

Introduction to Computational Advertising

MS&E 239

Stanford University

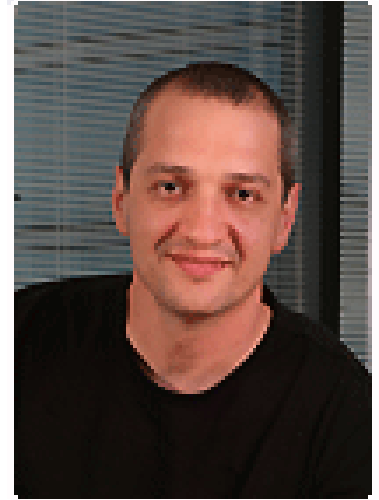
Autumn 2011

Instructors: Dr. Andrei Broder and Dr. Vanja
Josifovski, Yahoo! Research

General course info

- Course Website:
<http://www.stanford.edu/class/msande239/>
- TA: **Krishnamurthy Iyer**
 - Office hours: Tuesdays 6:30pm-7:00pm, Huang
- Course email lists
 - Staff: [msande239-aut1112-staff](#)
 - All: [msande239-aut1112-students](#)
 - Please use the staff list to communicate with the staff
- Lectures: 10am ~ 12:30pm Fridays in HP
- Office Hours:
 - After class
 - Andrei and Vanja will be on campus for 2 times each to meet and discuss with students. Feel free to come and chat about even issues that go beyond the class.

Instructor



- **Dr. Vanja Josifovski**
 - Senior Director at Yahoo! Research
 - Research Area: Computational Advertising, Search
 - Previously at IBM Research working on databases and enterprise search
 - M.Sc. from University of Florida, PhD from Linköpings University in Sweden
 - vanjaj@yahoo-inc.com
 - http://research.yahoo.com/Vanja_Josifovski

Instructor



- **Dr. Andrei Broder**

- Fellow and Vice President for Computational Advertising in Yahoo! Research
- Chief Scientist of Yahoo's Advertising Technology Group
- Research interests: computational advertising, web search, context-driven information supply, and randomized algorithms
- B. Sc. Summa cum Laude from the Technion, M.Sc. and Ph.D. in Computer Science at Stanford University under Don Knuth
 - broder@yahoo-inc.com
 - [http://research.yahoo.com/Andrei Broder](http://research.yahoo.com/Andrei_Broder)

Course Overview (subject to change)

1. 09/30 Overview and Introduction
2. 10/07 Marketplace and Economics
3. 10/14 Textual Advertising 1: Sponsored Search
4. 10/21 Textual Advertising 2: Contextual Advertising
5. 10/28 Display Advertising 1
6. 11/04 Display Advertising 2
7. 11/11 Targeting
8. 11/18 Recommender Systems
9. 12/02 Mobile, Video and other Emerging Formats
10. 12/09 Project Presentations

This lecture

1. Logistics of the course
2. Introduction & Overview of Computational Advertising
3. Detailed discussion of the requirements, goals of the course, questions, anything else

Logistics

General lecture structure

- Overview of the area: 1:15 hour
- Break: 10 minutes
- Student presentation 15 min
- In depth, discussion, and occasional quizzes: 1 hour

Class requirements

- Homework 40%
- Project 40%
- Quizzes 10%
- In-class short presentation 10%
- Active attendance strongly encouraged!

Grading and Requirements

Projects – 40%

- Group assignment
 - Preferred group size is 3
- Start forming groups!
- Project commenced at the second class
- Chose one of two projects
 - Real world online advertising
 - Algorithmic project

Real world advertising project

- Craft a campaign for a real world company
- Find a company that can benefit from online advertising
 - Local business
 - Upcoming company (startup)
- Funding provided (\$300)
- Craft campaigns
 - Major search engines
 - Display advertising Facebook/Google/Yahoo/MSFT
 - Smaller outlets that allow for customization/optimization
- Adjust campaign mid flight
- Important to learn something insightful!
- Write report and present at the last lecture (15min)

Algorithmic project

- Implement a small scale similarity search engine
- We provide a dataset based on a real world ad corpus
- Build an inverted index
- Implement two algorithms
- Explore multiple variants of the algorithms
- Write a report and present at the last class (15min)

Homework – 40%

- Total of 3 homework assignments
- Take home, open book one week to finish
- Third assignment to be a take-home, final with problems covering the whole course
- Homework are individual assignments
 - Please do not work together on the homework assignments
 - If you have any questions please contact the staff

Quizzes – 10%

- Total of 2 quizzes
- In class after the break for 10 minutes
- A few short questions covering the salient points from the class
- The goal of the quizzes is to motivate attendance

In-class presentation – 10%

- Every student should participate in a in-class presentation
- after the break, 15 minutes
- Describe the business model and the technology of a given company in the computational advertising field
- Kris, our TA will do a presentation in the next class to demonstrate the format
- Please start forming groups (3-4 people per group), and contact Kris about the scheduling.
- We might need to have 2 presentations in some classes to accommodate everybody.

Introduction to Computational Advertising

Disclaimers

- This talk presents the opinions of the authors. It does not necessarily reflect the views of Yahoo! inc or any other entity.
- Algorithms, techniques, features, etc mentioned here might or might not be in use by Yahoo! or any other company.
- These lectures benefitted from the contributions of many colleagues and co-authors at Yahoo! and elsewhere. Their help is gratefully acknowledged.

Lecture 1 plan

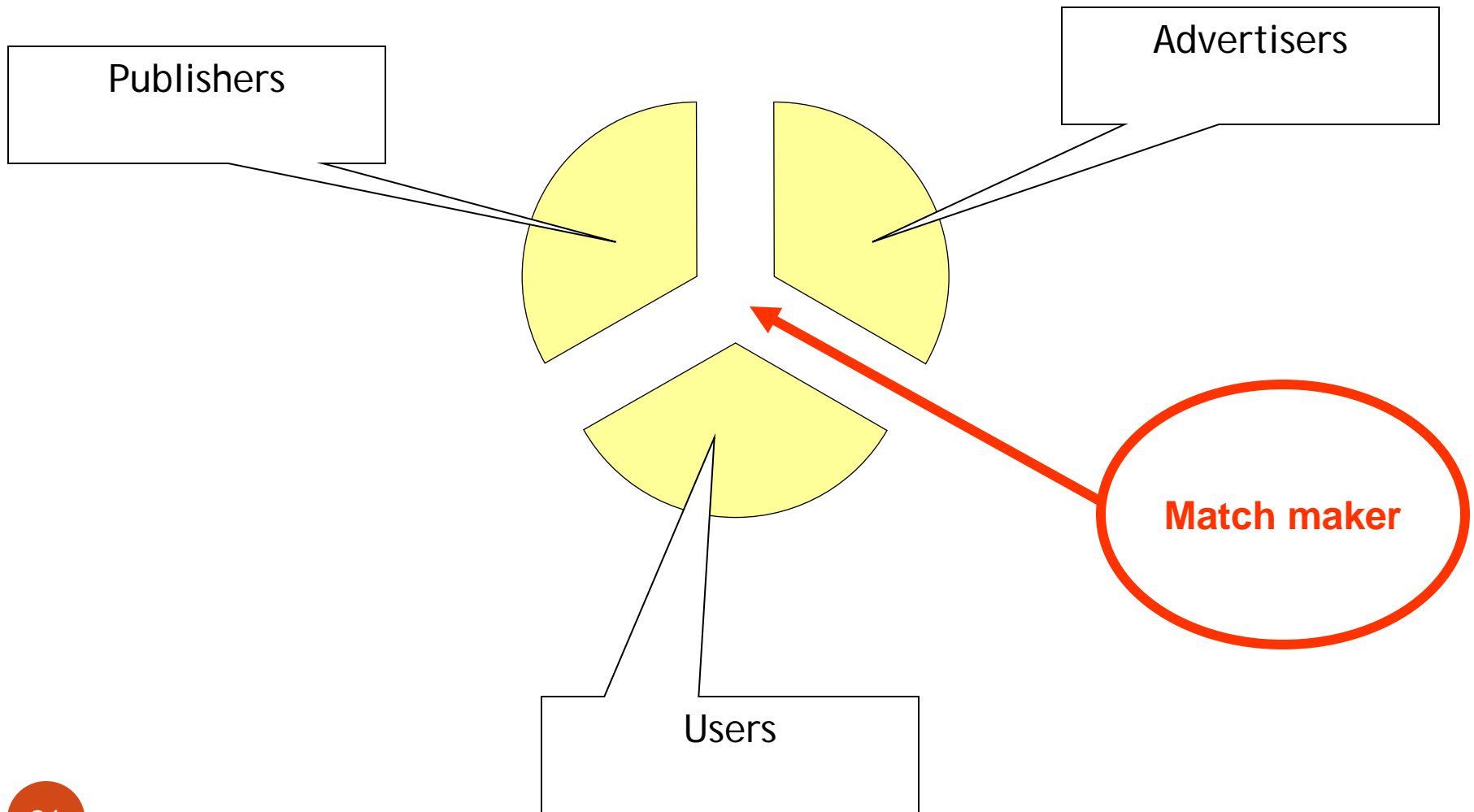
- Overview and key messages
- Classical advertising
 - Difference between classic and computational
- The opportunity: revenue and beyond
- Computational Advertising landscape
 - Graphical ads: guaranteed delivery, performance delivery, exchanges
 - Textual ads
- Ad selection
 - Textual ad selection
 - Performance graphical ads
- Mobile advertising
- Closing remarks

Computational advertising – the central challenge

Find the "best match" between a given user in a given context and a suitable advertisement.

- Examples
 - Context = Web search results → Sponsored search
 - Context = Publisher page → Content match, banners
 - Other contexts: mobile, video, newspapers, etc

Participants: Publishers, Advertisers, Users, & “Matcher”



What is “Computational Advertising”?

- New scientific sub-discipline bringing together
 - Information retrieval
 - Large scale search and text analysis
 - Statistical modeling
 - Machine learning
 - Microeconomics
 - Game theory, auction theory, mechanism design
 - Classification
 - Optimization
 - Recommender systems
 -

Establishing a new discipline...

The screenshot shows a Yahoo! search results page for the query "computational advertising". The page layout includes the Yahoo! logo, navigation tabs (Web, Images, Video, Local, Shopping, More), a search bar with the query, and a "Search" button. On the left sidebar, there is a "Search Pad" icon, a "SafeSearch - On" indicator, and a result count of "8,020,000 results for computational advert...". The main content area displays four search results, each with a blue link title, a snippet of text, and a green URL with a "Cached" link.

Web Images Video Local Shopping More ▾

computational advertising

Search Options ▾

Search Pad

SafeSearch - On

8,020,000 results for computational advert...

[Computational Advertising | Yahoo! Research](#)
Computational advertising is a new scientific sub-discipline, at the ... Introduction to Computational Advertising, Course at Management Sciences and ...
[research.yahoo.com/Computational_Advertising](#) - [Cached](#)

[Phil Windley's Technometria | Computational Advertising](#)
Organizations Get the ... Computational Advertising. Andrei Broder of Yahoo! Research ... Computation Advertising on, what else, computational advertising. ...
[windley.com/archives/2008/04/computational_advertising.shtml](#) - [Cached](#)


[Computational advertising - Lorcan Dempsey's weblog](#)
A new discipline - Computational Advertising - has recently emerged, which ... Computational advertising poses numerous challenges and open research problems ...
[orweblog.oclc.org/archives/001633.html](#) - [Cached](#)

[Computational Advertising](#)
Trends in ideas, concepts, and gestalt ... The central challenge of computational advertising is to find the "best match" ... Personalized advertising is a ...
[kk.org/ct2/2008/01/computational-advertising.php](#) - [Cached](#)


[Key Scientific Challenges: Computational Advertising | Yahoo! ...](#)
Computational Advertising. By By Andrei Broder and Vanja Josifovski. Computational advertising is a new scientific sub-discipline, at the ...
[research.yahoo.com/Computational_Advertising](#) - [Cached](#)

Establishing a new discipline...

Web [Images](#) [Videos](#) [Maps](#) [Finance](#) [Translate](#) [Gmail](#) [more](#) ▼ [Web History](#) | [Search settings](#) | [Sign in](#)

 SafeSearch strict ▼

About 956,000 results (0.18 seconds) [Advanced search](#)

 **Everything**
☐ [More](#)

The web
[Pages from Hong Kong](#)

Standard view
[Wonder wheel](#)
☐ [More search tools](#)

[Computational Advertising](#) | [Yahoo! Research](#)
Computational advertising is a new scientific sub-discipline, at the intersection of information retrieval, machine learning, optimization, ...
[research.yahoo.com/Computational_Advertising](#) - [Cached](#) - [Similar](#) - [Filter](#)

[Stanford University - Introduction to Computational Advertising](#)
14 Nov 2009 ... **Computational advertising** is a new scientific discipline, at the intersection of information retrieval, machine learning, optimization, ...
[Course Information](#) - [Course Schedule](#) - [Lecture Handouts](#)
[www.stanford.edu/class/msande239/](#) - [Cached](#) - [Similar](#) - [Filter](#)

[Geeking with Greg: Lectures on Computational Advertising](#)
4 Jan 2010 ... Slides from all the lectures of Andrei Broder's recent **Computational Advertising** class at Stanford University now are available online. ...
[glinden.blogspot.com/.../lectures-on-computational-advertising.html](#) - [Cached](#) - [Similar](#) - [Filter](#)

[\[PDF\] Introduction to computational advertising](#)
File Format: PDF/Adobe Acrobat - [Quick View](#)
computational advertising. Bogdan Cautis. Athens week, March 2009 page 1. Outline. ■
From IR to IS. ■ Advertising on the Web ...
[pierre.senellart.com/enseignement/2008-2009/.../4.../CompAdv.pdf](#) - [Similar](#) - [Filter](#)

[Computational advertising](#)
by A Broder - 2008 - [Cited by 3](#) - [Related articles](#)

Sponsored links
[Semantic R&D](#)
Web **advertising**, NLP, IR, machine learning: papers, software, data
[sites.Google.com/site/massiciara](#)

[Tutorial at SIGIR 2010](#)
Information Retrieval Challenges in **Computational Advertising**
[research.yahoo.com/tutorials/sigir](#)

[See your ad here »](#)

Establishing a new discipline...

