

Big Data Analytics: Applications & Opportunities in On-line Predictive Modeling

Usama Fayyad, Ph.D.

Chairman & CTO

ChoozOn Corporation

Twitter: @usamaf

August 12, 2012

BigMine: BigData Mining Workshop

KDD-2012 – Beijing, China



Outline

- Big Data all around us
- Introduction to Data Mining and Predictive Analytics
- On-line data and facts
- Case studies from multiple verticals:
 1. Yahoo! Big Data
 2. Social Network Data
 3. Case Study from nPario Applications
 4. ChoozOn Big Data from offers
- High-level view: don't forget the basics
- Summary and conclusions

What Matters in the Age of Analytics?

1. Being Able to exploit all the data that is available
 - not just what you've got available
 - what you can acquire and use to enhance your actions
2. Proliferating analytics throughout your organization
 - make every part of your business smarter
3. Driving significant business value
 - embedding analytics into every area of your business can help you drive top line revenues and/or bottom line cost efficiencies

What Organizations Are Struggling With

- Data Strategy - how much data? why data? how does it impact my business?
- Prioritization conducted based on business need, not IT
 - Business justifications for Big Data
 - Demonstrating value of data in impacting the business
 - Looking at specialized stores to reduce TCO
 - File systems for grid computing (Hadoop)
- We do need to stay on top of our basic business ops
 - billing, monitoring, inventory management, etc...
 - Most can be handled by stream processing and traditional BI
- But, a new generation of requirements are becoming a priority for data-driven business
 - Predictive analytics, advanced forecasting, automated detection of events of interest

Why Big Data?

A new term, with associated “Data Scientist” positions:

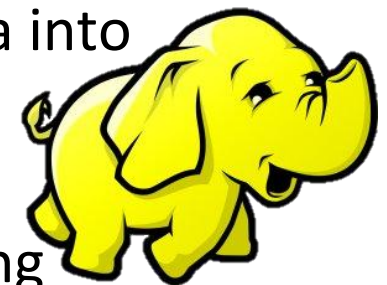
- *Big Data: is a mix of structured, semi-structured, and unstructured data:*
 - Typically breaks barriers for traditional RDB storage
 - Typically breaks limits of indexing by “rows”
 - Typically requires intensive pre-processing before each query to extract “some structure” – usually using Map-Reduce type operations
- Above leads to “messy” situations with no standard recipes or architecture: hence the need for “data scientists”
 - conduct “Data Expeditions”
 - Discovery and learning on the spot

What Makes Data “Big Data”?

- **Big Data is Characterized by the 3-V's:**
 - **Volume:** *larger than “normal” – challenging to load/process*
 - Expensive to do ETL
 - Expensive to figure out how to index and retrieve
 - Multiple dimensions that are “key”
 - **Velocity:** *Rate of arrival poses real-time constraints on what are typically “batch ETL” operations*
 - If you fall behind catching up is extremely expensive (replicate very expensive systems)
 - Must keep up with rate and service queries on-the-fly
 - **Variety:** *Mix of data types and varying degrees of structure*
 - Non-standard schema
 - Lots of BLOB's and CLOB's
 - DB queries don't know what to do with semi-structured and unstructured data.

The Distinction between “Data” and “Big Data” is fast disappearing

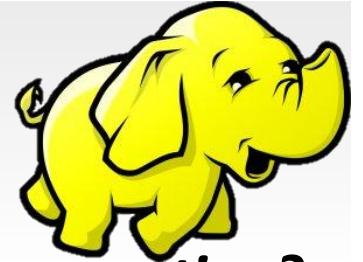
- Most real data sets nowadays come with a serious mix of semi-structured and unstructured components:
 - Images
 - Video
 - Text descriptions and news, blogs, etc...
 - User and customer commentary
 - Reactions on social media: e.g. Twitter is a mix of data anyway
- Using standard transforms, entity extraction, and new generation tools to transform unstructured raw data into semi-structured analyzable data
- Hadoop vs. Not Hadoop - when to use what kind of techniques requiring Map-Reduce and grid computing



Text Data: The Big Driver

- While we speak of “big data” and the “Variety” in 3-V’s
- **Reality:** biggest driver of growth of Big Data has been text data
- In fact Map-Reduce became popularized by Google to address the problem of processing large amounts of text data:
 - Indexing a full copy of the web
 - Frequent re-indexing
 - Many operations with each being a simple operation but done at large scale
- Most work on analysis of “images” and “video” data has really been reduced to analysis of surrounding text

