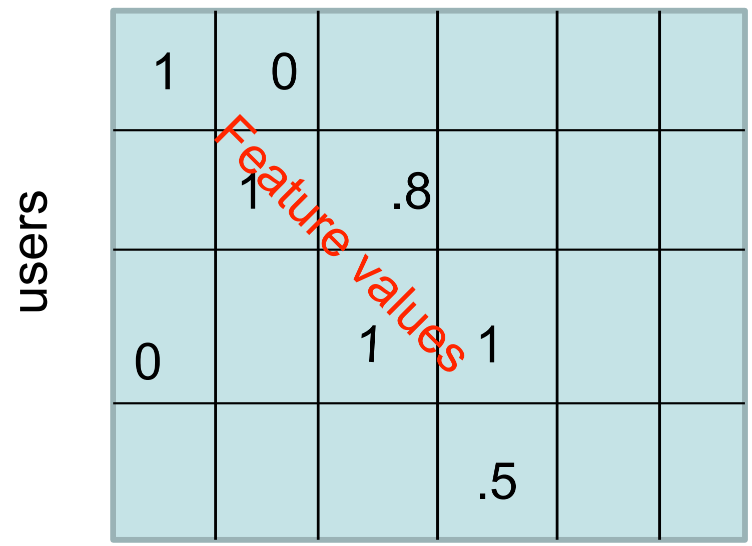
We compute rankings for each user based on his/her attributes

Collaborative Filtering

Movies



User features
e.g. ?sports, ?music,

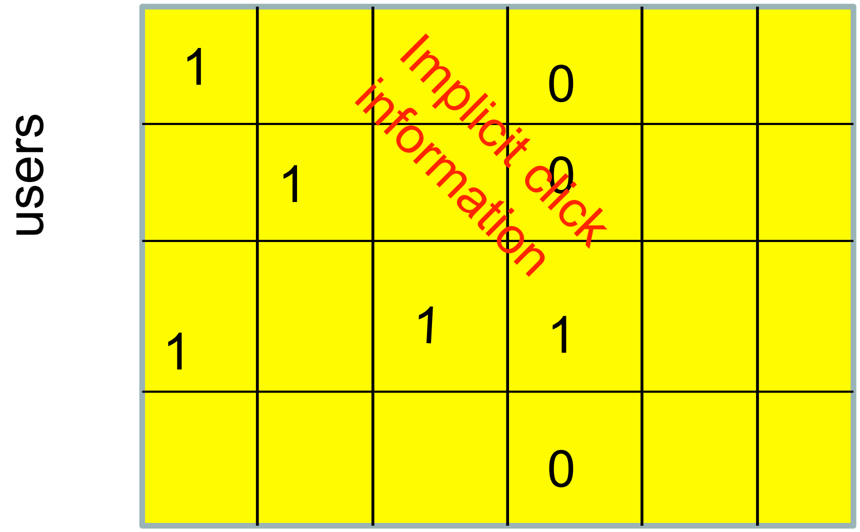


Q: Can we use additional feature information to better predict explicit ratings?

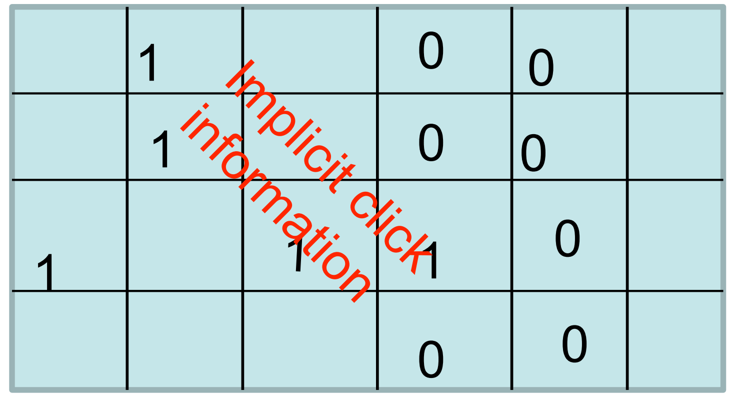
Feature information maybe incomplete, uncertain

Content Recommendation

Articles on Yahoo front pg



Articles on Yahoo news



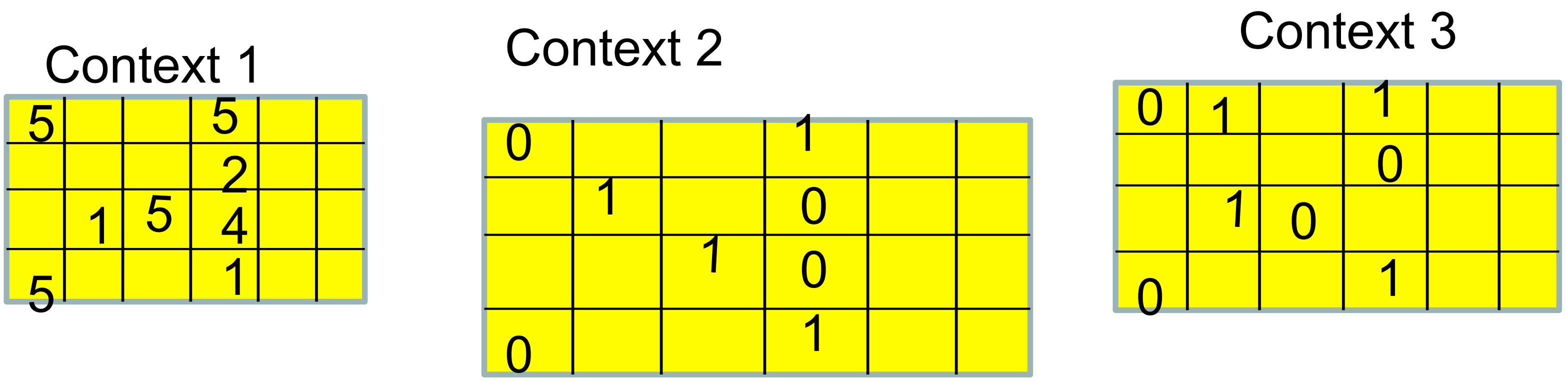
Users in two contexts are different but some are common

Articles in the two contexts maybe different with no overlap

Q: Can we make click prediction better in both contexts?

Problem Setup

- User x item matrices in *n* different contexts



- Items
  - Item-ids, feature values
- Response/ratings
  - Explicit/implicit ratings, feature values
- Matrices could be incomplete/complete

Localized Matrix Factorization (LMF)

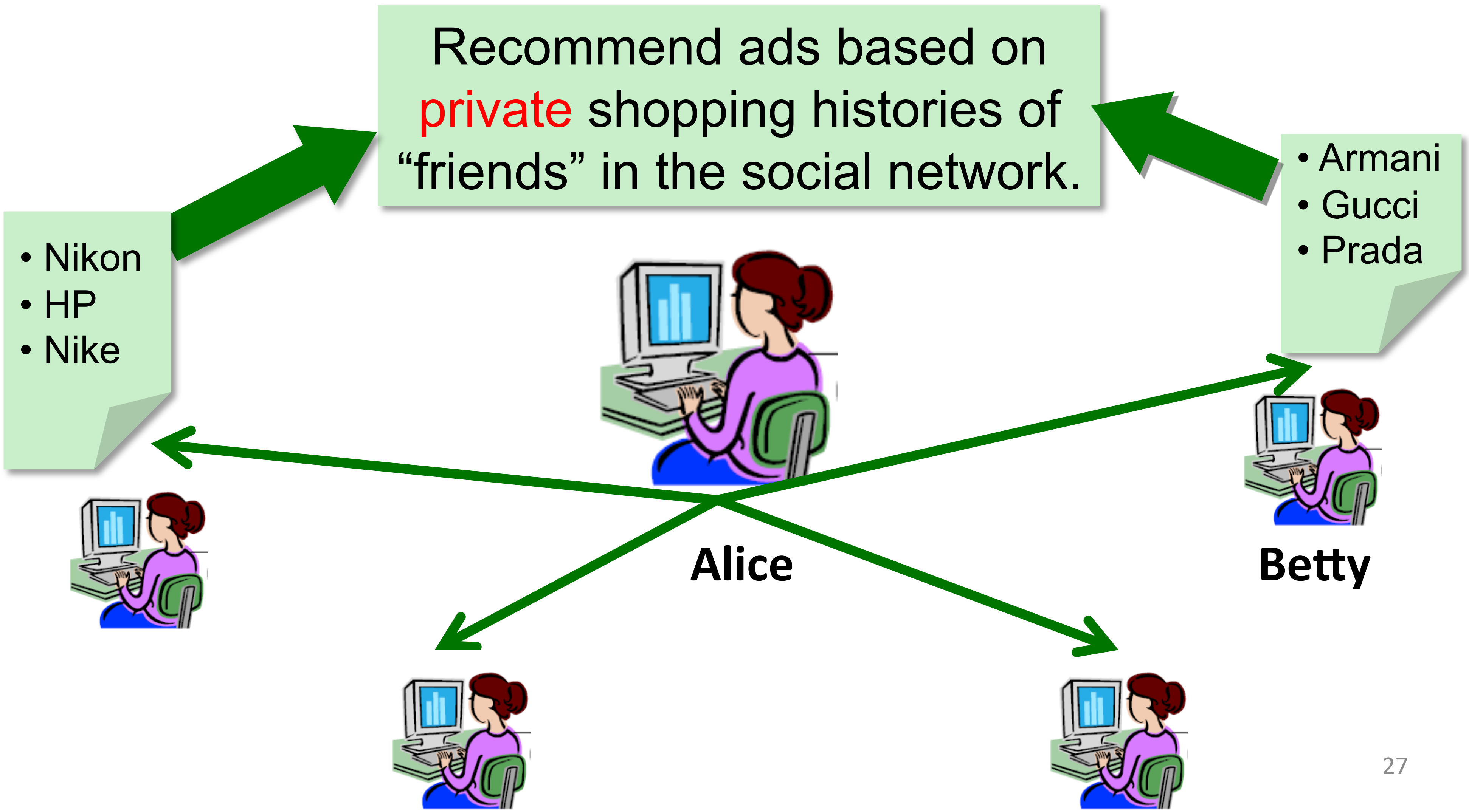
- Assume a joint distribution on user factors across contexts
- E.g. k = 2: $$\begin{pmatrix} \mathbf{z}_{i1} \\ \mathbf{z}_{i2} \end{pmatrix} \sim \text{iid} \quad \text{MVN}\left(\begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \begin{pmatrix} \sigma_{z,1}^2 I & \mathbf{\Sigma}_{12} \\ \mathbf{\Sigma}_{12}^{'} & \sigma_{z,2}^2 I \end{pmatrix} \right)$$