The screenshot shows a Bing search results page for the query "computational advertising". The top navigation bar includes links for Web, Images, Videos, Shopping, News, Maps, More, MSN, and Hotmail. The Bing logo is on the left, and the search bar contains the text "computational advertising" with a magnifying glass icon. Below the search bar, the results are displayed in a list. The first result is titled "Computational Advertising | Yahoo! Research" and describes it as a new scientific sub-discipline at the intersection of information retrieval, machine learning, optimization, and microeconomics. The second result is titled "Computational advertising and recommender systems" and mentions the ACM SIGKDD. The third result is titled "Computational Advertising - John Battelle's Searchblog" and discusses connecting brands with people. The fourth result is titled "Phil Windley's Technometria | Computational Advertising" and mentions Phil's tweet about registering his son. On the left side of the page, there is a "SEARCH HISTORY" section showing "computational advertising" and links to "See all" and "Clear all | Turn off". At the bottom left, there is a red circle with the number 25.

Web Images Videos Shopping News Maps More MSN Hotmail

bing

computational advertising

ALL RESULTS 1-10 of 676,000 results · [Advanced](#)

SEARCH HISTORY
[computational advertising](#)
[See all](#)
[Clear all](#) | [Turn off](#)

[Computational Advertising | Yahoo! Research](#)
Computational advertising is a new scientific sub-discipline, at the intersection of information retrieval, machine learning, optimization, and microeconomics.
[research.yahoo.com/Computational_Advertising](#) · [Cached page](#)

[www.cs.washington.edu](#)
[www.cs.washington.edu/htbin-post/mvis/mvis?ID=676](#)

[Computational advertising and recommender systems](#)
ACM: Association for Computing Machinery SIGKDD: ACM Special Interest Group on Knowledge Discovery in Data
[doi.acm.org/10.1145/1454008.1454009](#) · [Cached page](#)

[Computational Advertising - John Battelle's Searchblog](#)
I think humans are required anytime you want to connect a brand with a person in any kind of meaningful way. But "**computational advertising**" is one way to optimize that connection ...
[battellemedia.com/archives/004204.php](#) · [Cached page](#)

[Phil Windley's Technometria | Computational Advertising](#)
Organizations Get the IT They Deserve ... Phil tweeted, "Registering my son, who's in Brazil, for classes at BYU. Fun times at midnight."
[www.windley.com/archives/2008/04/computational_advertising.shtml](#) · [Cached page](#)

Technology | DOI:10.1145/1506409.1506415

David Essex

Matchmaker, Matchmaker

Computational advertising seeks to place the best ad in the best context before the right customer.

THE RAPIDLY CHANGING advertisements that appear on Web pages are often chosen by sophisticated algorithms that match ad keywords to words on a Web page. Take the Chevy ad, for example, that frequently appears on your favorite news site. A real-time ad network at one of the major search engines—Google, MSN, and Yahoo!—might place it on a page of automotive news. But what if the news page's featured article is about a tragic accident caused by a mechanical failure in a Chevy SUV? That's not a page General Motors wants to be associated with, let alone pay good money to advertise on.

Costly mishaps like this could be avoided by a new discipline called computational advertising, which seeks to put the best ad in the best context before the right customer. It draws from numerous fields, including information retrieval, machine learning, natural-language processing, microeconomics, and game theory, and tries to match ads with a variety of user scenarios, such as querying a search engine, reading a Web page, watching a video on YouTube, or instant messaging a friend.

Computational advertising could spur the Web's growth as a medium of mass customization. Better ad matching could quicken the trend toward personalization, making highly specialized magazines, Web sites, and TV channels more financially viable. "Advertising has been the engine that has powered the huge development of the Web," says Andrei Broder, fellow and vice president for computational advertising at Yahoo! Research. "Without advertising, you would not have blogs and search engines."

Computational advertising is a type of automation that tries to replicate what humans might do if they had the time to read Web pages to discern their content and find relevant



Andrei Broder, vice president for computational advertising at Yahoo! Research, presenting a tutorial on Web search and advertising at the 30th Annual International ACM SIGIR Conference in Amsterdam.

ads among the millions available. "In the old world of advertising, they deal with few choices and large amounts of money for each choice," Broder says. "We deal with maybe a hundred million potential ads, each worth a fraction of a cent."

A Perfect Match

There are basically three kinds of Web ads. Sponsored search ads are matched to the results of search engine queries; banner ads target particular demographics and venues, typically without regard to a page's content; and contextual advertising, also called context match, applies to other types of Web pages, such as the home page of a financial news site. Computational advertising addresses all three types of ads.

Google, MSN, and Yahoo! use electronic auctions to assign ads to their own results pages and the pages of other Web sites. "Google is a vendor, or matchmaker, says Google chief econo-

mist Hal Varian. "The goal is to get a perfect match."

In sponsored search, advertisers bid to place ads that contain keywords correlated to words in a user's search string. For contextual advertising, the keywords are related to words on the entire page, and the search engine's advertising service places the ads. For banner ads, online ad networks place ads on sites whose topics and audiences match the advertiser's criteria.

Before the advent of computational advertising, ad engines could make mistakes more simple-minded than the Chevy SUV scenario. Suppose, for example, a news page contains the word "flowers." If the article isn't about flowers but instead revisits the Rolling Stones' underrated 1967 record *Flowers*, the reader is unlikely to want ads from florists. The old method of analyzing co-occurring words and phrases doesn't help much, and neither does frequency. "You could extract a word used many times in the ar-

ticle and it still is not what the article is about," Broder says.

Therefore, Broder and the 30 researchers who work for him are finding ways to glean the meaning of a page. One promising avenue combines semantic and syntactic features. A semantic phrase categorizes the page and the ads into a 6,000-node topic taxonomy and compares the proximity of the two types of classes as a factor in ranking ads. The hierarchical taxonomy also improves the matching of ads that don't fit a page's exact topic. Keyword matching is still needed to capture more granular content, such as a specific brand of automobile. "We decided that what the article is about should count for about 80% and the words should count for 20%," Broder says.

Another area of interest is using statistical analysis to measure the effect of exogenous events on browsing behavior and adjust the advertisements accordingly. Varian cites short-lived examples, such as this year's rare snowfall in England, or longer-term ones such as the worldwide recession. "In the last few months, there is a big increase in interest in price-sensitive products," Varian says. "The advertisers, in turn, are trying to respond."

All three companies are close-lipped about which of their research has been commercialized, but say that new ideas for algorithms are quickly incorporated into their bidding mechanisms and advertiser tools. Bottom-line results are secret, but the search engines all collect metrics such as revenue per search.

Machine learning, another major focus, concentrates on training algorithms to scan pages for meaning, a technique employed successfully on single-topic documents with the aid of machine-generated labels, but trickier to perform on Web pages, with their assortment of graphics, text, and topics. Microsoft researchers have learned how to employ a type of multiple instance learning to automate classification of sub-documents on pages with incomplete labels and to detect the presence of certain types of content.

"Most of what we do can be boiled down to understanding intent," says Eric Brill, general manager of Microsoft adCenter Labs. By analyzing search strings, for example, algorithms can predict if a person is interested in ads. Some strings are pure attempts at finding information, while others, such as "buy Canon digital camera," have clear commercial intent. "When consumers don't have commercial intent, you don't want to put ads in front of them," Brill says.

Much work focuses on ensuring that new bidding mechanisms don't have incentives for advertisers to misrepresent click-through rates to get better ad placement. In the decentralized economy of the Internet, truthfulness is a currency reinforced by carefully crafted algorithms. "People are out there to make money," says Thore Graepel, a senior researcher at Microsoft Research. "We need to build mechanisms where everyone benefits."

One might expect the speed and volume of data to create a capacity problem, but the researchers express mixed opinions. Graepel says semantic analysis creates an extra burden. "You will hit a computational bottleneck, that's pretty clear," he says. To avoid this, researchers optimize algorithms to make the best decisions with the smallest possible data sets. But they also have faith in engineers' ability to exploit techniques such as parallel processing. "It's surprising how they are always able to scale to deal with these new algorithms," Varian says.

Privacy regulations remain an obstacle to personalizing ads, says Graepel. The existing opt-in, opt-out model lets users choose to reveal personal data in exchange for discounts and other incentives. Researchers are also investigating aggregating data on Web traffic to more accurately match ad categories with coarsely defined groups of users who identify their interests simply by visiting certain types of Web sites.

Fortunately, there is hope for avoiding embarrassments like the ill-placed Chevy ad. Researchers at Microsoft adCenter Labs claim their sub-document classification methods can prevent incompatible ads and Web sites from ever hooking up. You might call it a reverse matchmaker, just the sort of odd little entity the Internet's inventors might never have imagined. ■

David Essex is a freelance science writer based in Penarth, South Wales.

© 2009 ACM 0007-0760/09/050000 \$5.00

Education

Computer Science Enrollment Increases

Enrollment in computer science classes in the United States has increased for the first time in six years, according to the Computing Research Association's (CRA's) annual Taubman Survey.

Total enrollment by majors and pre-majors in computer science is up 6.2% per department over last year. If only majors are considered, the increase is 8.1%, according to the CRA survey, which collected enrollment data in fall 2008 from computer

science and computer engineering departments at 192 Ph.D.-granting universities.

"The upward surge of student interest is real and bigger than anyone expected," says Peter Lee, incoming chair of CRA. "The fact that computer science graduates usually find themselves in high-paying jobs accounts for part of the reversal. Increasingly students also are attracted to the intellectual depth and societal benefits of computing technology."

Computer science graduates on average earn 13% more than the average college graduate, according to the U.S. Department of Labor, and future job prospects for computer science graduates are higher than for any other science or engineering field.

The average number of new students per department majoring in computer science is up 9.5% over last year. Computer science departments are replenishing the freshman and sophomore ranks with larger

groups than they are graduating as seniors, and computer science graduation rates should increase in two to four years as these new students graduate.

The total number of Ph.D. graduates among responding departments grew to 1,877 for the period July 2007 to June 2008, a 5.7% increase over the previous year. One area that didn't show improvement is the number of women pursuing computer science degrees, which held steady at 11.8%.

Key messages


1. **Computational advertising = A principled way to find the "best match" between a user in a context and a suitable ad.**
2. **The financial scale for computational advertising is huge**
 - Small constants matter
 - Expect plenty of further research
3. **Advertising is a form of information.**
 - Adding ads to a context is similar to the integration problem of other types of information
 - Finding the “best ad” is a type of information retrieval problem with multiple, possibly contradictory utility functions
4. **New application domains and new techniques are emerging every day**
 - Good area for research + new businesses!

Classic Advertising

Long history: Advertising for coffee, London 1657




The Vertue of the *COFFEE* Drink..

First publiquely made and sold in England, by *Pasqua Rosee*. 

THE Grain or Berry called *Coffee*, groweth upon little Trees, only in the *Deserts of Arabia*.

It is brought from thence, and drunk generally throughout all the Grand Seigniors Dominions.

It is a simple innocent thing, composed into a Drink, by being dried in an Oven, and ground to Powder, and boiled up with Spring water, and about half a pint of it to be drunk, fasting an hour before, and not Eating an hour after, and to be taken as hot as possibly can be endured; the which will never fetch the skin off the mouth, or raise any Blisters, by reason of that Heat.

The Turks drink at meals and other times, is usually *Water*, and their Dyets consist: much of *Erug*, the *Crudities* whereof are very much corrected by this Drink. 

The quality of this Drink is cold and Dry; and though it be a Dryer; yet it neither heats, nor inflames more then *hot Posset*.

It secludeth the Orifice of the Stomack, and fortifies the heat with- its very good to help digestion, and therefore of great use to be hout 3 or 4 a Clock afternoon, as well as in the morning.

It quickens the Spirits, and makes the Heart *Lightsome*.

It is good against sore Eys, and the better if you hold your Head over it, and take in the Steem that way.

It suppresseth Fumes exceedingly, and therefore good against the Head-ach, and will very much stop any *Defluxion of Rheums*, that distil from the Head upon the Stomack, and so prevent and help *Consumptions*, and the *Cough of the Lungs*.

It is excellent to prevent and cure the *Dropsy*, *Gout*, and *Scurvy*.

It is known by experience to be better then any other Drying Drink for People in years, or Children that have any running humors upon them, as the *Kings Evil*. &c.

It is very good to prevent *Mis-carryings* in *Child-bearing Women*.

It is a most excellent Remedy against the Spleen, *Hypocondriack Winds*, or the like.

It will prevent *Drowsyness*, and make one fit for business, if one have occasion to Watch, and therefore you are not to Drink of it after Supper, unless you intend to be watchful, for it will hinder sleep for 3 or 4 hours.