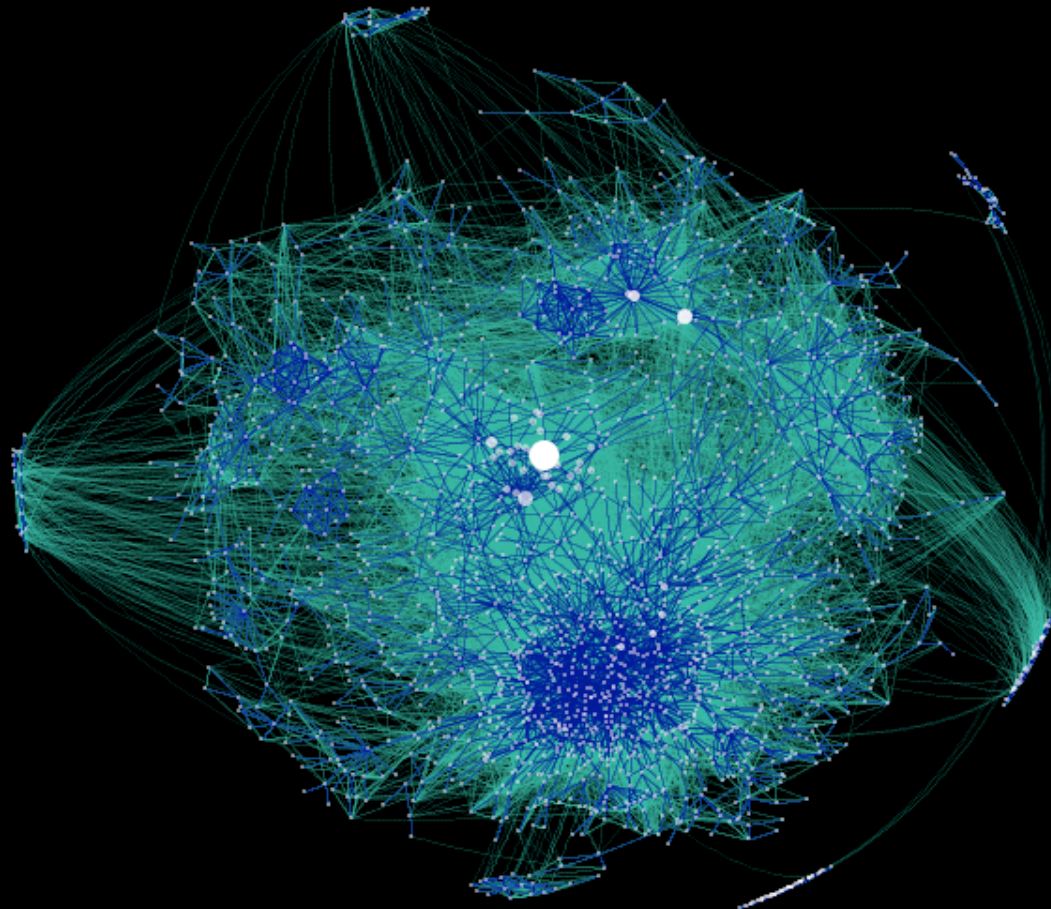
To Hadoop or not to Hadoop?



when to use techniques requiring Map-Reduce and grid computing?

- Typically organizations try to use Map-Reduce for everything to do with Big Data
 - This is actually very inefficient and often irrational
 - Certain operations require specialized storage
 - Updating segment memberships over large numbers of users
 - Defining new segments on user or usage data
- Map-Reduce is useful when a very simple operation is to be applied on a large body of unstructured data
 - Typically this is during entity and attribute extraction
 - Still need Big Data analysis post Hadoop
- Map-Reduce is not efficient or effective for tasks involving deeper statistical modeling
 - good for gathering counts and simple (sufficient) statistics
 - E.g. how many times a keyword occurs, quick aggregation of simple facts in unstructured data, estimates of variances, density, etc...
 - Mostly pre-processing for Data Mining

This is data of people blogging and tweeting about your products on the internet? How do you react to this?