Factor covariance

- Estimate factor covariance from data and borrow information across contexts through conditionals
  - Amount of information borrowed depends on covariance
  - Measurement error taken care of through joint modeling

(Localized Factor Models: Agarwal et al., KDD 2011)

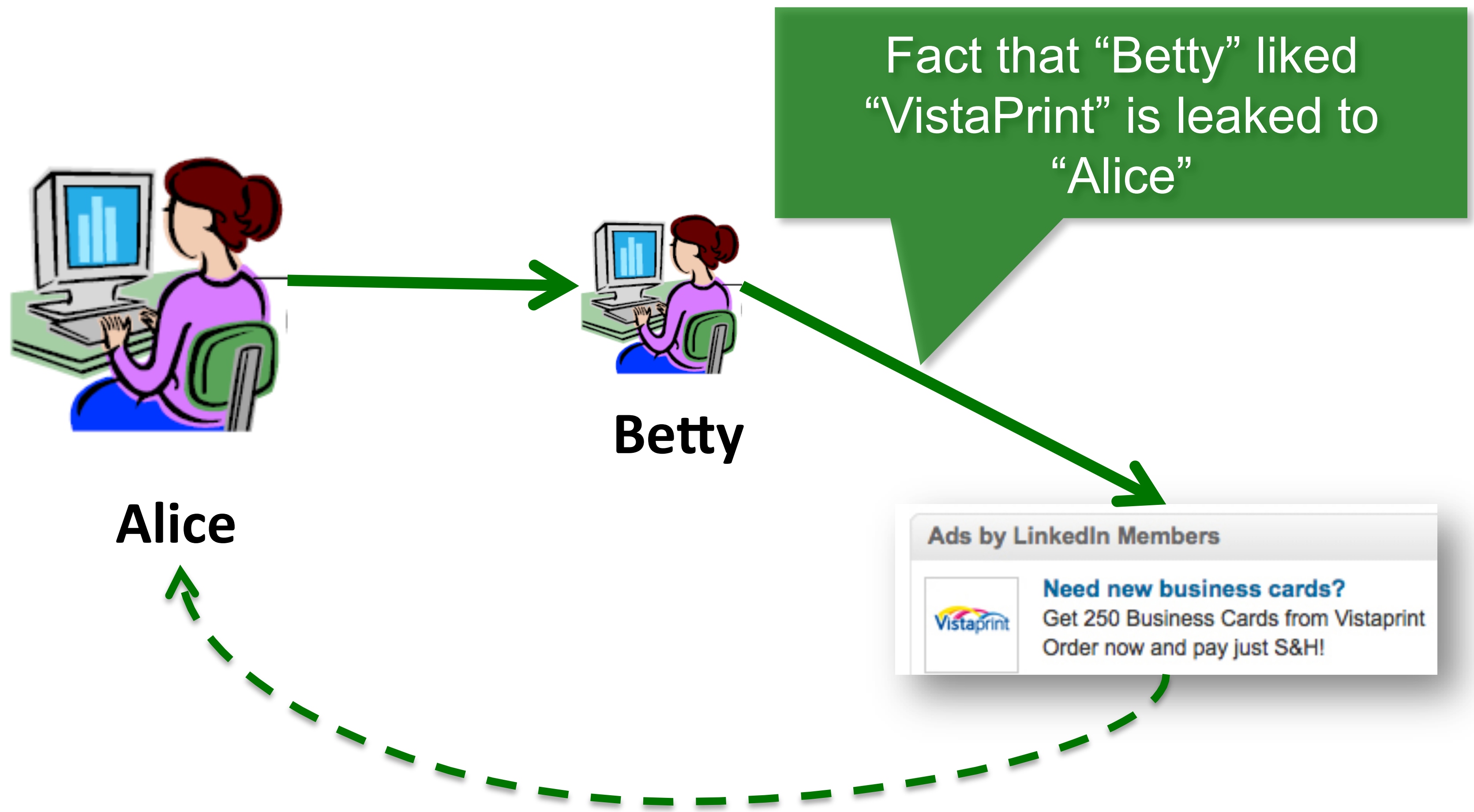
Social Advertising



(Personalized Social Recommendations: Machanavajjhala et al., VLDB 2011)

Privacy in Social Advertising

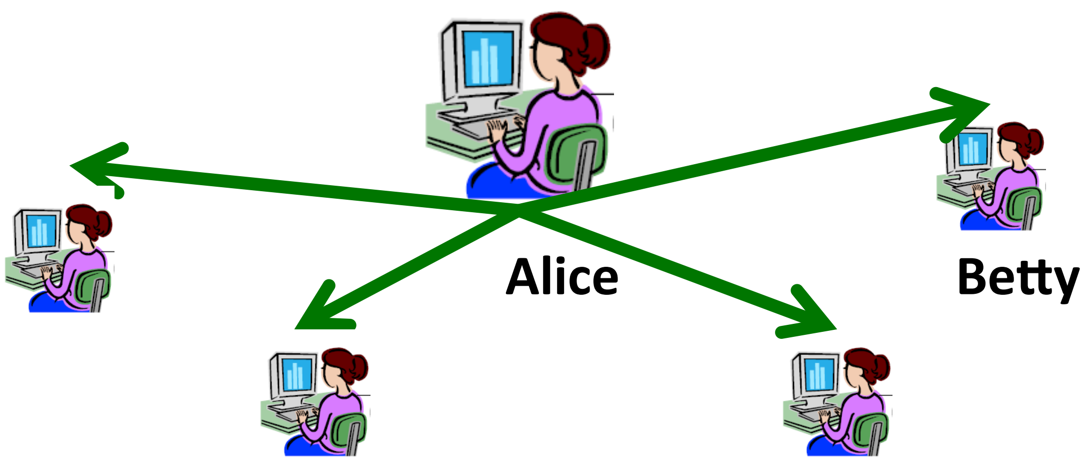
Items (products/people) liked by Alice's friends are better recommendations for Alice



28

Privacy in Social Advertising

Alice is recommended ‘X’



Can we provide *accurate* recommendations to Alice
while ensuring that
Alice *cannot* deduce that Betty likes ‘X’ ?

Takeaway …

- “For majority of the nodes in the network, recommendations must either be inaccurate or violate differential privacy!”
  - Maybe this is a “bad idea”
  - Or, Maybe **differential privacy is too strong a privacy definition to shoot for**.

30

System Challenges

- Highly dynamic system characteristics:
  - Short article lifetimes, pool constantly changing, user population is dynamic, CTRs non-stationary
  - Quick adaptation is key to success
- Scalability:
  - 1000’s of page views/sec; data collection, model training, article scoring done under tight latency constraints