It is observed that in Turkey, where this is generally drunk, that they are not troubled with the Stone, Gout, Dropsie, or Scurvy, and that their Skins are exceeding clear and white.

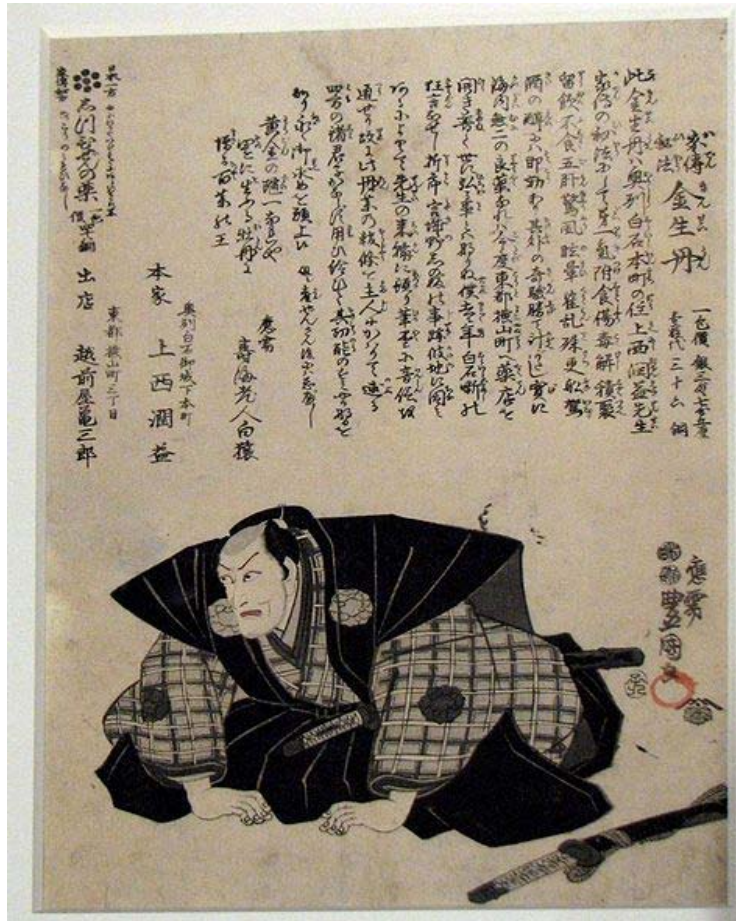
It is neither Laxative nor Restraining.



Made and Sold in *St. Michaels Alley* in *Cornhill*, by *Pasqua Rosee*, at the Signe of his own Head.

The ad explains “*The Vertue of the COFFEE drink*” what coffee is, how it grows, how it cures numerous maladies, including Dropsy, Gout, and Scurvy, ...

Long history....



Japan ,1806



USA,1890

Brand advertising

Goal: create a distinct favorable image



Direct marketing

Advertising that involves a "direct response": buy, subscribe, vote, donate, etc, now or soon

Round Table PIZZA

PALO ALTO/EAST PALO ALTO
263 University Avenue
(Downtown/Delivery to Stanford)
650-322-2893

KING ARTHUR'S SUPREME

Round Table PIZZA

We Deliver

20% OFF Any Order
Offer excludes beverages, Manager's Specials, Kids Meal or any promotional items.
Offer valid on Drive-In, Carry-out, or Delivery. Limited delivery area & hours. Minimum delivery fee may apply. Not valid with any other offer or discounts. Expires 7/1/06.

\$15.99 Any Large Specialty Pizza
Original or Skinny Crust only.
Offer valid on Drive-In, Carry-out, or Delivery. Limited delivery area & hours. Minimum delivery fee may apply. Not valid with any other offer or discounts. Expires 7/1/06.

\$11.99 Any Large 1-Topping Pizza
Original or Skinny Crust only.
Offer valid on Drive-In, Carry-out, or Delivery. Limited delivery area & hours. Minimum delivery fee may apply. Not valid with any other offer or discounts. Expires 7/1/06.

\$12.99 Any Large 2-Topping Pizza
Original or Skinny Crust only.
Offer valid on Drive-In, Carry-out, or Delivery. Limited delivery area & hours. Minimum delivery fee may apply. Not valid with any other offer or discounts. Expires 7/1/06.

\$5.00 OFF ANY X-LARGE PIZZA
\$4.00 OFF ANY LARGE PIZZA
\$2.00 OFF ANY MEDIUM PIZZA
Offer valid on Drive-In, Carry-out, or Delivery. Limited delivery area & hours. Minimum delivery fee may apply. Not valid with any other offer or discounts. Expires 7/1/06.

FREE Medium 1 Topping Pizza
with the purchase of any large or X-Large Specialty at regular menu price.
Offer valid on Drive-In, Carry-out, or Delivery. Limited delivery area & hours. Minimum delivery fee may apply. Not valid with any other offer or discounts. Expires 7/1/06.

ADVERTISE WITH MONEY MAILER OF PALO ALTO/08/ALTO/SMITH VIEW (250) 660-1234
226-07-0082 226-07-0082



H.O.T! Coupons Web Ad • 226-07-0082F

©2006 Money Mailer LLC
<http://www.hotcoupons.com>

Why “computational” advertising?

Lots of computational this and that ...

- Computational Linguistics
- Computational Biology
- Computational Chemistry
- Computational Finance
- Computational Geometry
- Computational Neuroscience
- Computational Physics
- Computational Mechanics
- Computational Economics
- ...

All are about

- a) Mixing an old science with computing capabilities
- b) Thinking algorithmically about an old challenge

Why go “computational” ?

- Classical:
 - Relatively few venues – magazines, billboards, newspapers, handbills, TV, etc
 - High cost per venue (\$3Mil for a Super Bowl TV ad)
 - No personalization possible
 - Targeting by the wisdom of ad-people
 - Hard to measure ROI
- Computational – almost the exact opposite:
 - Billions of opportunities
 - Billions of creatives
 - Totally personalizable
 - Tiny cost per opportunity
 - Much more quantifiable

- "Half the money I spend on advertising is wasted; the trouble is I don't know which half."

John Wanamaker, ~ 1875



Computational advertising – the central challenge

Find the "best match" between a given user in a given context and a suitable advertisement.

- Examples
 - Context = Web search results → Sponsored search
 - Context = Publisher page → Content match, banners
 - Other contexts: mobile, video, newspapers, etc
- Related challenge 1: Design markets and exchanges that help in this task, and maximize value for users, advertisers, and publishers
- Related challenge 2: Build the infrastructure to support this process

The central challenge decomposed

Find the "best match" between a given user in a given context and a suitable advertisement.

1. **Representation** = represent the user, the context, and the ads in an **effective & efficient** way
2. **Definition** = define the mathematical optimization problem to capture the actual marketplace constraints and goals
3. **Solution** = solve the optimization problem in an **effective & efficient** way

Computational Advertising and Market Design



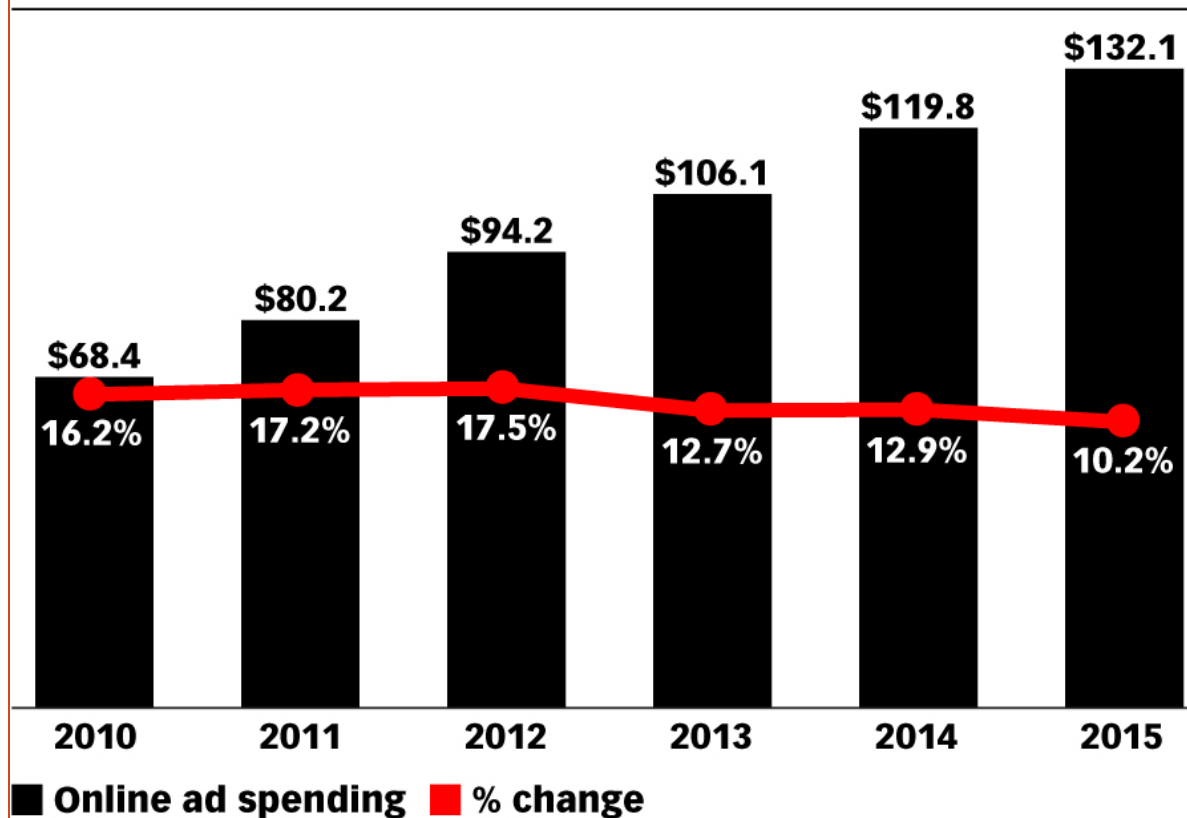
The Opportunity

The money and beyond

Online advertising is a huge business

Online Ad Spending Worldwide, 2010-2015

billions and % change



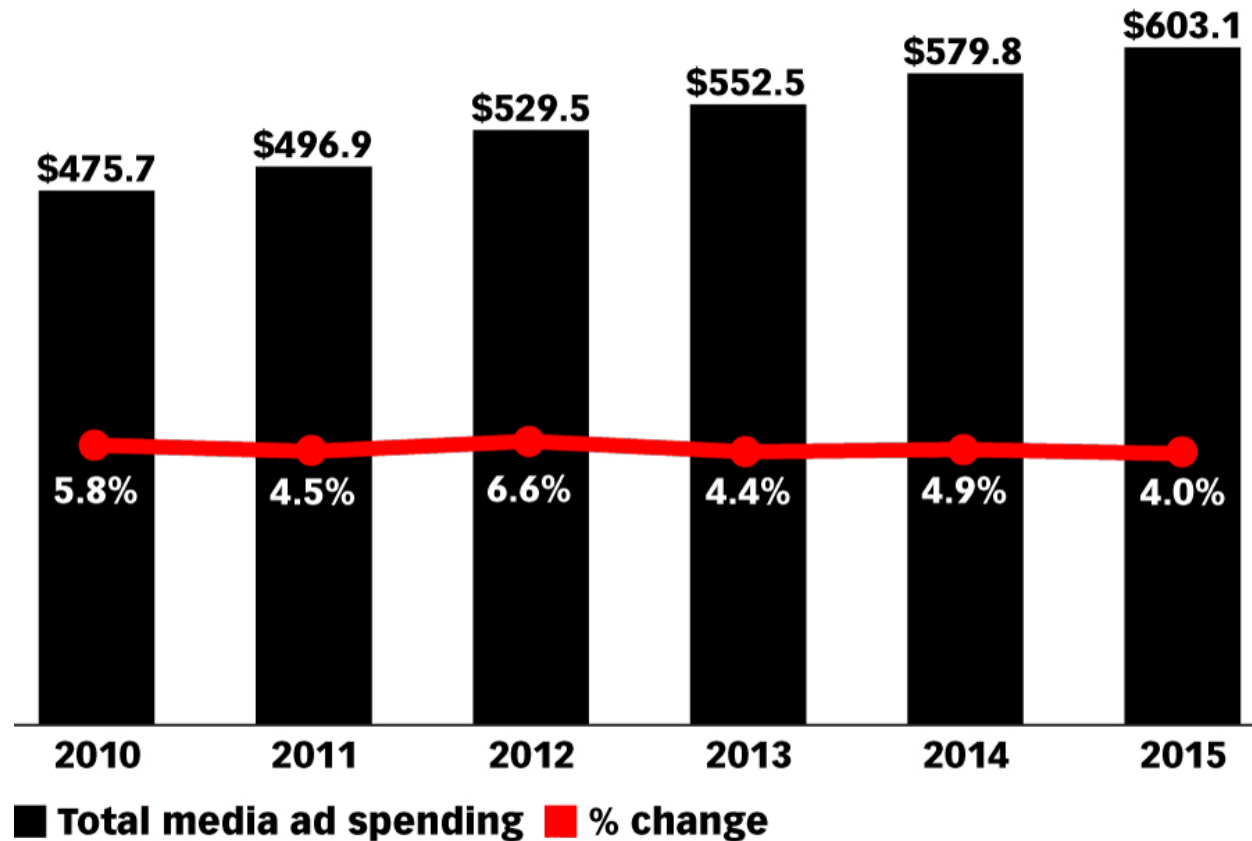
Note: includes banner ads, classifieds, email (embedded ads only), lead generation, rich media, search, sponsorships and video; includes mobile ads within the existing formats