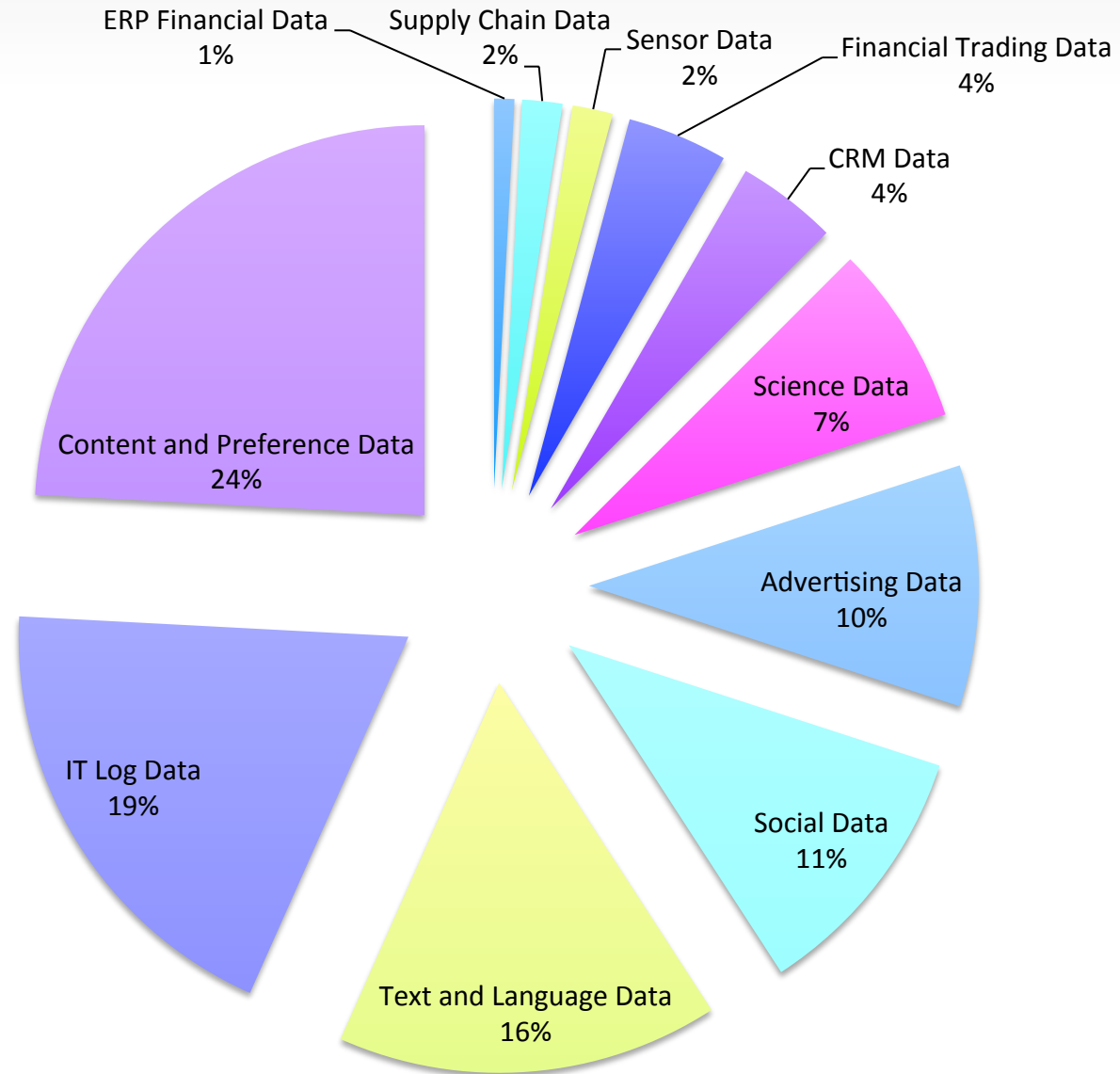


Can you analyze data from these companies and integrate the results with your strategy?







How about in real time and then change your business on the fly?