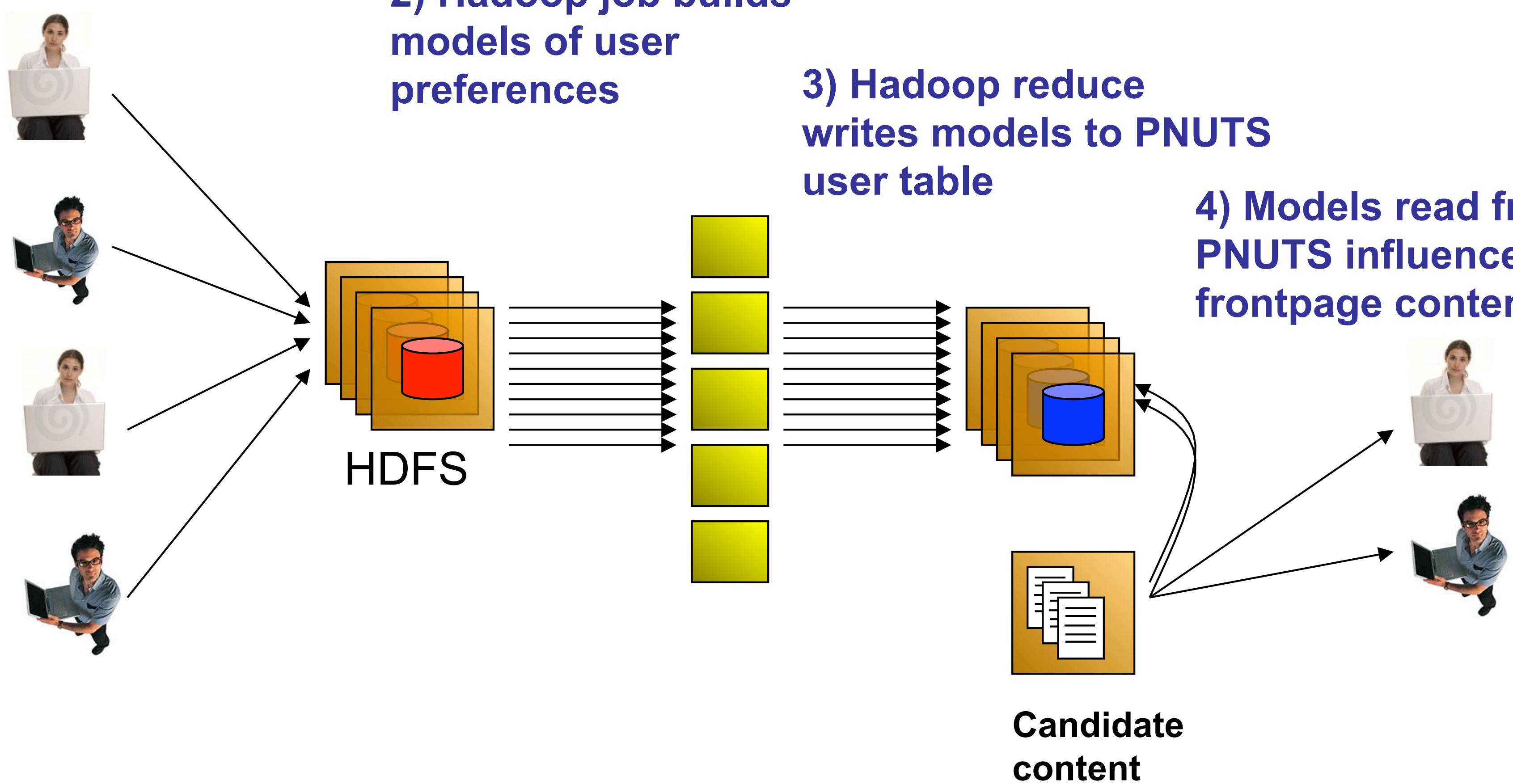
Data Management in CORE

1) User click history logs stored in HDFS

2) Hadoop job builds models of user preferences

3) Hadoop reduce writes models to PNUTS user table

4) Models read from PNUTS influence users' frontpage content



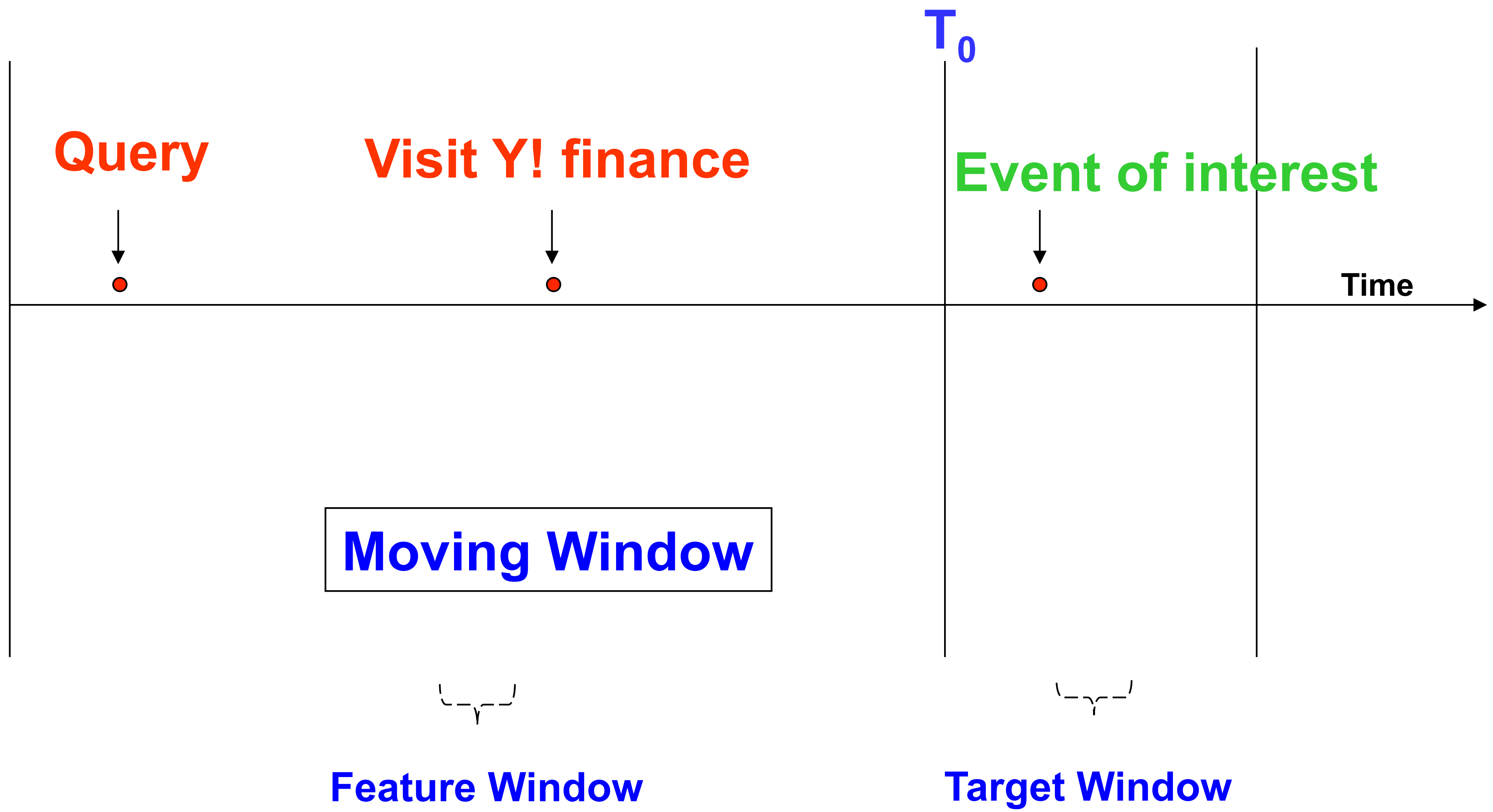
User Activity Modeling

- Large dimensionality vector describing possible user activities
- But a typical user has a sparse activity vector

Attribute	Possible Values	Typical values per user
Pages	~ MM	10 – 100
Queries	~ 100s of MM	Few
Ads	~ 100s of thousands	10s

33

Feature and Target Windows



User Modeling Pipeline

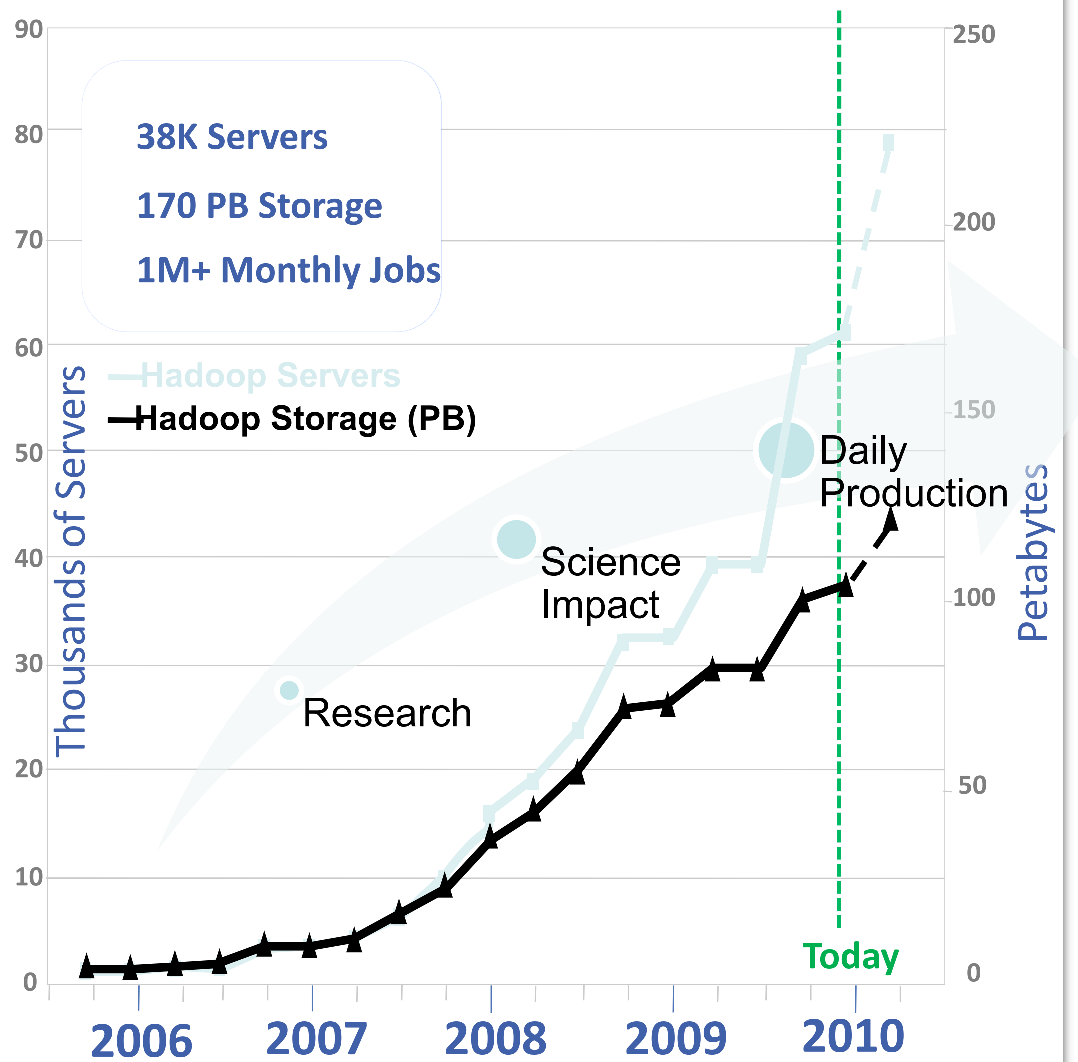
Component	Data Processed	Time
Data Acquisition	~ 1 Tb per time period	2 – 3 hours
Feature and Target Generation	~ 1 Tb * Size of feature window	4 - 6 hours
Model Training	~ 50 - 100 Gb	1 – 2 hours for 100’s of models
Scoring	~ 500 Gb	1 hour