Source: eMarketer, June 2011

Total worldwide advertising spend

Total Media Ad Spending Worldwide, 2010-2015

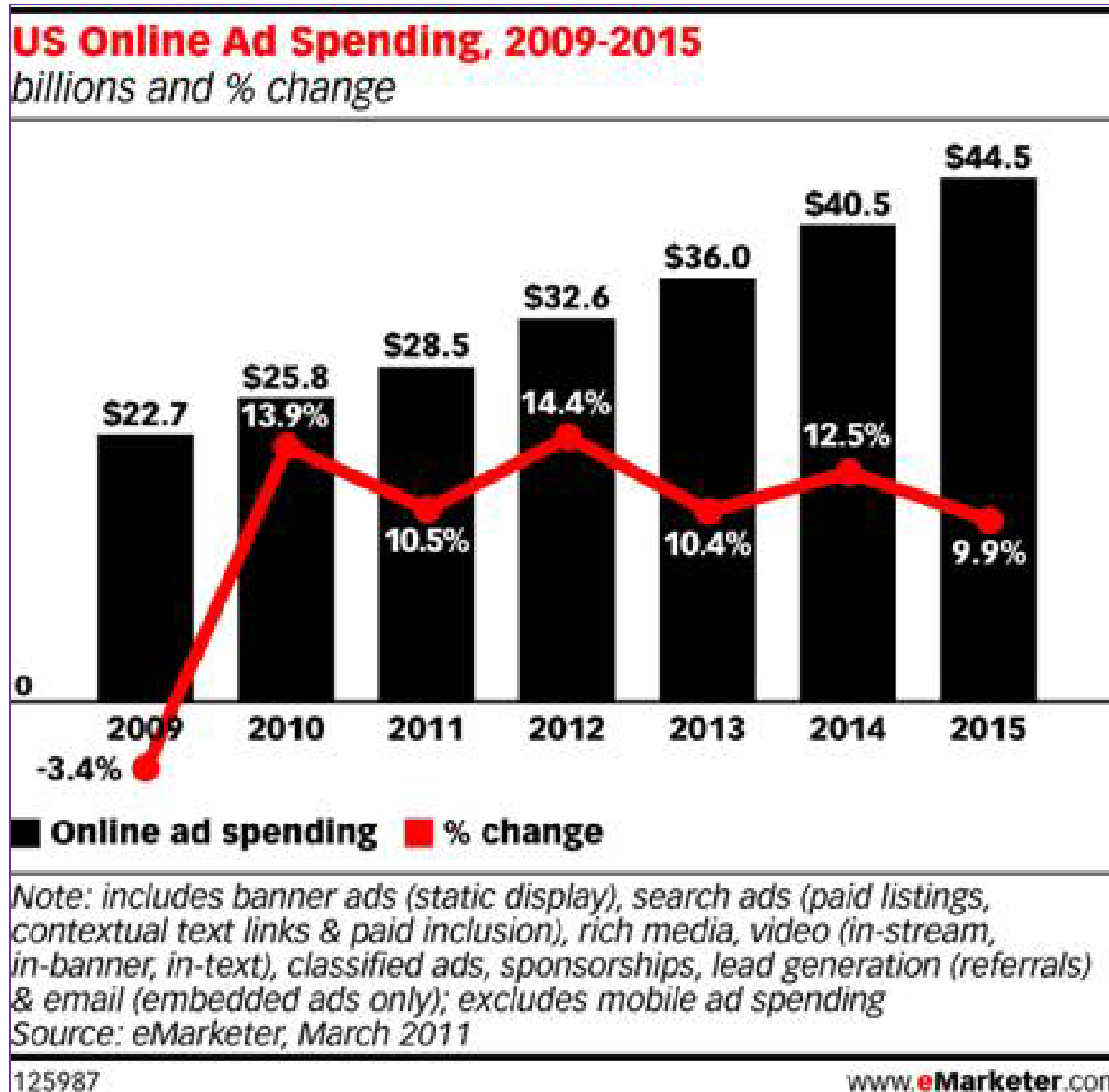
billions and % change



Note: includes directories, internet, magazines, newspapers, outdoor, radio and TV

Source: eMarketer, June 2011

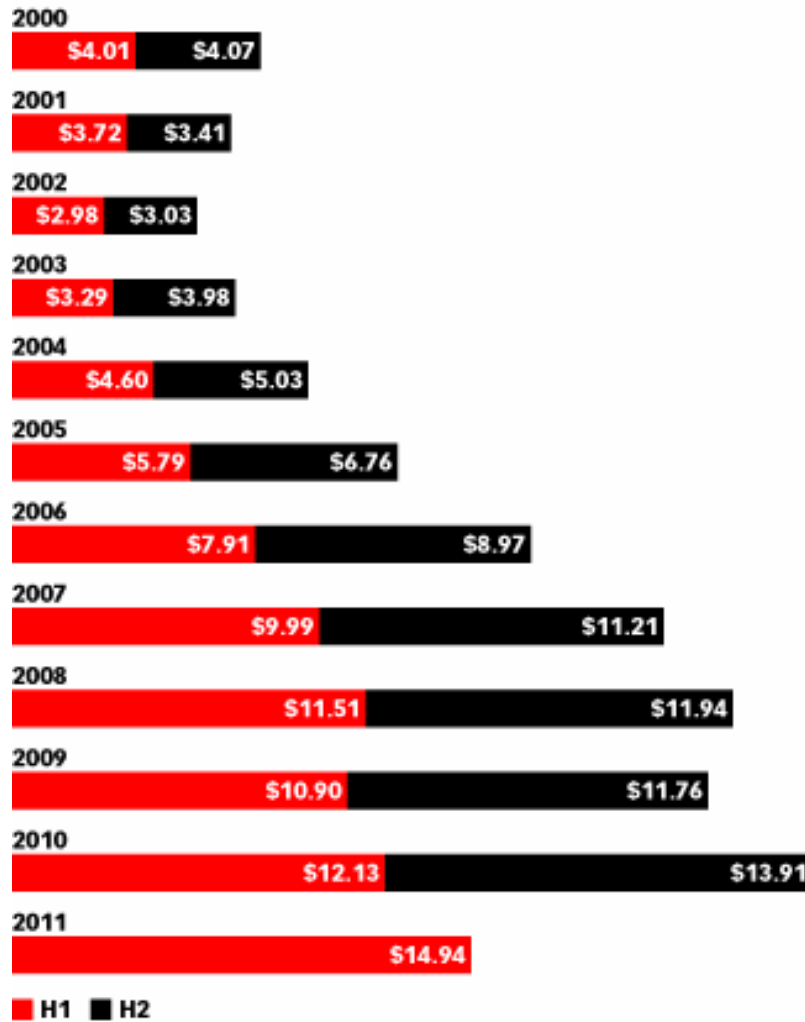
US Online Ad Spending



Hot off the press!

US Online Ad Revenues, H1 2000-H1 2011

billions

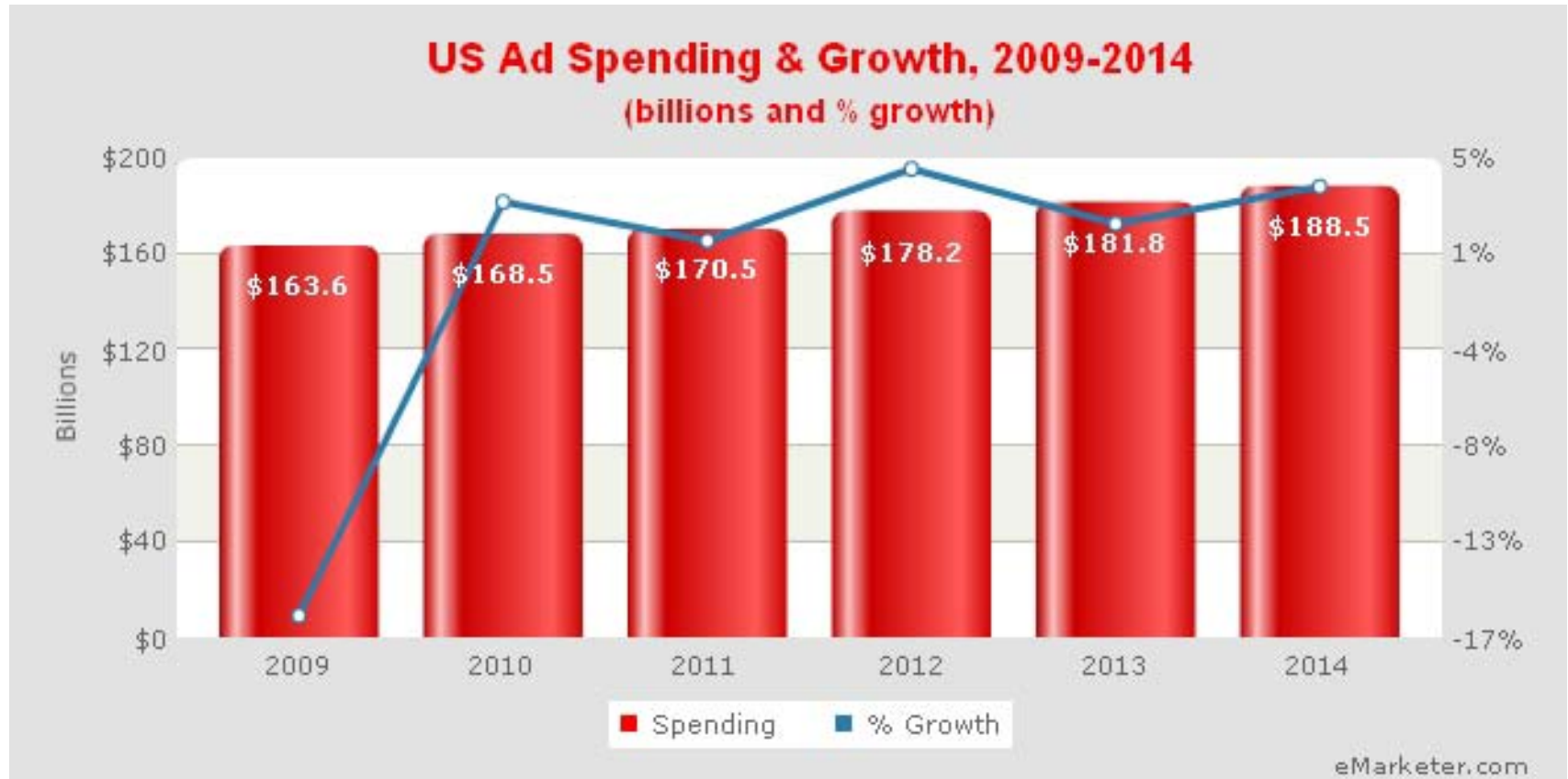


Source: Interactive Advertising Bureau (IAB) and PricewaterhouseCoopers (PwC), "IAB Internet Advertising Revenue Report: 2011 First Six Months Results," Sep 28, 2011

132938

www.eMarketer.com

Total US advertising spend



Spending by format

US Online Ad Revenues, by Format, H1 2010 & H1 2011 *millions and % change*

	H1 2010	H1 2011	% change
Search	\$5,747	\$7,286	26.8%
Display	\$4,356	\$5,535	27.1%
—Banner ads	\$2,744	\$3,414	24.4%
—Rich media	\$743	\$763	2.7%
—Digital video	\$627	\$891	42.1%
—Sponsorship	\$242	\$467	93.0%
Classifieds	\$1,262	\$1,237	-2.0%
Lead generation	\$642	\$805	25.4%
Email	\$120	\$79	-34.2%
Total	\$12,127	\$14,942	23.2%

Source: Interactive Advertising Bureau (IAB) and PricewaterhouseCoopers (PwC), "IAB Internet Advertising Revenue Report: 2011 First Six Months Results," Sep 28, 2011

Who is spending the money?