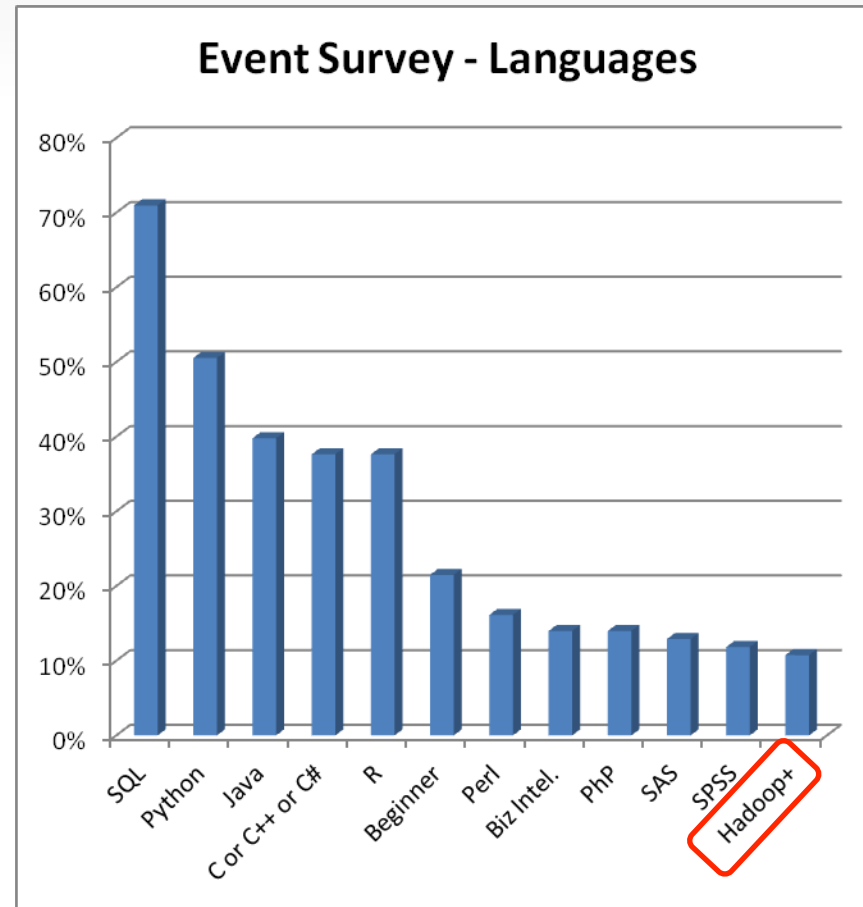
Hadoop Use Cases by Data Type



Big Data Applications and Uses

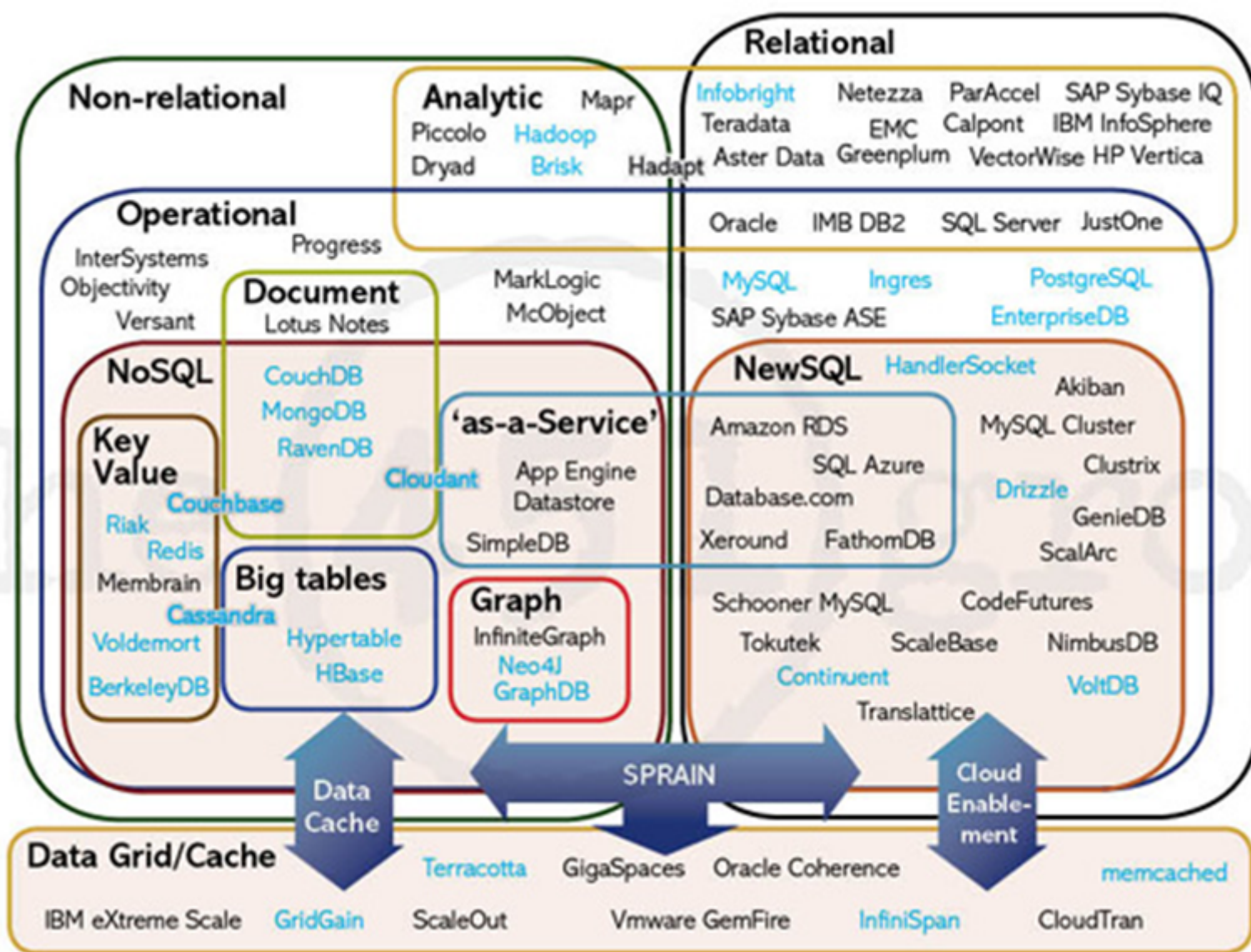
	<p>IT Log & Security Forensics & Analytics</p>	<p>100% Capture</p> <p>Find New Signal Predict Events</p>
	<p>Automated Device Data Analytics</p>	<p>Product Planning</p> <p>Failure Analysis</p> <p>Proactive Fixes</p>
	<p>Advertising Analytics</p>	<p>Recommendation</p> <p>Segmentation</p> <p>Social Media</p>
	<p>Big Data Warehouse Analytics</p>	<p>Hadoop + MPP + EDW</p> <p>Cost Reduction Ad Hoc Insight</p> <p>Predictive Analytics</p>

Analysis & Programming Software



disco
massive data - minimal code

RevolutionAnalytics / RHadoop



From Basic Dashboards to Advanced Analytics

- Data Reduction to get
 - Advanced views oriented by customer or product
 - Segmentation
 - Pattern analysis and summaries
- Predictive Analytics
 - Data Mining
 - Statistical analysis
 - Optimization of processes and spend

The **same** analytics technique apply **across many** industries: fraud detection is fraud detection, is fraud detection

What is Data Mining?