Hadoop: Stability at Scale

Hadoop powers the Yahoo! Network: must be rock-solid

We fix bugs before you see them

- We run very large clusters
- We have a large QA effort
- We run a huge variety of workloads
- Recent spinoff (HortonWorks)



Related Work

- Amazon, Netflix, Y! Music, etc.:
  - Collaborative filtering with large content pool (see KDD Cup, 2011)
  - Achieve lift by eliminating bad articles
  - We have a small number of high quality articles
- Search, Advertising
  - Matching problem with large content pool
  - Match through feature based models

Summary of Approach

- Offline models to initialize online models
- Online models to track performance

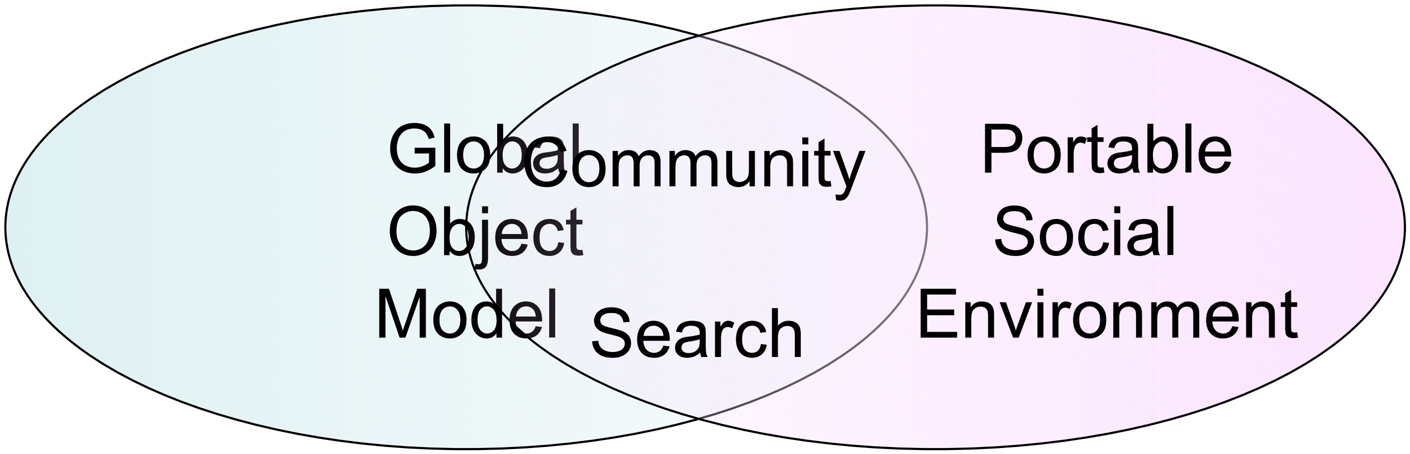
- Explore/exploit to converge fast

- Study user visit patterns and behavior; program content accordingly

The Social Web

Social networks and online communities will result in personalized experiences in a number of ways

PeopleWeb:
Site-Centric ⇨ People-Centric



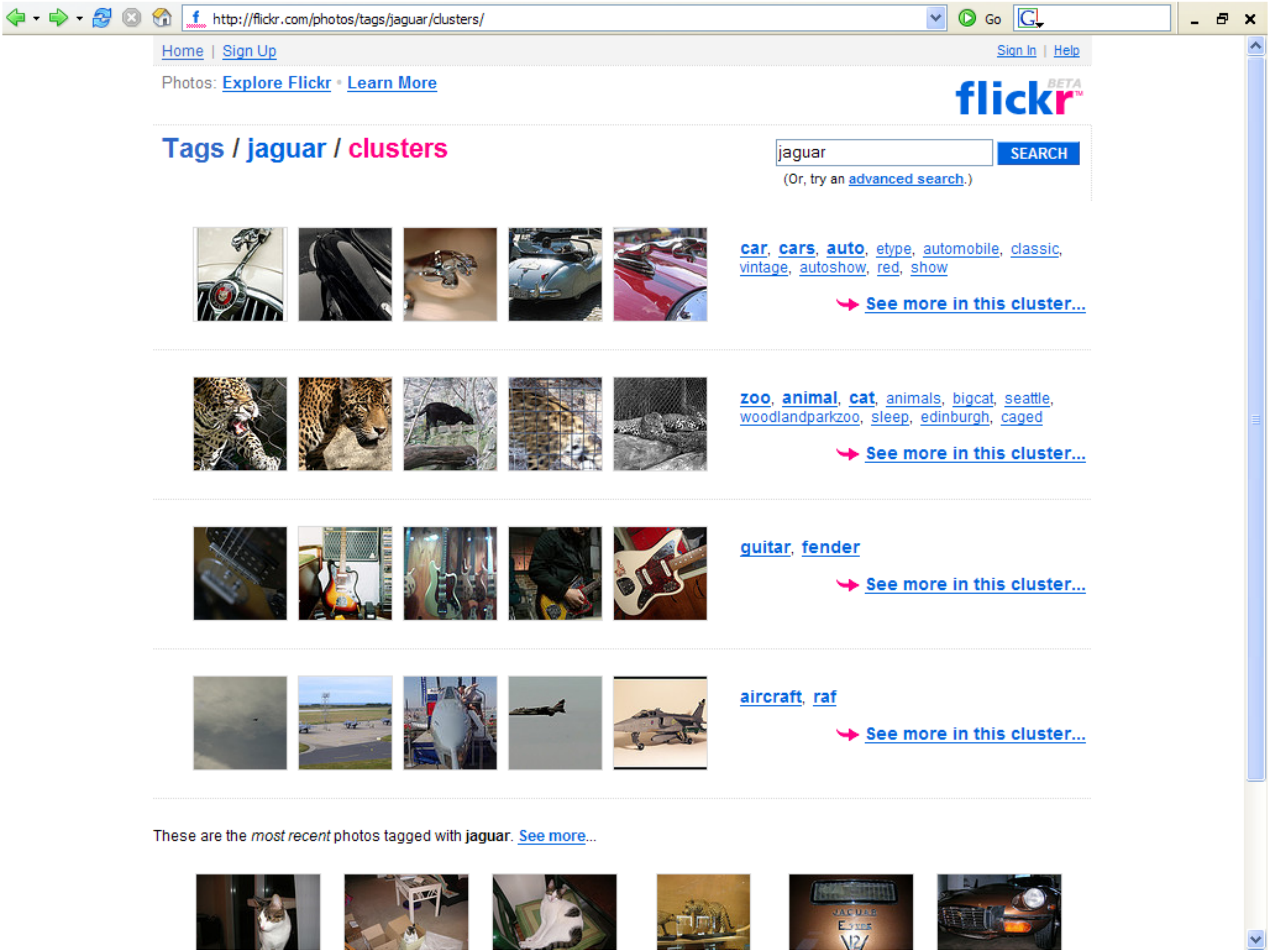
- **Common web-wide id for objects (incl. users)**
- As users move across sites, their personas and social networks will be carried along
- **Increased semantics on the web through community activity (another path to the goals of the Semantic Web)**

(Towards a PeopleWeb, Ramakrishnan & Tomkins, IEEE Computer, August 2007)