US Online Ad Revenues, by Major Industry Category, H1 2010 & H1 2011

millions and % change

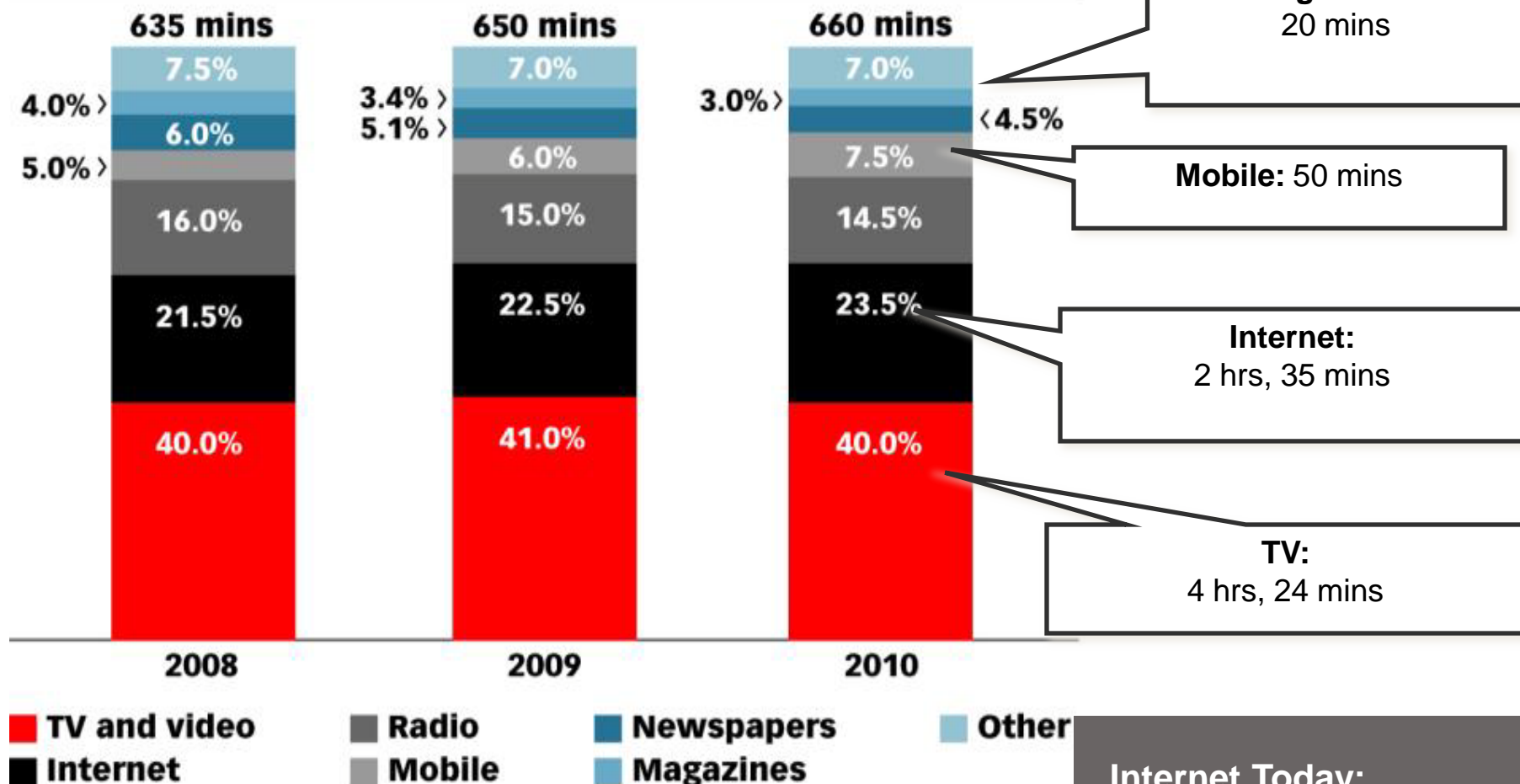
	H1 2010	H1 2011	% change
Retail	\$2,425	\$3,437	41.7%
Telecom	\$1,698	\$2,092	23.2%
Financial services	\$1,455	\$1,942	33.5%
Auto	\$1,334	\$1,644	23.2%
Computing products	\$1,213	\$1,494	23.2%
Consumer packaged goods	\$970	\$897	-7.6%
Leisure travel	\$849	\$1,195	40.8%
Pharma & healthcare	\$606	\$598	-1.4%
Media	\$485	\$598	23.2%
Entertainment	\$485	\$598	23.2%
Other	\$606	\$448	-26.1%
Total	\$12,127	\$14,942	23.2%

Note: industry definitions may have changed over the time period depicted, both within the survey process and definitionally by survey respondents; numbers do not add up to 100% because minor categories are not displayed

Source: Interactive Advertising Bureau (IAB) and PricewaterhouseCoopers (PwC), "IAB Internet Advertising Revenue Report: 2011 First Six Months Results," Sep 28, 2011

Share of Time Spent per Day with Major Media by US Adults, 2008-2010

% of total



Note: time spent with each medium includes all time spent with that medium, regardless of multitasking; for example, 1 hour of multitasking on the internet and watching TV was counted as 1 hour for TV and 1 hour for internet
 Source: eMarketer, Dec 2010

Internet Today:

23.5% of time
 15.3% of dollars
 8.2 p.p. gap

Over Half of Consumers are Simultaneously Online and Watching TV

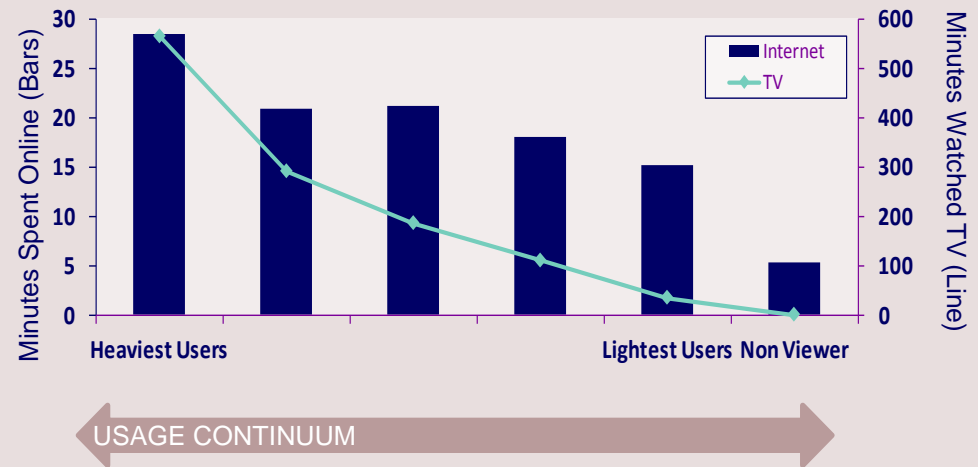
M-F 6p-9p / Adults 25-54¹

57%

Simultaneous
Internet + TV



Media Consumption Level By Usage Quintile²



Dominant players

Net US Online Ad Revenues at Top 5 Ad-Selling Companies, 2009-2012

billions

	2009	2010	2011	2012
Google	\$7.90	\$10.03	\$12.77	\$16.53
Yahoo!	\$3.66	\$3.47	\$3.46	\$3.57
Facebook	\$0.56	\$1.21	\$2.19	\$2.87
Microsoft	\$1.18	\$1.49	\$1.92	\$2.66
AOL	\$0.99	\$0.88	\$0.85	\$0.86
Total top 5	\$14.28	\$17.07	\$21.19	\$26.49
Total internet	\$22.66	\$26.04	\$31.30	\$36.80

Note: net ad revenues after companies pay traffic acquisition costs (TAC) to partner sites

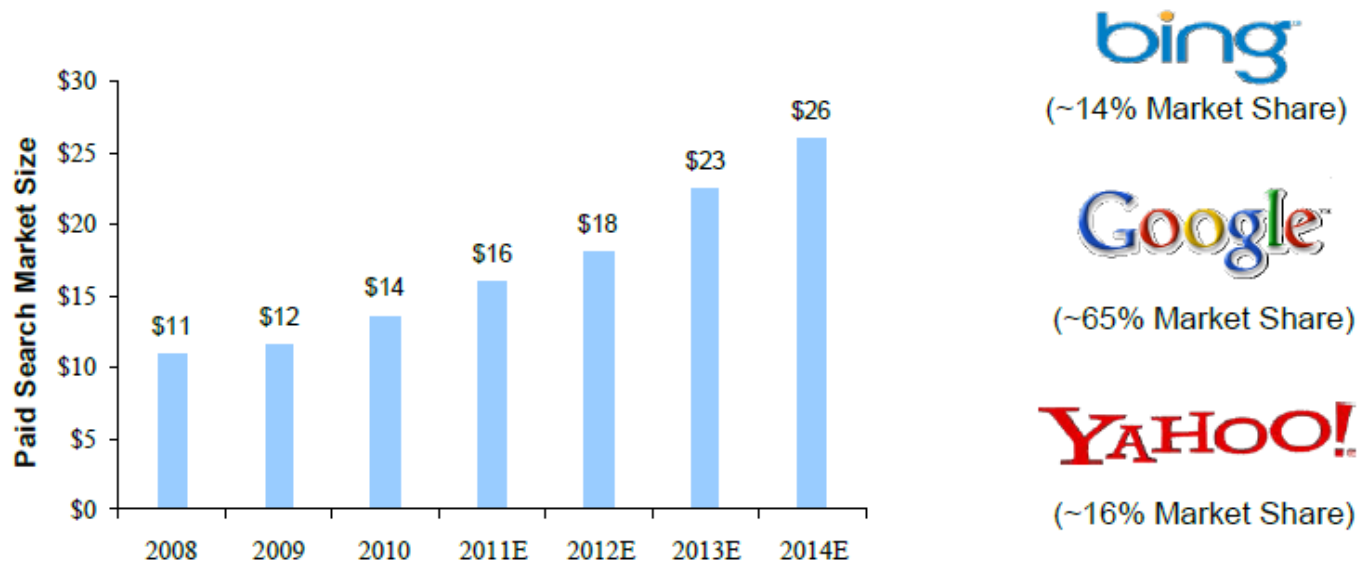
Source: company reports, April-May 2011; eMarketer, Jan & June 2011

Paid search marketshare

- Worldwide Internet users conduct over 200 billion searches each month
- Google drives roughly 46% of overall online ads with annual growth estimated at +20%
- Search should continue to be driven by leaders Google, Yahoo! and Bing
- Search expected to benefit from growth of the mobile Internet

Leading the Search

(\$ in billions)



Source: Baird Equity Research, Gridley estimates and Search Engine Journal

Why spend on-line?

“Economic unrest means more spending for online advertising—even as ad budgets remain tight. Marketers see online advertising as safe because the ads are typically more **measurable**, more **efficient** and more **targetable**, making the internet a more accountable medium than nearly all traditional channels”

-- E-marketeer, December 29, 2010

The Value of Web Advertising

Advertising supports a vast eco-system on the Web

1. Publisher revenue makes macro & micro publishers viable
2. Focused reach & targeting makes niche interests businesses possible
3. Advertiser revenue makes large scale “free” services possible: Facebook, Google, Twitter, Yahoo, ...



- The web would be a lot smaller without advertising → Advertising creates huge direct and indirect value for consumers !

Computational Advertising Landscape

Marketplace basics

- What do advertisers pay?
 - CPM = cost per thousand impressions
 - Typically used for graphical/banner ads (brand advertising)
 - Could be paid in advance → “Guaranteed delivery”
 - CPC = cost per click
 - Typically used for textual ads
 - CPT/CPA = cost per transaction/action a.k.a. referral fees or affiliate fees
 - Typically used for shopping (“buy from our sponsors”), travel, etc.
 - ... but now also used for textual ads (risk mitigation)

The central challenge decomposed

Find the "best match" between a given user in a given context and a suitable advertisement.

1. **Representation** = represent the user, the context, and the ads in an **effective & efficient** way
2. **Definition** = define the mathematical optimization problem to capture the actual marketplace constraints and goals
3. **Solution** = solve the optimization problem in an **effective & efficient** way

Graphical ads

Yahoos: The Y! Sports Team is enhancing our article pages and navigation and we'd like to hear what you think. Please click this link to give us your feedback. Leaving feedback sooner is better so we can consider it as part of our first production launch. Thanks!

Yahoo! | My Yahoo! | Mail | More

Welcome, jayavel.shanmugas... | Sign Out | Help

Web Search

YAHOO! SPORTS

Sports Search

Presented by: PEPSI STUFF

Home | NFL | MLB | NHL | NASCAR | NCAAF | NCAAB | Golf | Tennis | Soccer | MMA | Boxing | More | Video | Shop | Stats.com | Fantasy

PEPSI STUFF amazonmp3

Go to PepsiStuff.com NOW!

NCAAF

INSIDE DISH: ARKANSAS WANTS MALLET TO PLAY IN...

Big Blue heaven

Super Bowl MVP Eli Manning threw a pair of fourth-quarter touchdowns.

Perfection belonged to the Giants on Sunday, as New York scored with 35 seconds left to shock New England 17-14 in Super Bowl XLII. [Recap Tyree's catch of a lifetime](#) | [Peyton pumped](#) | [Wetzel: Winners & losers](#)

More on Super Bowl
[Giants Super Bowl gear](#)
[Super Bowl history](#)

Belichick shows no class

Our Top 5 Super Bowl ads

O-line collapses for Patriots

Bill Belichick couldn't face the music. [MJD blog](#) | [Halftime show](#)

Watch the best Super Bowl ads, according to Jamie Mottram. [Blog](#)

The Giants made life miserable for Tom Brady. [Story](#) | [Clinton crows](#)

Headlines

- Giants end Pats' perfect season 17-14
- Woods rallies, wins Dubai Desert Classic
- Police: Saints DE Grant stabbed in neck
- No Gasol, but Bryant brings Lakers win
- Hansbrough's 22 and 21 lift Heels in OT
- Wallace powers Pistons past Mavs 90-67
- Gordon's 29 lift Indiana by Northwestern
- Holmes beats Lefty in FPR Open playoff

More News | Calendar | My Sports News | Video

ADVERTISEMENT

Go to PepsiStuff.com NOW!

CLICK HERE FOR MORE ON WYCLEF

Scoreboard

NFL | NBA | NCAAB | GOLF | My Teams

Full Scoreboard

Graphical ads – Guaranteed Delivery

- Two types of online graphical advertising
 - Guaranteed delivery (GD)
 - Performance graphical advertising (non-guaranteed delivery, NGD)
- Guaranteed delivery
 - Contract booked based on targeting attributes of an impression: age, income, location,...
 - Each contract has a duration and a desired number of impressions
 - Issues in GD
 - Contract pricing
 - Traffic forecasting
 - Impression allocation to the active contracts
 - ...