Finding *interesting structure* in data

- *Structure*: refers to statistical patterns, predictive models, hidden relationships
- *Interesting*: Accurate predictions, associated with new revenue potential, associated with cost savings, enables optimization

- Examples of tasks addressed by Data Mining
 - Predictive Modeling (classification, regression)
 - Segmentation (Data Clustering)
 - Affinity (Summarization)
 - relations between fields, associations, visualization

Data Mining and Databases

Many interesting analysis queries are difficult to state precisely

- Examples:
 - **which records represent fraudulent transactions?**
 - **which households are likely to prefer a Ford over a Toyota?**
 - **Who is a good credit risk in my customer DB?**
 - **Why are these automobiles in need of unusual repairs?**
- Yet database contains the information
 - **good/bad customer, profitability**
 - **did/didn't respond to mailout/survey/campaign/...**
 - **automobile repair and warranty records**

Many Business Uses

Analytic technique	Uses in business
Marketing and sales	Identify potential customers; establish the effectiveness of a campaign
Understanding customer behavior	model churn, affinities, propensities, ...
Web analytics & metrics	model user preferences from data, collaborative filtering, targeting, etc.
Fraud detection	Identify fraudulent transactions
Credit scoring	Establish credit worthiness of a customer requesting a loan
Manufacturing process analysis	Identify the causes of manufacturing problems
Portfolio trading	optimize a portfolio of financial instruments by maximizing returns & minimizing risks
Healthcare Application	fraud detection, cost optimization, detection of events like epidemics, etc...
Insurance	fraudulent claim detection, risk assessment
Security and Surveillance	intrusion detection, sensor data analysis, remote sensing, object/person detection, link analysis, etc...

So this Internet thing is going to
be big!

Big opportunity, Big Data, Big
Challenges!

Stats about on-line usage

- How many people are on-line today?
 - 2.1 Billion (per Comscore estimates)
 - 30% of world Population
- How much time is spent on-line per month by the whole world?
 - 4M person-years per month
- How many hours per month per Internet User?
 - 16 hours (global average)
 - 32 hours (U.S. Average)

**Sources: Feb.2012 - from Go-Gulf.com compiled from Comscoredatamine.com, Nielsen.com, thisDigitalLife.com, PewInternet.org*

How Are Users Distributed Geographically



78.90 %	North America 272,066,000
37.74 %	Latin America 215,939,400
64.50 %	Europe 476,213,935
22.14 %	Asia 922,329,554
11.60 %	Africa 118,609,620
58.19 %	Oceania 21,293,830

	85.00% United Kingdom 52,996,180
	81.85% Germany 66,825,986
	80.10% France 51,879,480
	80.00% Japan 102,063,316
	79.00% United States 243,542,822



	43.00% Russia 59,937,788
	40.65% Brazil 81,748,504
	34.30% China 456,238,464
	28.43% Nigeria 45,944,229
	7.50% India 87,983,101

*Sources: Feb.2012 - from Go-Gulf.com compiled from Comscoredatamine.com, Nielsen.com, thisDigitalLife.com, PewInternet.org

How Do People Spend Their On-line Time?

- On-line Shopping? **5%**
- Searches? **21%**
- Email/Communication? **19%**
- Reading Content? **20%**
- Social Networking? **22%**
- Multimedia Sites? **13%**

Most Popular Activities On-Line?



*Sources: Feb.2012 - from Go-Gulf.com compiled from Comscore.datamine.com, Nielsen.com, thisDigitalLife.com, PewInternet.org

Top 10 Sites Visited?

- **Google:** 153.4M visitors/month
 - each spending 1h 47mins
- **Facebook:** 137.6M visitors
 - each spending 7h 50mins

	Unique Visitors Per Month	Time Spent Per Person Per Month in hh:mm:ss
	130,121,000	2:12:08
	115,890,000	1:43:45
	106,692,000	1:41:27
	83,691,000	0:45:05
	74,633,000	2:52:52
	62,097,000	0:18:03
	61,608,000	1:06:15
	60,552,000	0:12:27

Interesting Events

- **Google:** How many queries per day?
 - More than 1 Billion
- **Twitter:** How many Tweets/day?
 - More than 250M
- **Facebook:** Updates per day?
 - More than 800M
- **YouTube:** Views/day
 - 4 Billion views
 - 60 hours of video uploaded every minute!
- **Social Networks:** users who have used sites for spying on their partners?
 - **56%**

**Sources: Feb.2012 - from Go-Gulf.com compiled from Comscore.datamine.com, Nielsen.com, thisDigitalLife.com, PewInternet.org*

Interesting Events

- Country with Highest online friends?
 - Brazil
 - 481 friends per user
 - Japan has least at 29
- Country with maximum time spent shopping on-line??
 - China: 5 hours/week

**Sources: Feb.2012 - from Go-Gulf.com compiled from Comscoredatamine.com, Nielsen.com, thisDigitalLife.com, PewInternet.org*

So Internet is a big place with lots happening?

Do we understand what each individual is trying to achieve?

- *What is user intent?*
- *Critical in monetization, advertising, etc...*

Do we understand what a community's sentiment is?

- What is the emotion?
- Is it negative or positive?
- What is the health of my brand online?

Do we understand context and content?

- What are appropriate ads?
- Is it Ok to associate my brand with this content?
- Is content sad?, happy?, serious?, informative?



Yahoo! – One of Largest Destinations on the Web



80% of the U.S. Internet population uses Yahoo!
– Over 600 million users per month globally!