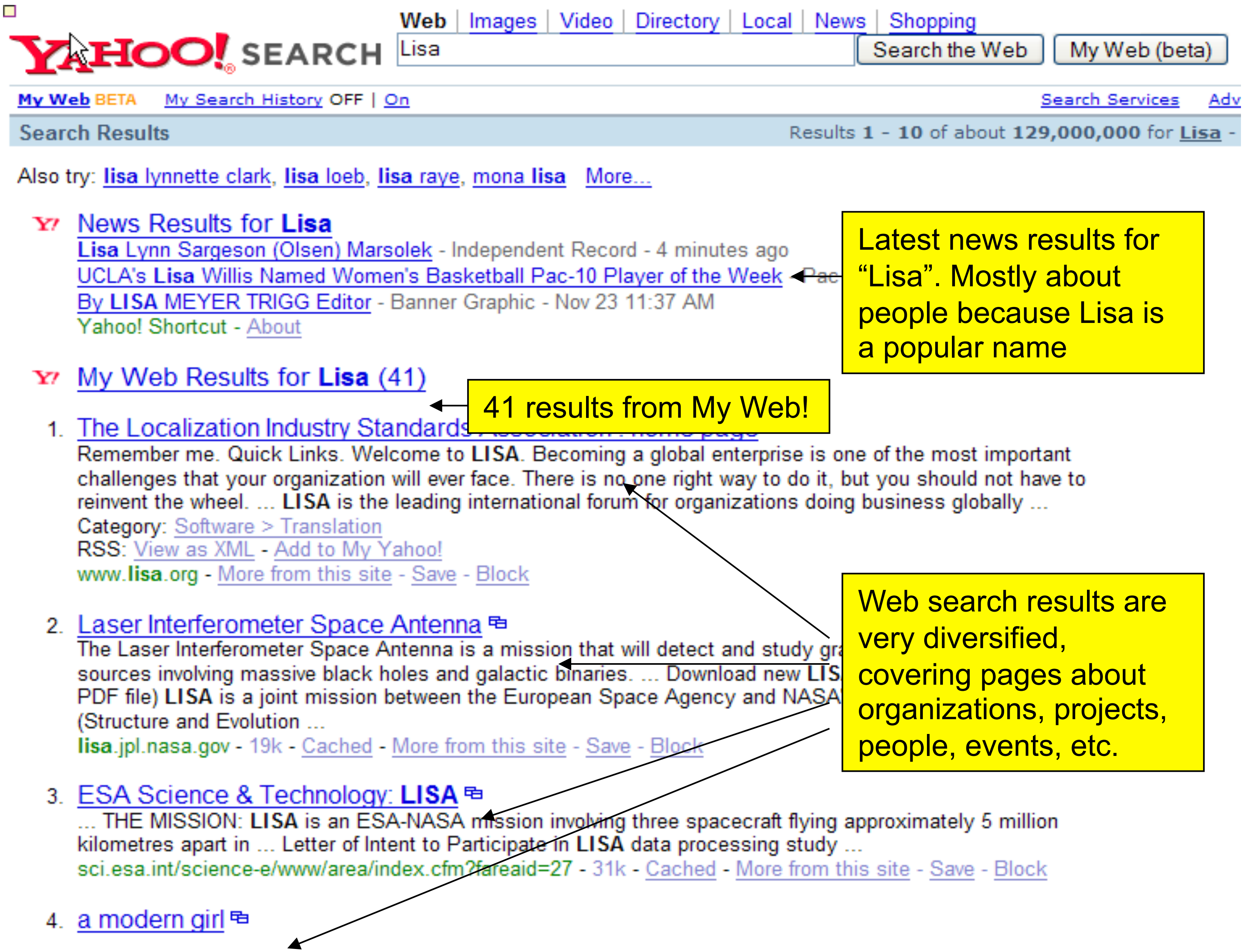
The Power of Social Media

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

(Courtesy: Prabhakar Raghavan)


http://flickr.com/photos/tags/jaguar/clusters/
Go
Home | Sign Up
Sign In | Help
Photos: Explore Flickr • Learn More
flickr BETA
Tags / jaguar / clusters
jaguar
SEARCH
(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...

Web Search Results for *Lisa*



“Social” Search Results for *Lisa*

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

Yahoo! My Web 2.0

My Web 2.0 Home Page

YAHOO! SEARCH
My Web 2.0 BETA

[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]**

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 knowledge

Community vitality
- Statistics
- Recent pages
- Popular pages

Public knowledge

Everyone's Web: 319,070 Saved Pages | 84,650 Tags

Recent Pages | Popular Pages

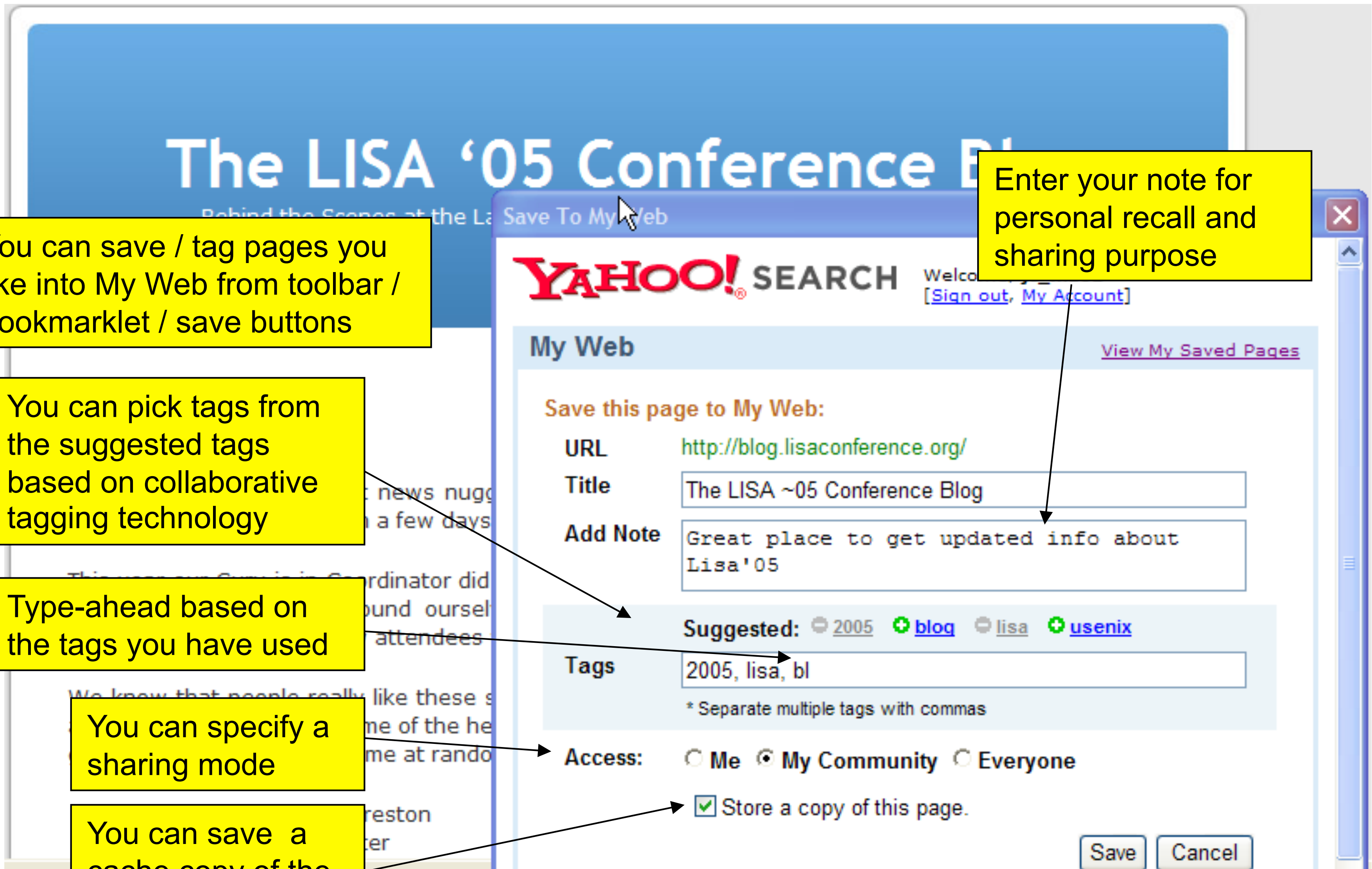
Tips for Working Securely from Hotspots
Shared by: hermosa1001 Tags: security, wireless 14 minutes ago - Details - Save

Public vitality
- Statistics
- Recent pages
- Popular pages

(Courtesy: Raymie Stata)

45

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

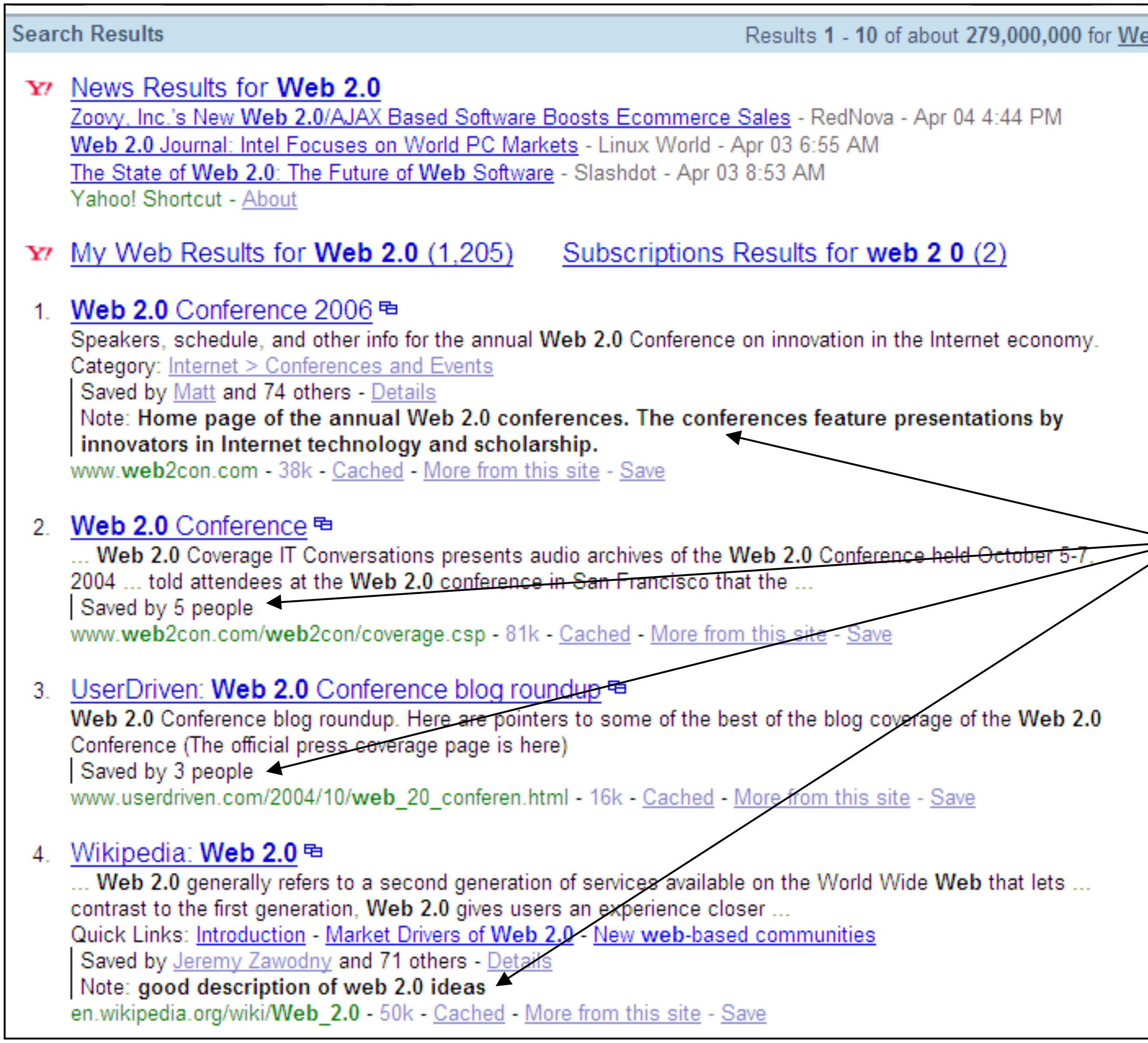
You can save a cache copy of the page content