Representation

Definition of the
Optimization
Problem

Performance graphical ads

- Graphical ads can also be placed based on performance – CPM/CPC/CPA
- Assume for the moment
Optimization Problem Definition = Max CTR
- Matching approaches:
 1. **Reactive**: explore the placement of a particular ad on different pages; for each page observe achieved CTR; once the CTRs are learned, given page, pick the ad with highest observed CTR
 2. **Predictive**: generate features for the ad using related ads (same advertiser), landing page, or advertiser metadata – predict performance based on page and ad features
 3. **Hybrid**: (1) and (2) are complementary and can be combined

Representation

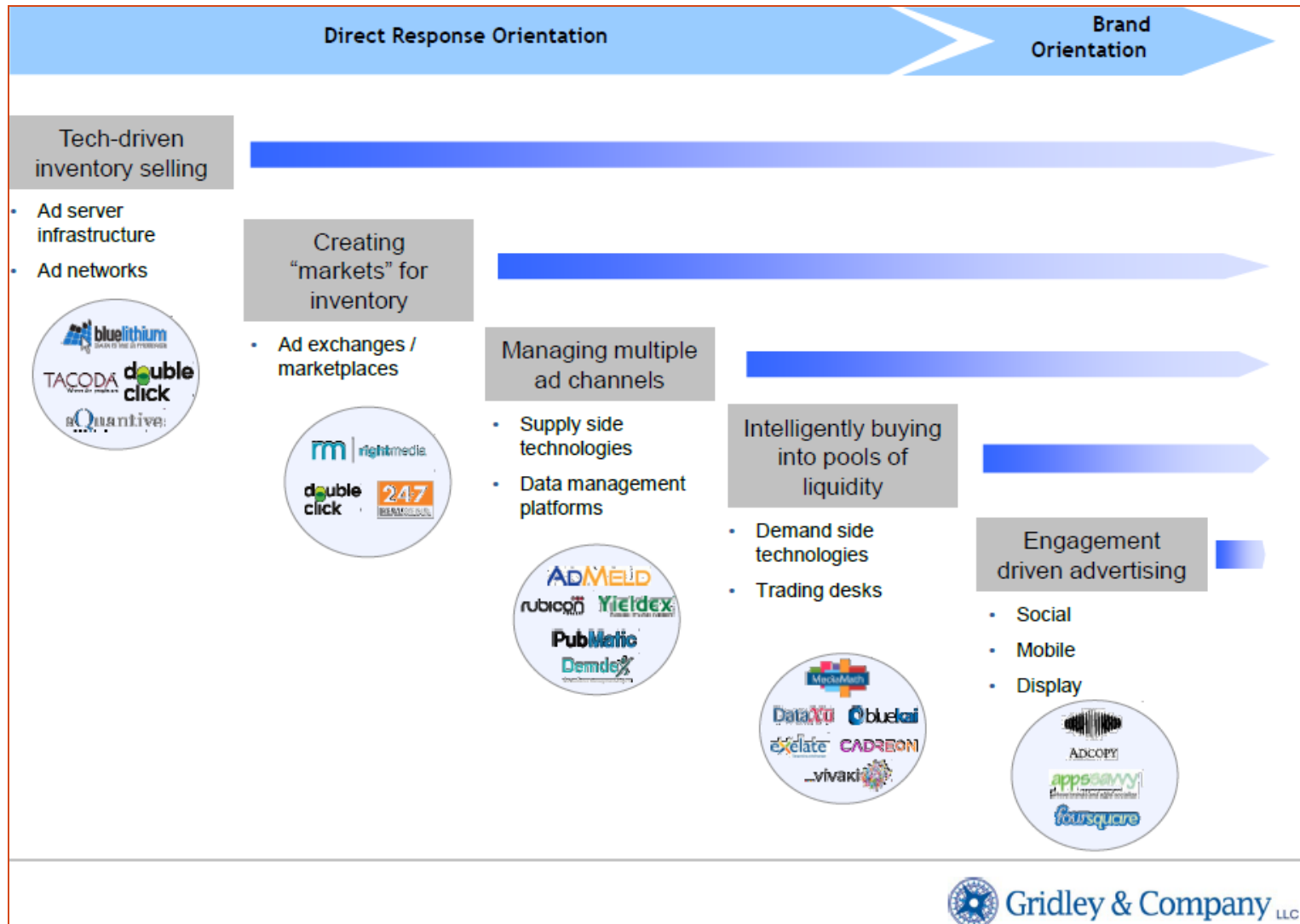
1. **Reactive**: explore the placement of a particular ad on different pages; for each page observe achieved CTR; once the CTRs are learned, given page, pick the ad with highest observed CTR
→ Ads represented by achieved CTR/page + weights
1. **Predictive**: generate features for the ad using related ads (same advertiser), landing page, or advertiser metadata – predict performance based on page and ad features
→ Ads (pages) represented by features of ads (resp. pages) + weights
2. **Hybrid**: (1) and (2) are complementary and can be combined
→ Combined representation

The world of display advertising

DISPLAY LUMAscape



Display ad technology evolution



Textual Ads

Textual ads

1. Ads driven by search keywords = “**sponsored search**”
(a.k.a. “keyword driven ads”, “paid search”, “adwords”, etc)
 - Advertiser chooses a “bid phrase” = query on which to display
 - Can also subscribe to “advanced match” = display me on related queries
 - Needed to achieve volume
 - Huge challenge
2. Ads driven by the content of a web page = “**content match**”
(a.k.a. “context driven ads”, “contextual ads”, “adsense”, etc)

Textual ads are heavily related to Search and IR

Historical view on textual advertising

- Late 1990s Alta Vista tried the Sponsored Search model
 - Rejected by the early search engine users
- Goto.com (acquired later by Overture) develops a search engine for paid ads
 - Users with commercial interest go to this engine
 - At the peak, a billion dollar business
- Google tries the Sponsored Search model again
 - This time a success
- Advertisers cannot get enough volume
 - Content match to provide more impressions

Textual ads anatomy

Search Yahoo for sigir 2010

Hi, **azbroder** | Sign Out | Help

Make Yahoo! your homepage | Mail

Web Images Video Local Shopping News More ▾

YAHOO! Options ▾

Search Pad

SafeSearch - On

600,017 results for **sigir 2010**

Also try: [sigir 2010 workshops](#), [more...](#)

[home \[ACM SIGIR 2010\]](#)
Deadline for the Elsevier **2010** App Challenge extended to July 9th (read more > ... 33rd Annual ACM **SIGIR** Conference. 19-23 July **2010**, Geneva, Switzerland. Cochairs: Stephane ...
www.sigir2010.org - [Cached](#)

[cfp \[ACM SIGIR 2010\]](#)
SIGIR is the major international forum for the presentation of new ... **SIGIR 2010** welcomes contributions related to any aspect of IR theory and foundation, techniques, and ...
www.sigir2010.org/doku.php?id=cfp - [Cached](#)

[ACM SIGIR Special Interest Group on Information Retrieval ...](#)
ACM **SIGIR** addresses issues ranging from theory to user demands in the application of ... Get ready for **SIGIR 2010** in Geneva! **SIGIR** members Ricardo Baeza-Yates ...
www.sigir.org - [Cached](#)

[ACM SIGIR - Upcoming Events](#)
SAPMIA **2010** - Social, Adaptive and Personalized Multimedia ... **SIGIR 2010** (Geneva, Switzerland) - 33rd Annual ACM **SIGIR** Conference on Research & Development on Information ...
www.sigir.org/events/events-upcoming.html - [Cached](#)

Sponsored Results

[Computational Advertising Tutorial](#)
SIGIR 2010 Tutorial in Computational Advertising
research.yahoo.com

[See your message here...](#)

Search Bing for sigir 2010

Web Images Videos Shopping News Maps More | MSN Hotmail

bing

Web Images

sigir 2010

Sign in
Santa Clara, California Settings

RELATED SEARCHES

- SIGIR 2009
- SIGIR Iraq
- SIGIR Grand Jury
- SIGIR Conference
- Drudge
- ACM
- Sigar
- Special Inspector General for Iraq Reconstruction

SEARCH HISTORY

- leonard cohen
- John Cocke
- andrei broder
- Andrei Broder House
- rick han
- See all

ALL RESULTS

1-10 of 164,000 results · [Advanced](#)

[home \[ACM SIGIR 2010\]](#)
Google grants for female computer scientists
[www.sigir2010.org](#) · [Cached page](#)

Keynotes Talks	CfP
ACM SIGIR 2010	Workshops
Local Info	Spc/Pc
Request	Sponsors

[Show more results from www.sigir2010.org](#)

[cfp \[ACM SIGIR 2010\]](#)
PDF version **SIGIR** is the major international forum for the presentation of new research results and for the demonstration of new systems and techniques in the broad field of ...
[sigir2010.org/doku.php?id=cfp](#) · [Cached page](#)

[ACM SIGIR Special Interest Group on Information Retrieval Home Page](#)
more **SIGIR** news... Upcoming **SIGIR** Sponsored Conferences. JCDL 2010 (Brisbane, Australia) - June 21-25, 2010; **SIGIR 2010** (Geneva, Switzerland) - July 18-23, 2010
[www.sigir.org](#) · [Cached page](#)

Sponsored sites


[SIGIR 2010 Tutorial](#)
Information Retrieval Challenges in Computational Advertising (Jul 19)
[research.yahoo.com/tutorials/sigir](#)

[See your message here](#)


Search Google for sigir 2010

[Web](#) [Images](#) [Videos](#) [Maps](#) [News](#) [Shopping](#) [Mail](#) [more ▼](#)

broder@acm.org | [Settings ▼](#) | [Sign out](#)



About 155,000 results (0.24 seconds) [Advanced search](#)

 Everything

All results
Related searches

[home \[ACM SIGIR 2010\]](#) ☆
33rd Annual ACM SIGIR Conference 19-23 July 2010, Geneva, Switzerland ...
Contact: Viper group (Computer Science Dept - CUI) - info@sigir2010.org ...
www.sigir2010.org/ - [Cached](#) - [Similar](#)

Papers	At-a-glance
Time information	Organization
Workshops	Tutorials
Posters and Demos	Registration

[More results from sigir2010.org »](#)

[cfp \[ACM SIGIR 2010\]](#) ☆
Feb 17, 2010 ... **SIGIR 2010** welcomes contributions related to any aspect of IR theory and foundation, techniques, and applications. Relevant topics include ...
www.sigir2010.org/doku.php?id=cfp - [Cached](#) - [Similar](#)

[ACM SIGIR Special Interest Group on Information Retrieval Home Page](#) ☆
Get ready for **SIGIR 2010** in Geneva! SIGIR members Ricardo Baeza-Yates (Yahoo! Research), Jeffrey Dean (Google Inc.), Urs Hoelzle (Google Inc.), David Karger ...
www.sigir.org/ - [Cached](#) - [Similar](#)

Sponsored links
[Tutorial at SIGIR 2010](#)
Information Retrieval Challenges in Computational Advertising
research.yahoo.com/tutorials/sigir
[See your ad here »](#)

Google campaign

Google AdWords

Home Campaigns Opportunities Reporting Billing My account

>> All online campaigns >

Campaign: Campaign #1

Last 7 days

Jun 28, 2010 - Jul 4, 2010

Enabled Budget: \$7.00/day Edit Targeting: Search Edit All devices Edit All languages Edit All countries and territories Edit