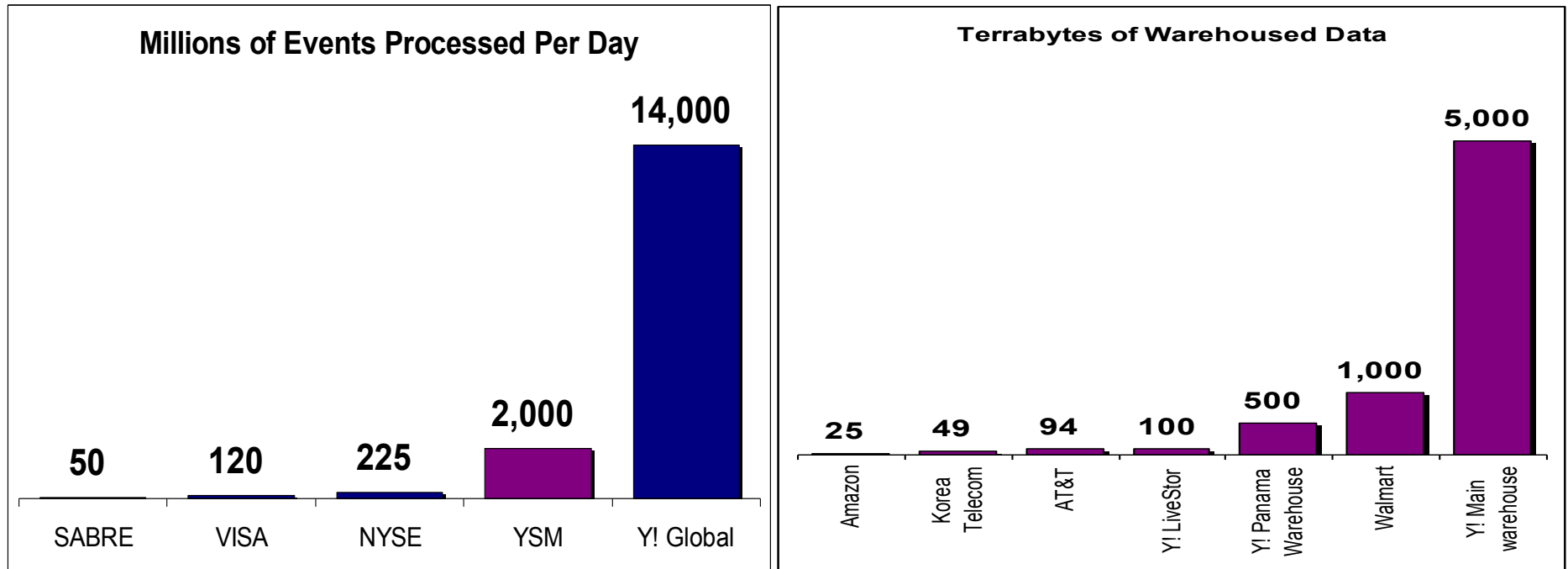
- Global network of content, commerce, media, search and access products
- 100+ properties including mail, TV, news, shopping, finance, autos, travel, games, movies, health, etc.
- **25+ terabytes of data collected each day**
 - Representing 1000's of cataloged consumer behaviors

Data is used to develop content, consumer, category and campaign insights for our key content partners and large advertisers

More people visited Yahoo! in the past month than:

- Use coupons
- Vote
- Recycle
- Exercise regularly
- Have children living at home
- Wear sunscreen regularly

Yahoo! Big Data – A league of its own...



GRAND CHALLENGE PROBLEMS OF DATA PROCESSING

TRAVEL, CREDIT CARD PROCESSING, STOCK EXCHANGE, RETAIL, **INTERNET**

Y! Data Challenge Exceeds others by 2 orders of magnitude

Behavioral Targeting (BT)

Content

YAHOO! FINANCE
Monday, June 4, 2007, 12:17
HOME INVESTING NEWS & OPINION PERSONAL FINANCE MY PORTFOLIOS
Enter Symbol(s) GET QUOTES Symbol Lookup Finance Search
Market Summary
Dow 13,642.97 **▲ 24.14 (0.18%)**
Nasdaq 2,610.51 **▲ 3.41 (0.13%)**
S&P 500 1,535.75 **▲ 6.59 (0.04%)**
10-Yr Bond 4.9310% **▲ 0.0250**
NYSE Volume 768,163,000
Nasdaq Volume 872,000,000

YAHOO! Web Images Video Local Shopping more
Search: Home Loan **Search**
My Yahoo! My Mail Page Options

YAHOO!
BT

YAHOO! SEARCH Home Loan Search Advanced Search
Search Results
Also try: [home loan calculator](#), [countrywide home loan](#), [home savings and loan](#) More...

- Countrywide® Home Loans**
[www.countrywide.com](#) - Less-than-perfect credit expertise from Countrywide®, 4 of 5 approved.
- Home Loans - LendingTree**
[www.lendingtree.com](#) - Refinance \$200,000 for \$667/Month. Refinance Offers, Quality Online.
- Home Loan**
[www.quickenloans.com](#) - Get a \$200,000 Home Loan for Only \$685 / Month - Apply Online Today.
- Home Loans - Ditech®**
[www.ditech.com](#) - Fast & Easy Home Loan Process from Ditech. No Hidden Fees - Apply Now.

Yahoo! Report bad results or ads. Bucket test: None [GSP?]