Enter your note for personal recall and sharing purpose

(Courtesy: Raymie Stata)

46

My Web Overlays



Search Results — Results **1** - **10** of about **279,000,000** for We

News Results for **Web 2.0**
Zoovy, Inc.'s New **Web 2.0**/AJAX Based Software Boosts Ecommerce Sales - RedNova - Apr 04 4:44 PM
Web 2.0 Journal: Intel Focuses on World PC Markets - Linux World - Apr 03 6:55 AM
The State of **Web 2.0**: The Future of **Web** Software - Slashdot - Apr 03 8:53 AM
Yahoo! Shortcut - About

My Web Results for **Web 2.0** (1,205) Subscriptions Results for **web 2 0** (2)

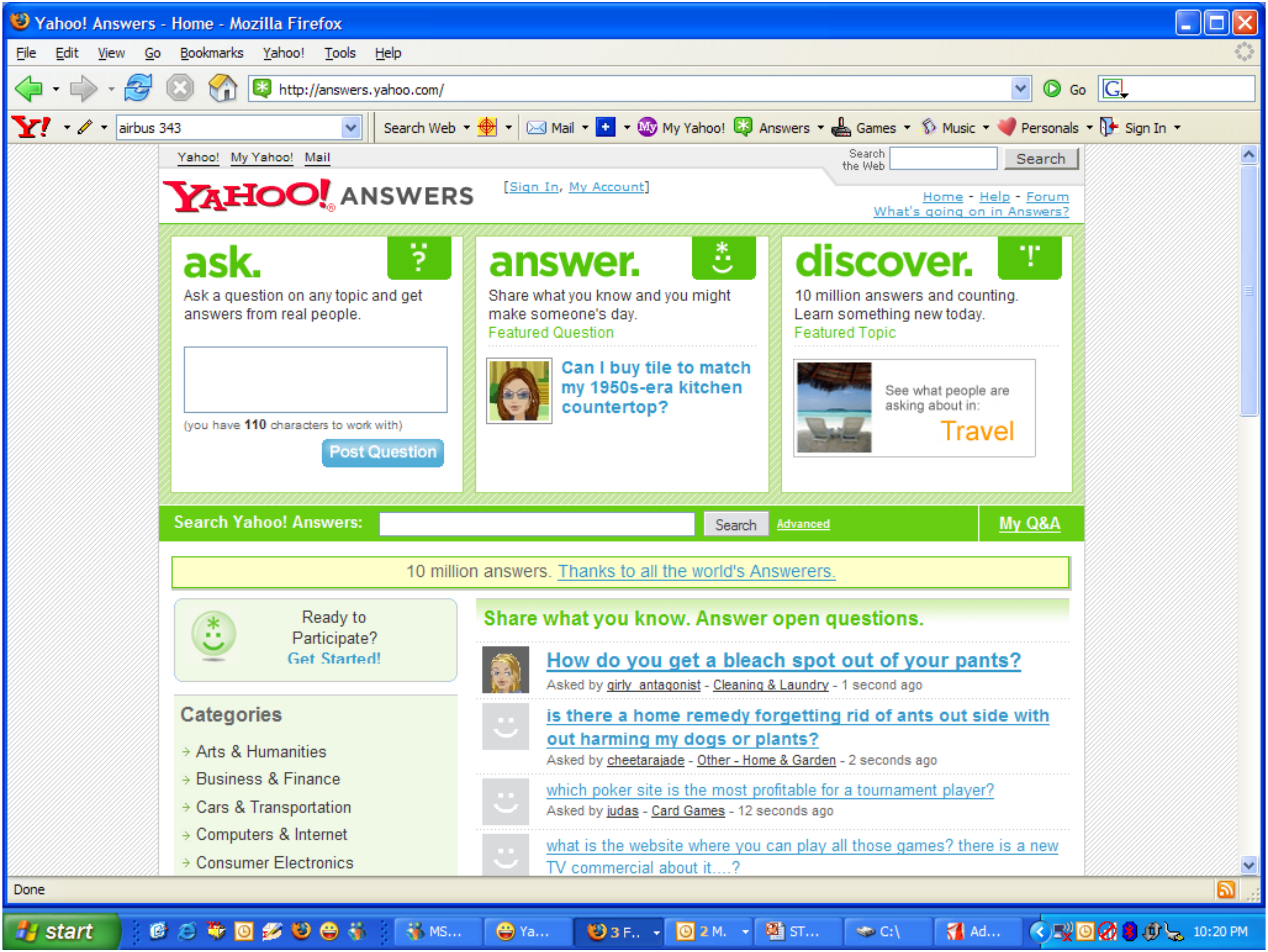
1. **Web 2.0** Conference 2006
Speakers, schedule, and other info for the annual **Web 2.0** Conference on innovation in the Internet economy.
Category: Internet > Conferences and Events
Saved by Matt and 74 others - Details
Note: **Home page of the annual Web 2.0 conferences. The conferences feature presentations by innovators in Internet technology and scholarship.**
www.**web**2con.com - 38k - Cached - More from this site - Save

2. **Web 2.0** Conference
... **Web 2.0** Coverage IT Conversations presents audio archives of the **Web 2.0** Conference held October 5-7, 2004 ... told attendees at the **Web 2.0** conference in San Francisco that the ...
Saved by 5 people
www.**web**2con.com/**web**2con/coverage.csp - 81k - Cached - More from this site - Save

3. UserDriven: **Web 2.0** Conference blog roundup
Web 2.0 Conference blog roundup. Here are pointers to some of the best of the blog coverage of the **Web 2.0** Conference (The official press coverage page is here)
Saved by 3 people
www.userdriven.com/2004/10/**web**_20_conferen.html - 16k - Cached - More from this site - Save

4. Wikipedia: **Web 2.0**
... **Web 2.0** generally refers to a second generation of services available on the World Wide **Web** that lets ... contrast to the first generation, **Web 2.0** gives users an experience closer ...
Quick Links: Introduction - Market Drivers of **Web 2.0** - New **web**-based communities
Saved by Jeremy Zawodny and 71 others - Details
Note: **good description of web 2.0 ideas**
en.wikipedia.org/wiki/**Web_2.0** - 50k - Cached - More from this site - Save