Ad groups Settings Ads Keywords Networks

All but deleted Keywords

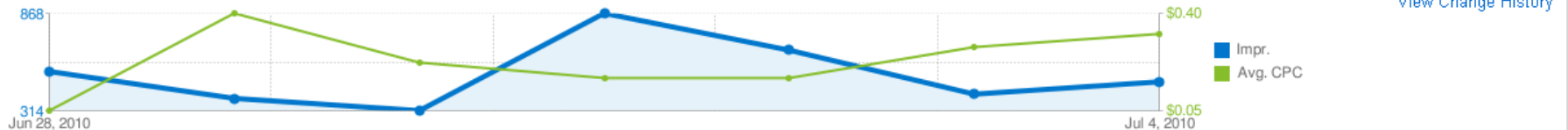
Segment

Filter

Columns



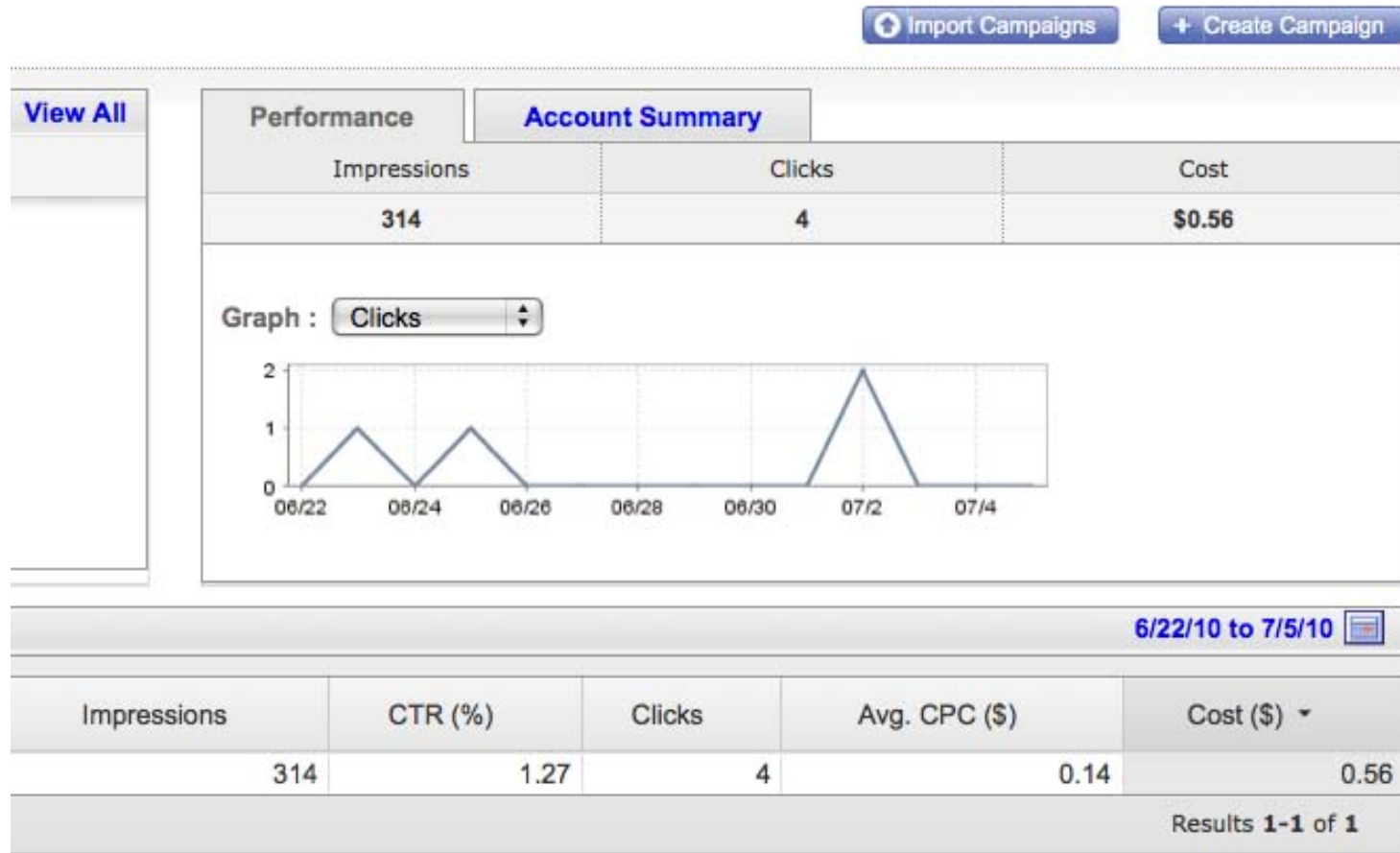
Search



+ Add keywords Edit Change status... See search terms... More actions...

<input type="checkbox"/>	<input checked="" type="radio"/>	Keyword	Ad group	Status	Max. CPC	Clicks	Impr.	CTR	Avg. CPC	Cost	Avg. Pos.
<input type="checkbox"/>	<input checked="" type="radio"/>	sigir	Computational Advertising	Below first page bid First page bid estimate: \$1.00	\$0.40	20	2,449	0.82%	\$0.16	\$3.19	1.2
<input type="checkbox"/>	<input checked="" type="radio"/>	sigir 2010	Computational Advertising	Eligible	\$0.40	3	647	0.46%	\$0.40	\$1.20	1
<input type="checkbox"/>	<input checked="" type="radio"/>	sigir2010	Computational Advertising	Eligible	\$0.40	1	25	4.00%	\$0.16	\$0.16	1
<input type="checkbox"/>	<input checked="" type="radio"/>	text retrieval conference	Computational Advertising	Rarely shown due to low quality score	\$1.00	0	108	0.00%	\$0.00	\$0.00	1.1
<input type="checkbox"/>	<input checked="" type="radio"/>	andrei z broder	Computational Advertising	Eligible	\$0.40	0	6	0.00%	\$0.00	\$0.00	1
<input type="checkbox"/>	<input checked="" type="radio"/>	sigir geneva	Computational Advertising	Eligible	\$1.00	0	3	0.00%	\$0.00	\$0.00	1
<input type="checkbox"/>	<input checked="" type="radio"/>	sigir forum	Computational Advertising	Rarely shown due to low quality score	\$1.00	0	4	0.00%	\$0.00	\$0.00	1
<input type="checkbox"/>	<input checked="" type="radio"/>	evgeniy gabrilovich	Computational Advertising	Eligible	\$0.40	0	12	0.00%	\$0.00	\$0.00	1

Yahoo! Search



Bing

Microsoft adCenter - Windows Internet Explorer

https://adcenter.microsoft.com/campaign/PricingBrowse.aspx?cid=5350128&aid=8721788&cmpld=30373055&adid=316741644

Google

Search

Share

Sidewiki

Bookmarks

Check

Translate

AutoFill

Sign In

Microsoft adCenter

Home

Campaigns

Accounts & Billing

Tools

Reports

Customer: Yahoo > Accounts: Yahoo > Campaigns: SIGIR 2010 Campaign > Ad Group: Ad group #1

Ad group selected: Ad group #1

Settings

Ad group name: Ad group #1 | [Change settings](#)

Ad group status: [Active](#) | [Pause](#)

Default search bid (USD): 0.30 | [Edit](#)

Additional settings

Campaign: SIGIR 2010 Campaign | [Change settings](#)

Current spend (USD): 0.00 of 10.00 | [Edit](#)

Keywords and ads

Keywords

Ads

Performance data for: Entire time | Start: 7/6/2008 | End: 7/6/2010 | [OK](#)

Keywords	Status	Match type	Current bid (max. CPC)	Spend	Clicks	Impressions	CTR (%)	Avg. position	Conversions	Avg. CPC
<input type="checkbox"/> asm sigir	Active	Broad	0.30	0.00	0	0	0.00	0.00	0	0.00
<input type="checkbox"/> asm sigir 2009	Active	Broad	0.30	0.00	0	0	0.00	0.00	0	0.00
<input type="checkbox"/> asm sigir 2010	Active	Broad	0.30	0.00	0	0	0.00	0.00	0	0.00
<input type="checkbox"/> andrei broder	Active	Broad	0.30	0.00	0	2	0.00	1.00	0	0.00
<input type="checkbox"/> andrei z broder	Active	Broad	0.30	0.00	0	0	0.00	0.00	0	0.00
<input type="checkbox"/> computational advertising	Active	Broad	0.30	0.00	0	2	0.00	1.00	0	0.00
<input type="checkbox"/> evgeniy gabrilovich	Active	Broad	0.30	0.00	0	3	0.00	1.00	0	0.00
<input type="checkbox"/> prabhakar raghevan	Active	Broad	0.30	0.00	0	2	0.00	1.00	0	0.00
<input type="checkbox"/> sigir	Active	Broad	0.30	0.00	0	8	0.00	1.00	0	0.00
<input type="checkbox"/> sigir 2009	Active	Broad	0.30	0.00	0	8	0.00	1.00	0	0.00
<input type="checkbox"/> sigir 2010	Active	Broad	0.30	0.00	0	13	0.00	1.00	0	0.00
<input type="checkbox"/> sigir 2011	Active	Broad	0.30	0.00	0	0	0.00	0.00	0	0.00
<input type="checkbox"/> sigir conference	Active	Broad	0.30	0.00	0	1	0.00	1.00	0	0.00
<input type="checkbox"/> sigir forum	Active	Broad	0.30	0.00	0	0	0.00	0.00	0	0.00
<input type="checkbox"/> sigir genesis	Active	Broad	0.30	0.00	0	0	0.00	0.00	0	0.00
<input type="checkbox"/> sigir tutorials	Active	Broad	0.30	0.00	0	0	0.00	0.00	0	0.00
<input type="checkbox"/> sigir10	Active	Broad	0.30	0.00	0	0	0.00	0.00	0	0.00
<input type="checkbox"/> sigir10	Active	Broad	0.30	0.00	0	0	0.00	0.00	0	0.00
<input type="checkbox"/> sigir2010	Active	Broad	0.30	0.00	0	1	0.00	1.00	0	0.00
<input type="checkbox"/> text retrieval conference	Active	Broad	0.30	0.00	0	0	0.00	0.00	0	0.00
<input type="checkbox"/> varja josifovski	Active	Broad	0.30	0.00	0	3	0.00	1.00	0	0.00
Search total	-	-	-	0.00	0	37	0.00	1.00	0	0.00
All sources total	-	-	-	0.00	0	37	0.00	1.00	0	0.00

Tuesday, July 06, 2010 (GMT-08:00) Pacific Time (U.S. & Canada): 7:19am

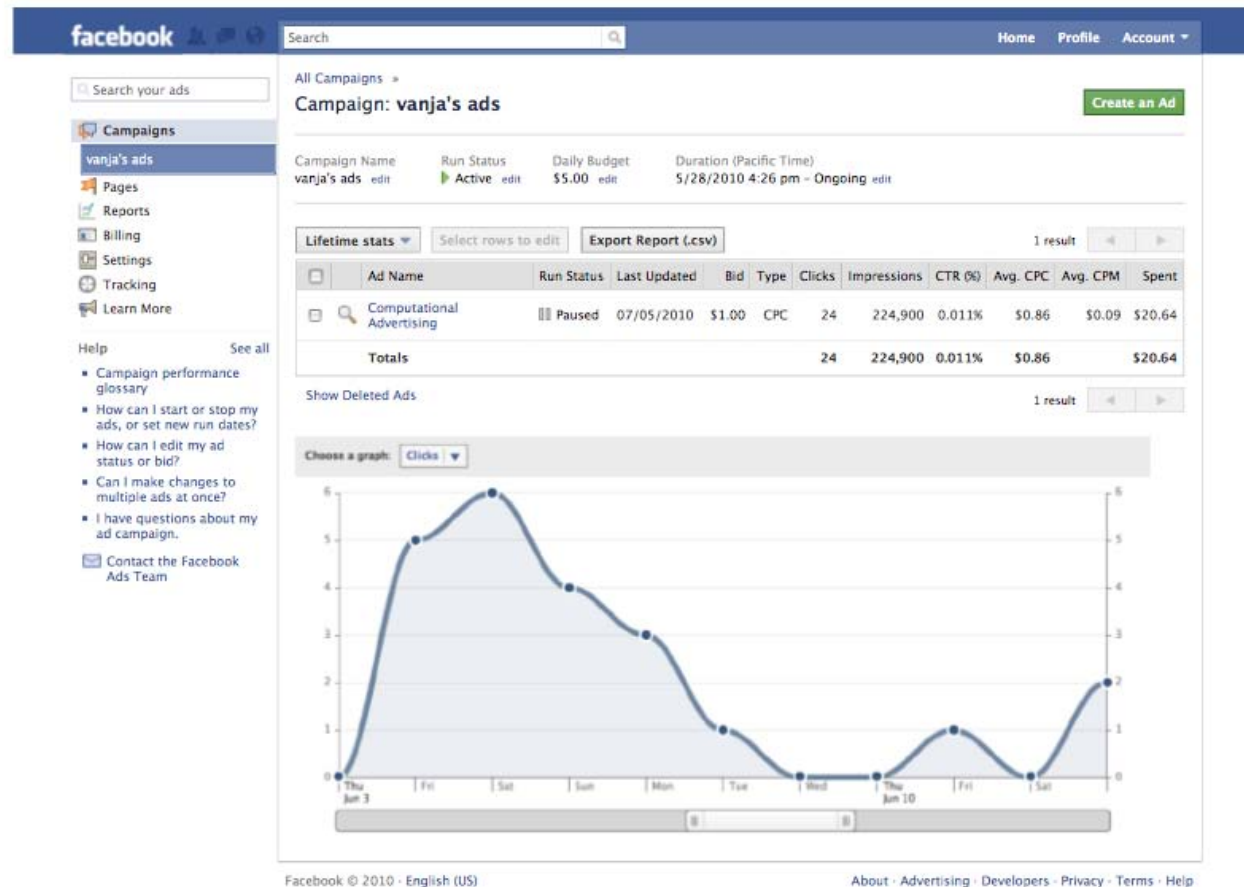
Use incremental bids and targeting to help make your ads more competitive and show them at specific locations and times. You can target at the campaign level on the [Change campaign settings](#) page or target at the ad group level on the [Change ad group settings](#) page.

Done

Local intranet

100%

Facebook



Visible and invisible parts

Bid phrase: sigir 2010
Bid: \$1

Title

{ **Tutorial at SIGIR 2010**

Creative

{ Information Retrieval Challenges in
Computational Advertising

Display URL

{ research.yahoo.com/tutorials/sigir

Landing URL:

http://research.yahoo.com/tutorials/sigir10_compadv/

Destination: the landing page

The image shows a landing page for the SIGIR 2010 Tutorial. At the top, a dark blue header bar contains the text "SIGIR 2010 - Geneva, Switzerland" on the left and "19-23 July, 2010" on the right. Below the header is a large photograph of a cityscape (Geneva) with a prominent fountain in the background. Overlaid on the photograph is the text "SIGIR 2010 Tutorial" in large white letters, and "Information Retrieval Challenges in Computational Advertising" in smaller white letters below it. At the bottom of the page, on a light gray background, is the text "Tutorial on Information Retrieval Challenges in Computational Advertising" followed by "In conjunction with the 33rd ACM SIGIR Conference" and "19-23 July, 2010 - Geneva, Switzerland". Below this is a section titled "OVERVIEW" with a paragraph of text about web advertising.