Search Clicks

Yahoo! My Yahoo! Mail Help
YAHOO! FINANCE Welcome, dang_97 [Sign Out, My Account]
Web Finance Pictures Video Audio SEARCH
HOME INVESTING NEWS & OPINION PERSONAL FINANCE MY PORTFOLIOS
Enter Symbol(s) GET QUOTES Symbol Lookup Finance Search
Loans Autos Education Home Personal & Small Biz
Banking & Budgeting Career & Work College & Education Family & Home Insurance Loans Real Estate Retirement

Home

Mortgage Brokers: Are They Your Friends or Foes?
Borrowers often see mortgage brokers as allies, but many brokers don't put customers' interests first. Here's what you should know... read more

HOW-TO GUIDES

- Mortgage Basics
- Financing a New Car
- Home Refinancing Basics
- College Financial Aid Programs
- View more how-to guides

ADVERTISMENT

Sponsored by: **Countrywide® HOME LOANS**
Call for a No Cost Refi**
America's #1 Home Loan Lender **1-800-687-9414**

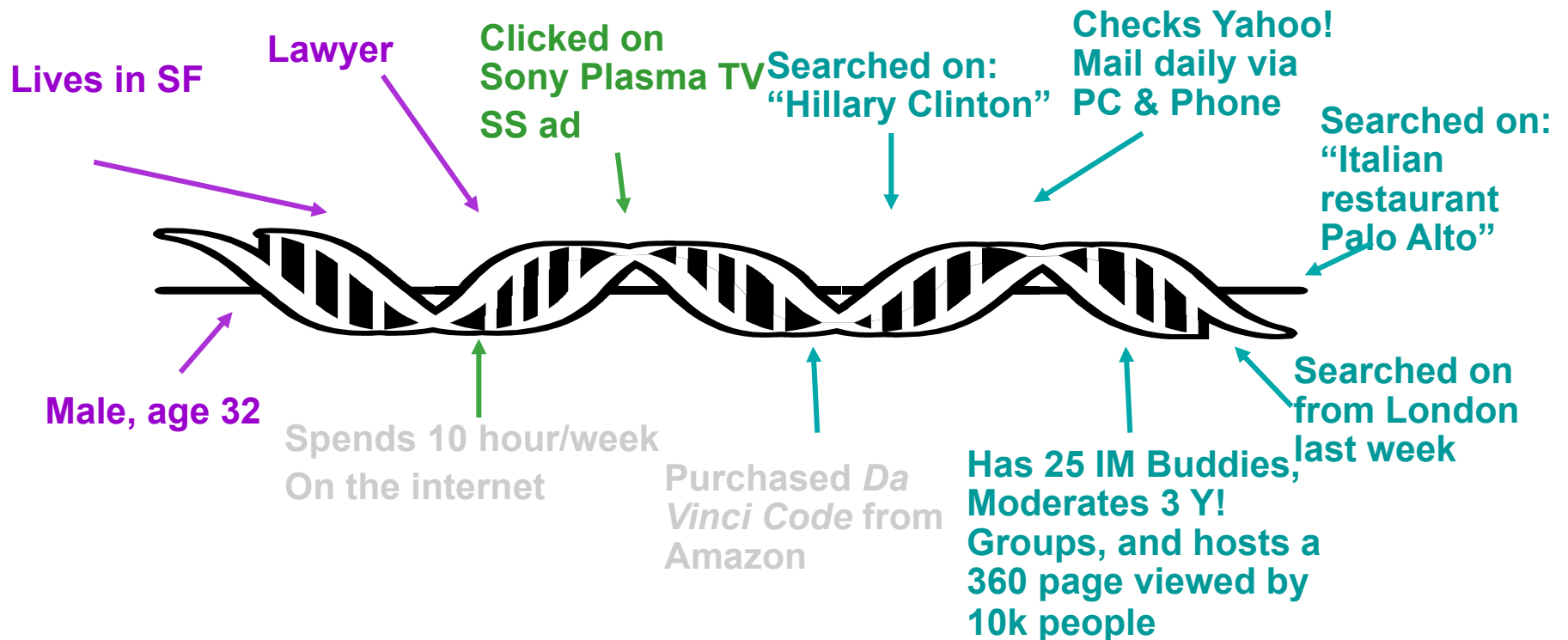
- Refinance
- Consolidate Debt
- Cash Out Refinance
- Loans for Self-Employed Borrowers
- Can I lock in a low rate now?
- How much cash can I get?
- Home Equity Loans
- Home Improvement
- Bad Credit Loans
- How much can I save each month?
- Flexible Payment Options
- Jumbo Loans

AD Clicks

Targeting your ads to consumers whose recent behaviors online indicate that your product category is relevant to them