Joining My Web data into Web Search results

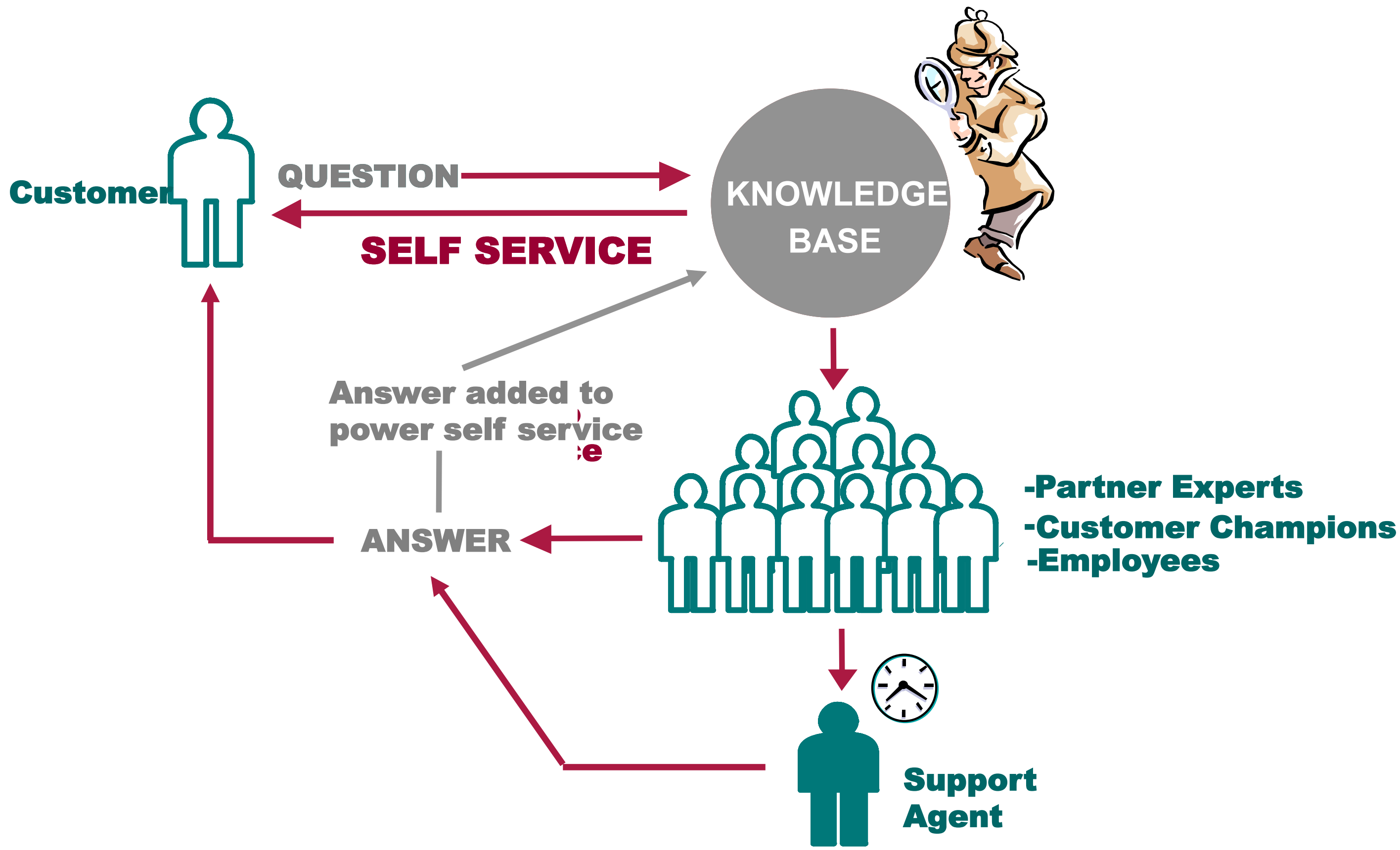

Yahoo! Answers - Home - Mozilla Firefox
File Edit View Go Bookmarks Yahoo! Tools Help
http://answers.yahoo.com/
Go
airbus 343
Search Web
Mail
My Yahoo!
Answers
Games
Music
Personals
Sign In
Yahoo! My Yahoo! Mail
Search the Web
Search
YAHOO! ANSWERS
[Sign In, My Account]
Home - Help - Forum
What's going 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.
Ready to Participate?
Get Started!
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_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....?
Done
start
10:20 PM

"In newsgroups, conversations disappear and you have to ask the same question over and over again. The thing that makes the real difference is the ability for customers to collaborate and have information be persistent. That's how we found QUIQ. It's exactly the philosophy we're looking for."

"Tech support people can't keep up with generating content and are not experts on how to effectively utilize the product … Mass Collaboration is the next step in Customer Service."

– Steve Young, VP of Customer Care, Compaq

How It


Customer
QUESTION
KNOWLEDGE BASE
SELF SERVICE
Answer added to power self service
ANSWER
-Partner Experts
-Customer Champions
-Employees
Support Agent

What is a Relevant Group? (7 M Y! groups)

- 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.

(Courtesy: Sihem Amer-Yahia)

Social Search

- Explicitly open up search
  - Enable communities, sites and consumers to explicitly re-define search results (e.g., SearchMonkey, Boss)
    - Right unit for a "search result"? Can we "stitch together" more informative abstracts from multiple sources?
    - Creation of specialized ranking engines for different tasks, or different user communities
- Implicitly leverage socially engaged users and their interactions
  - Learning from shared community interactions, and leveraging community interactions to create and refine content
- Expanding search results to include sources of information
  - E.g., Experts, sub-communities of shared interest, particular search engines (in a world with many, this is valuable!)

Reputation, Quality, Trust, Privacy

Challenges in Social Search

- How do we use annotations for better search?
- How do we cope with spam?
- Ratings? Reputation? Trust?
- What are the incentive mechanisms?

iana baltimore - Yahoo! Search Results - Mozilla Firefox

story Bookmarks Yahoo! Tools Help

http://search.yahoo.com/search;_ylt=A0oGklHEA_RLAyYBgxJXNyoA?p=eggplant+parmigiana+baltimore&fr2=sb-top&fr=moz35&sao=1

hummus ba

iana baltimore - Yah...

Web Images Video Local Shopping News More

YAHOO!

eggplant parmigiana baltimore

Search

Options

Start typing to see suggestions.

Explore related concepts:

Pizza	mushrooms
veal	tomato sauce
Chicken	Little Italy
Italian Restaurant	Crab Cake

Settings

Search Pad

SearchScan - On

30,300 results for
eggplant parmigiana ...:

Show All

Los Angeles Times

Local Business Sites

Eggplant Parmigiana Restaurants near Baltimore, MD

Neighborhood

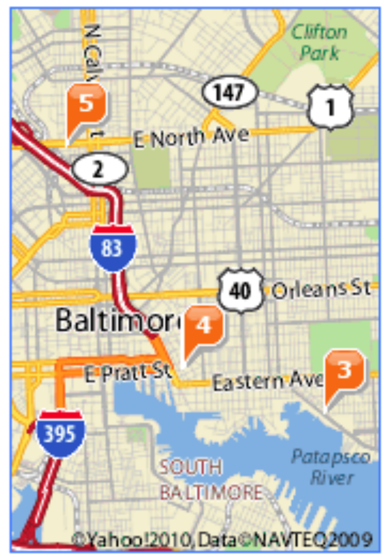
All (36)	Canton (2)	Charles North (1)	Downtown (1)
Abell (1)	Central Bal... (4)	Chinquapin ... (1)	Federal Hill (1)

1 Ciao Bella - **Baltimore** ★★★★★ (11)
local.yahoo.com
(410) 685-7733 - 236 S High St, **Baltimore**, MD
Menu: **eggplant parmigiana**
4 Reviews | Overview | 2 Photos | Directions

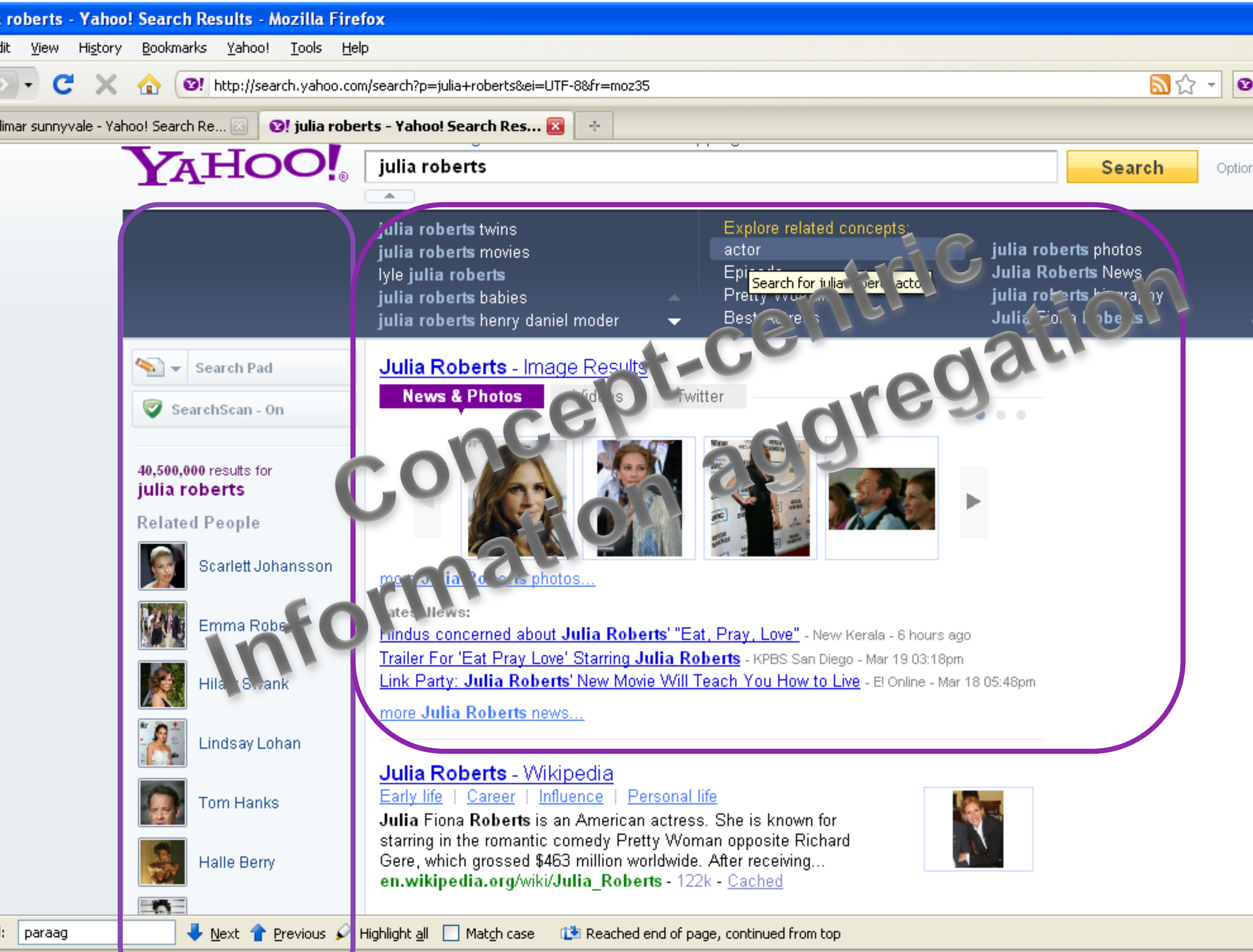
2 Amicci's ★★★★★ (20)
amiccis.com
(410) 528-1096 - 231 S High St, **Baltimore**, MD
Menu: **eggplant parmigiana**
14 Reviews | Overview | 23 Photos | Directions

3 Pasticcio ★★★★☆ (8)
local.yahoo.com
(410) 522-7700 - 2400 Boston St, **Baltimore**, MD
Menu: **eggplant parmigiana**
5 Reviews | Overview | 3 Photos | Directions