SIGIR 2010 - Geneva, Switzerland 19-23 July, 2010

SIGIR 2010 Tutorial

Information Retrieval Challenges in Computational Advertising

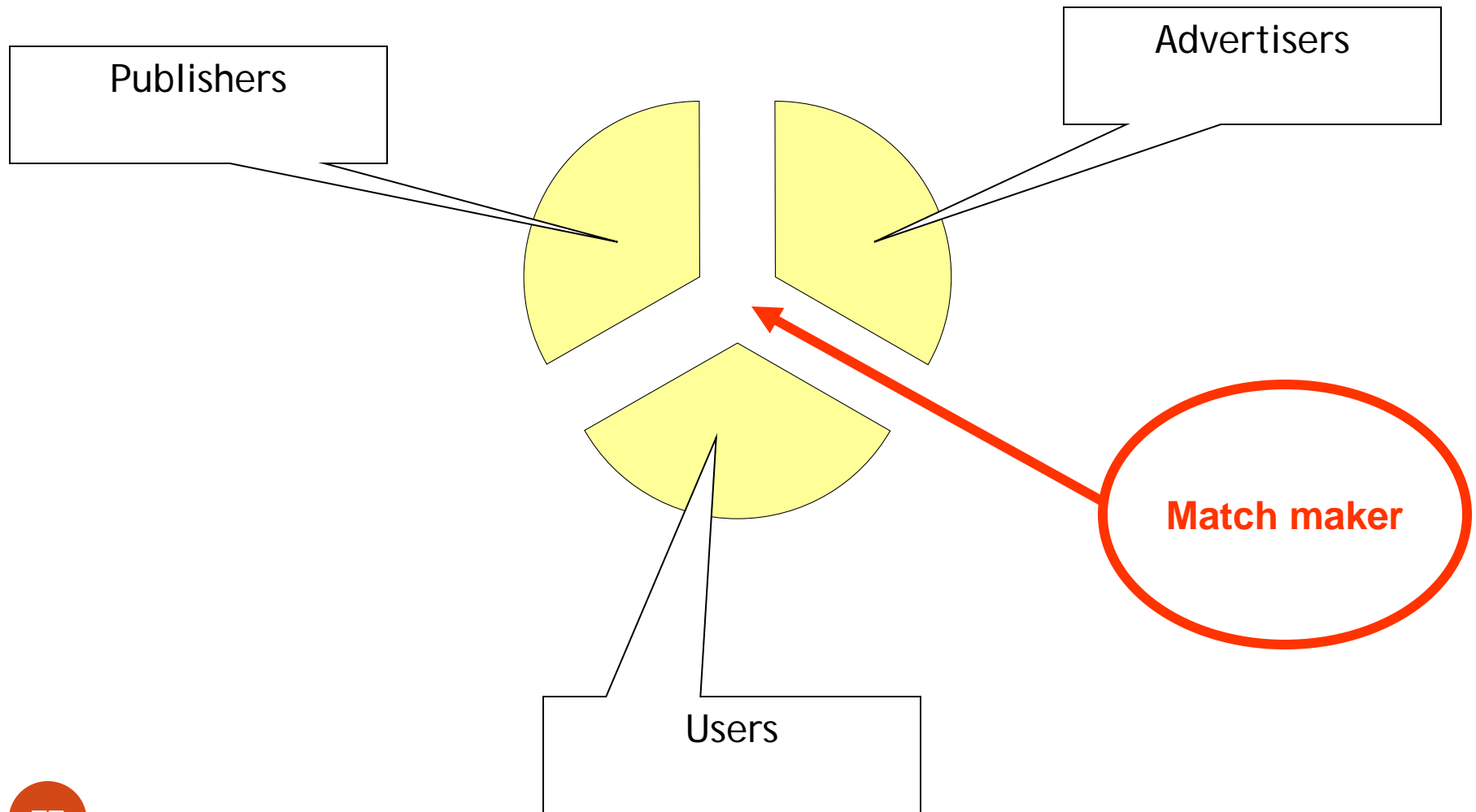
Tutorial on Information Retrieval Challenges in Computational Advertising
In conjunction with the 33rd ACM SIGIR Conference
19-23 July, 2010 - Geneva, Switzerland

OVERVIEW

Web advertising supports a large swath of the Internet ecosystem. It brings revenue to countless publishers that rent space on their pages for advertising: from small mom-and-pop shops to major Internet companies. It also provides valuable traffic to numerous commercial Web sites and has fueled the development of Web search engines. Today, Web advertising is increasingly impacting the world outside the Internet by shaping the

Ad Selection

Participants: Publishers, Advertisers, Users, & “Match maker”



Problem definition: Ad selection objective

- Each participant has its own utility
 1. Advertisers wants ROI and volume
 2. User wants relevance
 3. Publisher wants revenue per impressions/search
 4. Ad network wants revenue and growth
- Ad selection: optimize for a goal that **balances the utilities of the four participants**
- Some tradeoffs are linked with the short term and long term business objectives:
 - Allow for easy adjustments based on periodical changes in objectives

Efficiency requirements: Scale and Cost of Serving



➤ **The Billions:**

- Billions of individual ads in sponsored search and content match
- Billions of unique queries/millions of searches per hour
- Trillions of page impressions (content match and graphical advertising)
- Billions of users



➤ **The Milliseconds:**

- Requests served while the user 'waits': no more than 100ms response time



➤ **The Money:**

- Serving each requests require some CPU amount
- Data usually needs to be in memory
- Per-request cost needs to be lower than the serving cost
- Low CTR make this a challenging problem

Textual ad selection

In the beginning: The database approach

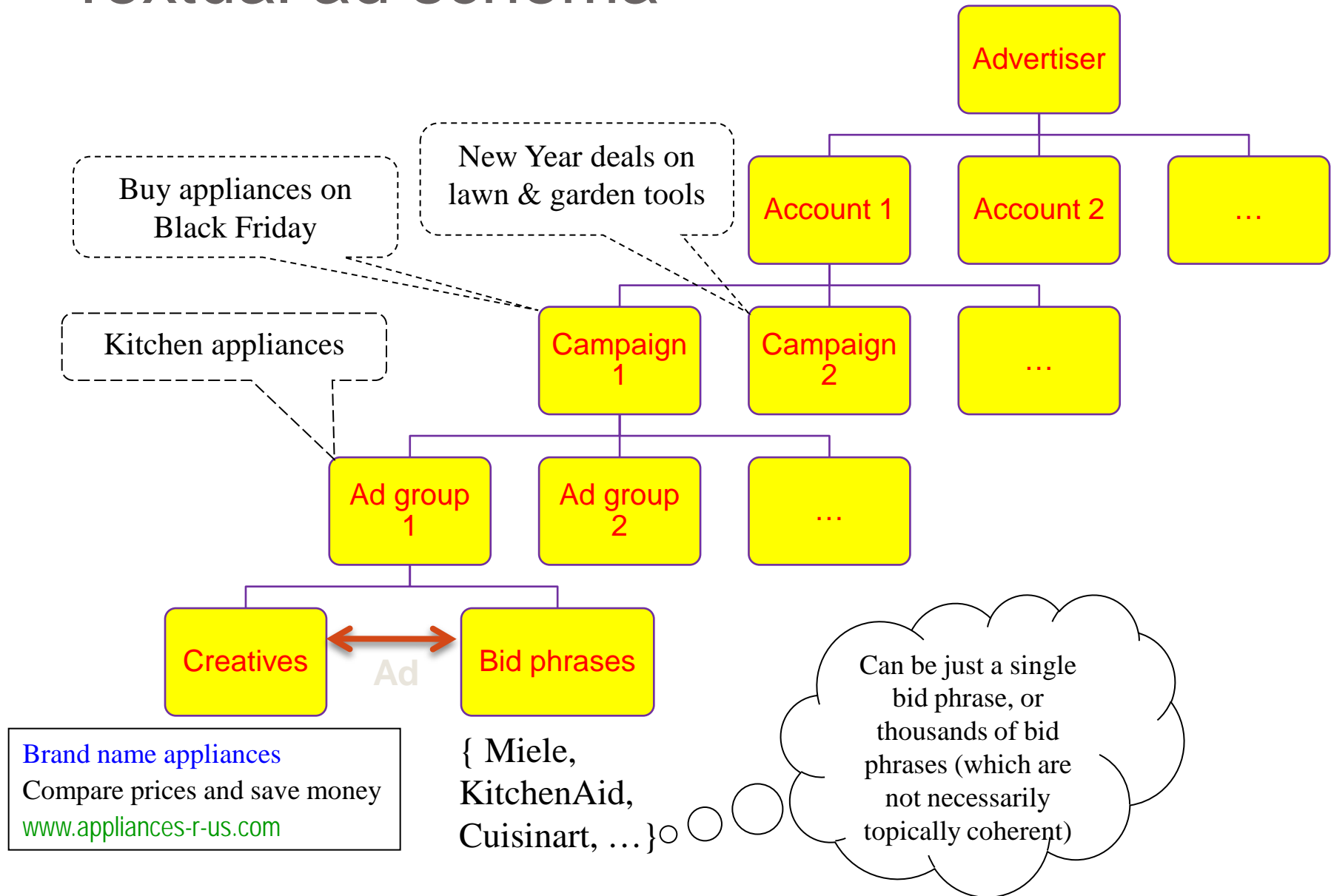
- **Thinking of SS as a data base problem**

- SELECT ads
FROM ad_table
WHERE bid_phrase = query

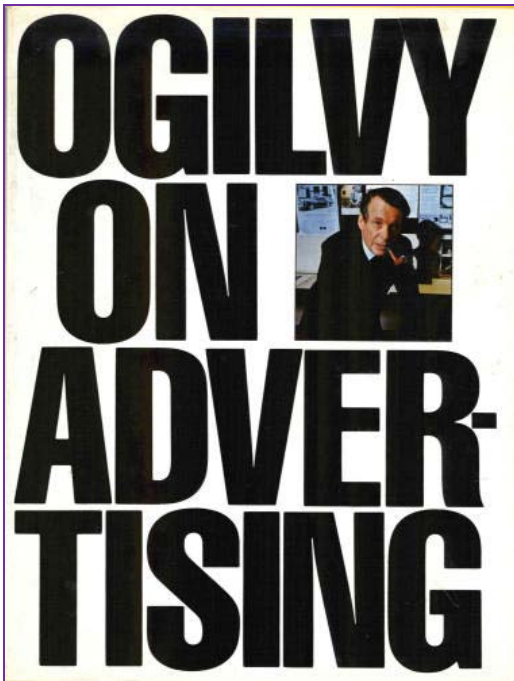
- **Implementation**

- Sponsored search
 - Match the query to the ad bid phrase (some normalization performed)
 - Advertisers cannot bid on all feasible queries (especially in the tail) → Need advanced match
- Advanced match → translate the query into bid phrases
 - Very difficult to capture context, relevance, etc.
 - Pricing is misleading – bid on original phrase has little to do with value of AM
- Content match → bid phrases from pages
 - very difficult to capture context, semantics, relevance, etc.

Textual ad schema



New old concept: advertising as information



- “I do not regard advertising as entertainment or an art form, but as a medium of information....” [David Ogilvy, 1985]
- “Advertising as Information” [Nelson, 1974]
- Irrelevant ads are annoying; relevant ads are interesting
 - Vogue, Skiing, etc are mostly ads and advertorials

Finding the best textual ad is an information retrieval problem with multiple, possible contradictory utility functions

Finding the “best ad” as an Information Retrieval (IR) problem

- **Representation:** Treat the ads as documents in IR
[Ribeiro-Neto et al. SIGIR 2005] [Broder et al. SIGIR2007] [Broder et al. CIKM2008]
- **Optimization/solution:** Retrieve the ads by evaluating the query over the ad corpus

Details

- Analyze the “query” and extract query-features
Query = full context (content, user profile, environment, etc)
- Analyze the documents (= ads) and extract doc-features
- Devise a scoring function = predicates on q-features and d-features + weights
- Build a search engine that produces quickly the ads that maximize the scoring function

Setting the ad retrieval problem

- Ads corpus =
 - Textual Ads: Bid phrase(s) + Title + Creative + URL + Landing Page + ...
 - Graphical Ads: Advertiser supplied meta data + URL + Landing Page + Content of pages with clicks...
- Query features =
 - Search Keywords + Outside Knowledge Expansion + Context features
- Context features (for sponsored search) =
 - Location + User data + Previous searches + ...
- Context features (for content match) =
 - Location + User data + Page topic + Page keywords ...

Summary

Key messages

1. **Computational advertising = A principled way to find the "best match" between a user in a context and a suitable ad.**
2. **Key sub-problems:**
 - I. Representation of user/context/ads
 - II. Definition of optimization problem
 - III. Efficient and effective solution
3. **The financial scale for computational advertising is huge**
 - Small constants matter
 - Expect plenty of further research
4. **Advertising is a form of information.**
 - Adding ads to a context is similar to the integration problem of other types of information
 - Finding the “best ad” is often a type of information retrieval problem with multiple, possibly contradictory utility functions
5. **New application domains and new techniques are emerging every day**
 - Good area for research + new businesses!

Conferences where most of the action happens

- WWW (World Wide Web)
- WSDM (Web Search and Data Mining)
- SIGIR (Information Retrieval)
- CIKM (Information and Knowledge Management)
- EC (Electronic Commerce)
- 4 workshops in 2009: SIGIR, KDD, EC, WINE
- First to sixth Workshop on Sponsored Search Auctions



Questions?

We welcome suggestions about all aspects of the course: msande239-aut0910-staff

Thank you!

broder@yahoo-inc.com
vanjaj@yahoo-inc.com

<http://research.yahoo.com>

This talk is Copyright 2011.
Authors retain all rights, including copyrights and
distribution rights. No publication or further
distribution in full or in part permitted without
explicit written permission