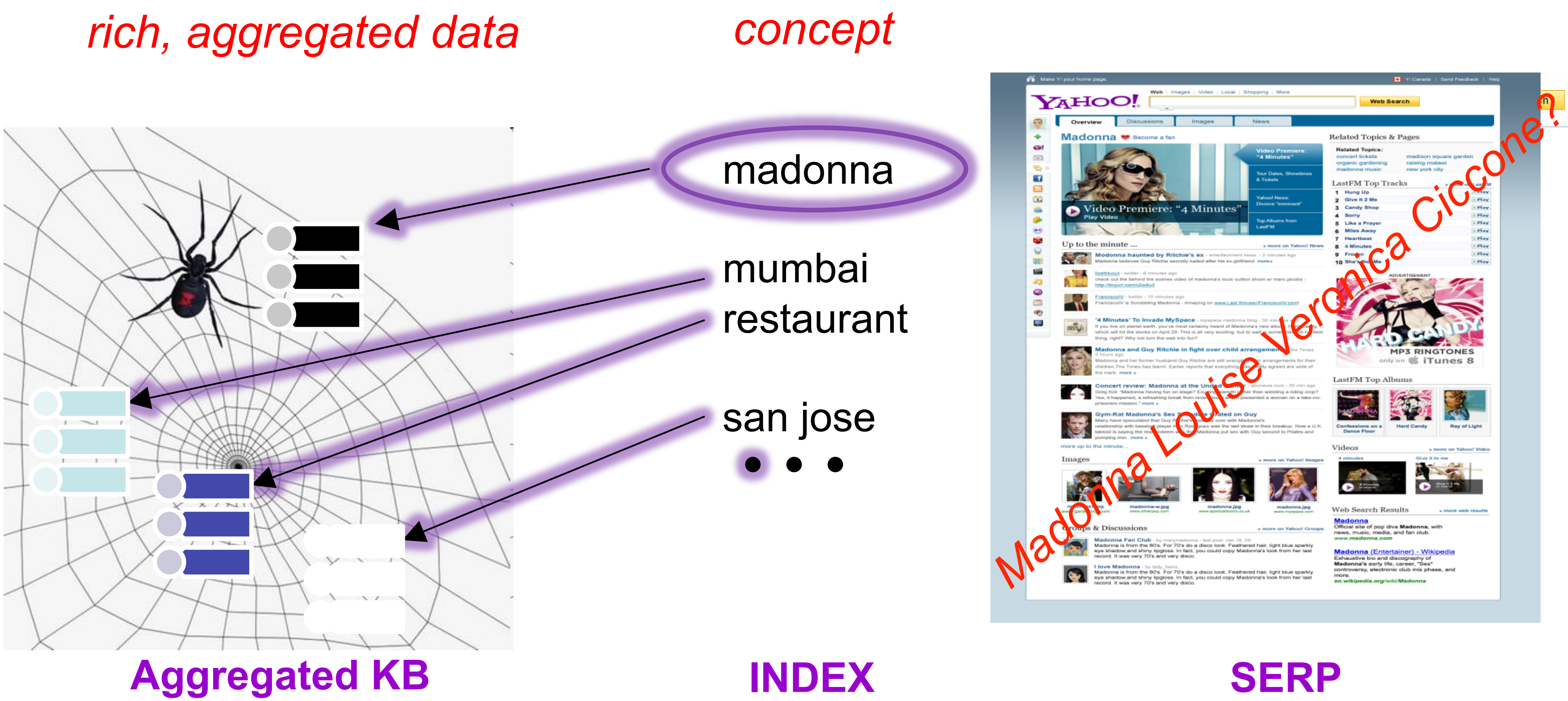
4 Caesar's Den ★★★★☆ (7)
caesarsden.com
(410) 547-0820 - 223 S High St, **Baltimore**, MD
Menu: **eggplant parmigiana**
4 Reviews | Overview | 11 Photos | Directions



ylt=A0oGki_1A_RLLt0AQ8ZXNyoA;_ylu=X3oDMTBzN2FtbDk3BHNlYwNzYwRjb2xvA3NrMQR2dGlkA0Y4NjJfOTE-/SIG=138m2fk10/EXP=1274369397/**http://local.yahoo.com/results?stx=eggplant+parmigiana

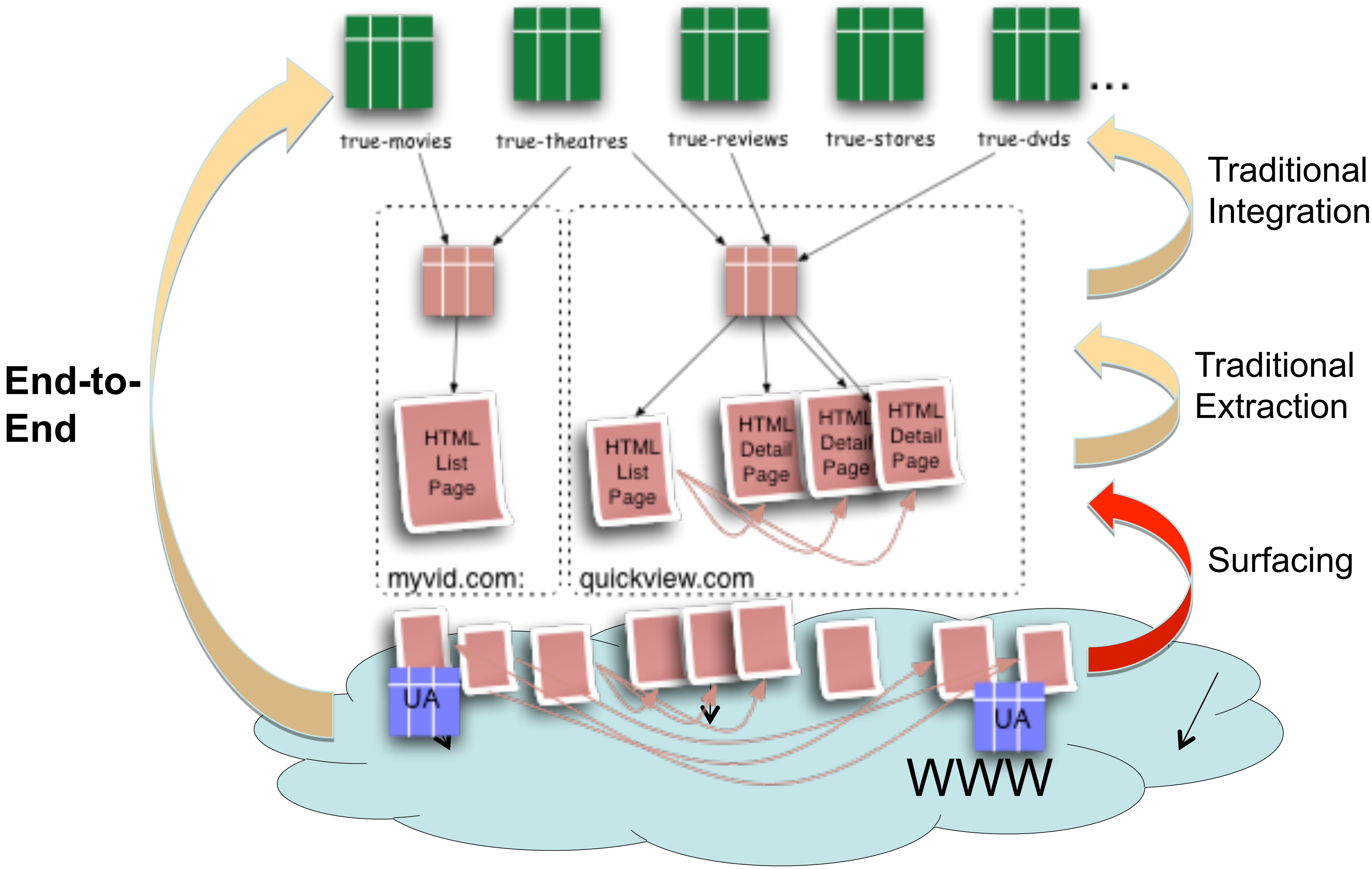

roberts - Yahoo! Search Results - Mozilla Firefox
View History Bookmarks Yahoo! Tools Help
http://search.yahoo.com/search?p=julia+roberts&ei=UTF-8&fr=moz35
limar sunnyvale - Yahoo! Search Re...
julia roberts - Yahoo! Search Res...
YAHOO!
julia roberts
Search
Concept-centric
Information aggregation
julia roberts twins
julia roberts movies
lyle julia roberts
julia roberts babies
julia roberts henry daniel moder
Explore related concepts:
actor
julia roberts photos
Julia Roberts News
Search Pad
SearchScan - On
40,500,000 results for
julia roberts
Related People
Scarlett Johansson
Tom Hanks
Halle Berry
Lindsay Lohan
Julia Roberts - Image Results
News & Photos
Twitter
Hindus concerned about Julia Roberts' "Eat, Pray, Love" - New Kerala - 6 hours ago
Trailer For 'Eat Pray Love' Starring Julia Roberts - KPBS San Diego - Mar 19 03:18pm
Link Party: Julia Roberts' New Movie Will Teach You How to Live - E! Online - Mar 18 05:48pm
more Julia Roberts news...
Julia Roberts - Wikipedia
Early life | Career | Influence | Personal life
Julia Fiona Roberts is an American actress. She is known for starring in the romantic comedy Pretty Woman opposite Richard Gere, which grossed $463 million worldwide. After receiving...
en.wikipedia.org/wiki/Julia_Roberts - 122k - Cached
paraag
Next
Previous
Highlight all
Match case
Reached end of page, continued from top

Web of Concepts



The “index” is keyed by concept instance, and organizes all relevant information (data describing the concept instance and its relationship to other instances), wherever it is drawn from, in semantically meaningful ways

Web IE: Surfacing, Extraction, Integration



Summary

- The Web will be increasingly personalized, but “personalization” will mostly happen within the context of
  - Content optimization
  - Semantic interpretation of web content and user intent
  - Socialization of the web