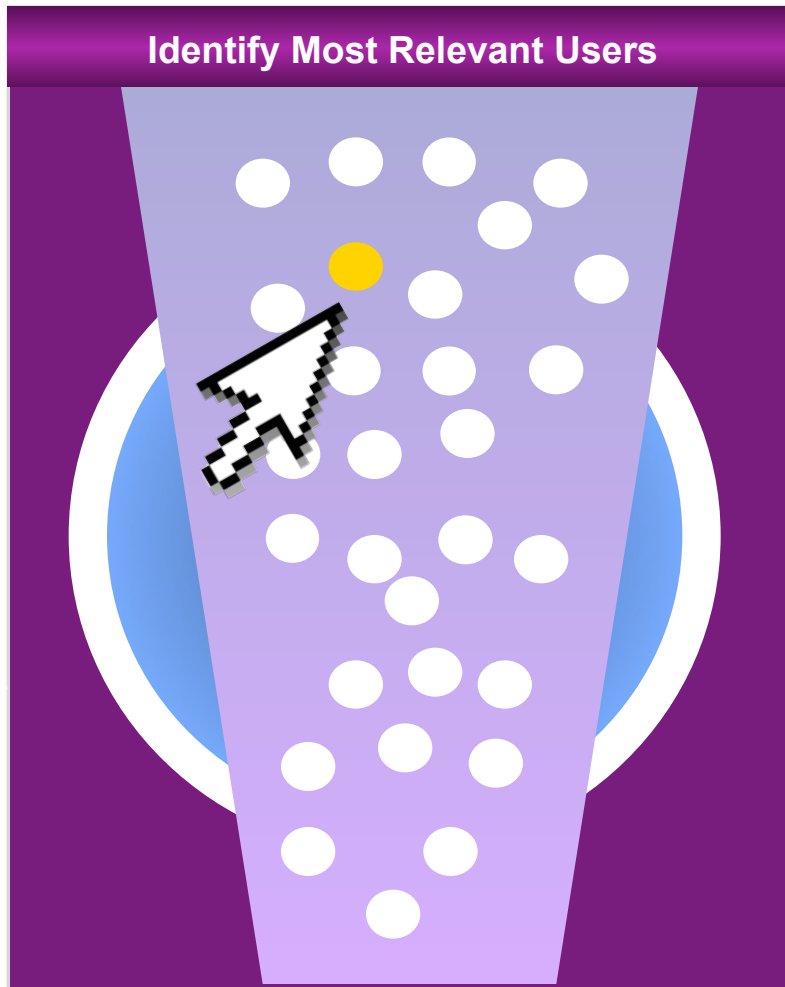
Yahoo! User DNA

■ Registration ■ Campaign ■ Behavior ■ Unknown



- *On a per consumer basis: maintain a behavioral/interests profile and profitability (user value and LTV) metrics*

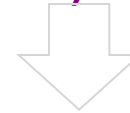
How it works | Network + Interests + *Modelling*



Analyze predictive patterns for purchase cycles in over 100 product categories



In each category, build models to describe behaviour most likely to lead to an ad response (i.e. click).



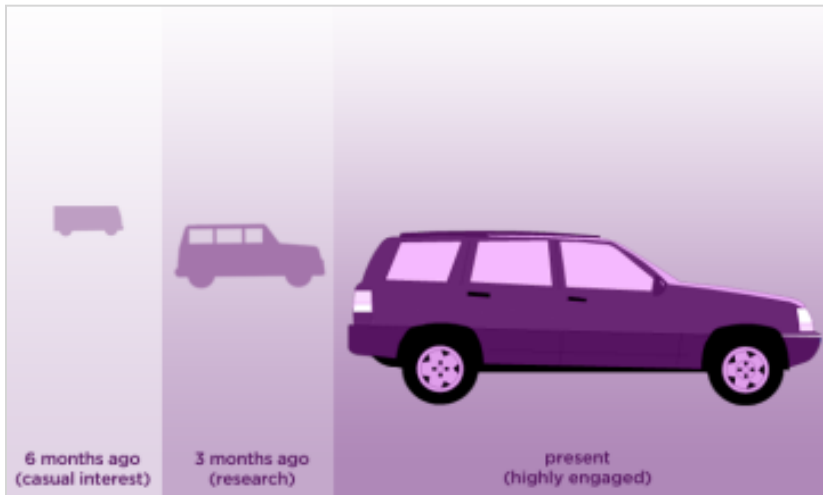
Score each user for fit with every category...daily.



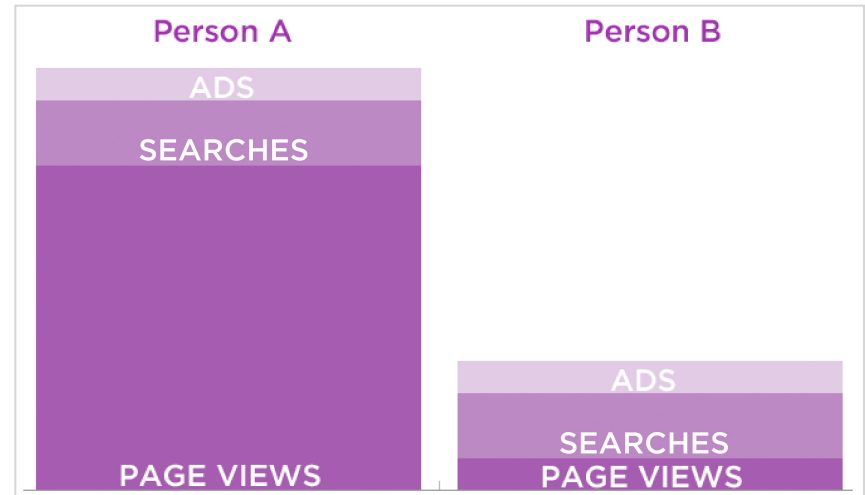
Target ads to users who get highest 'relevance' scores in the targeting categories

Recency Matters, So Does Intensity

Active now...

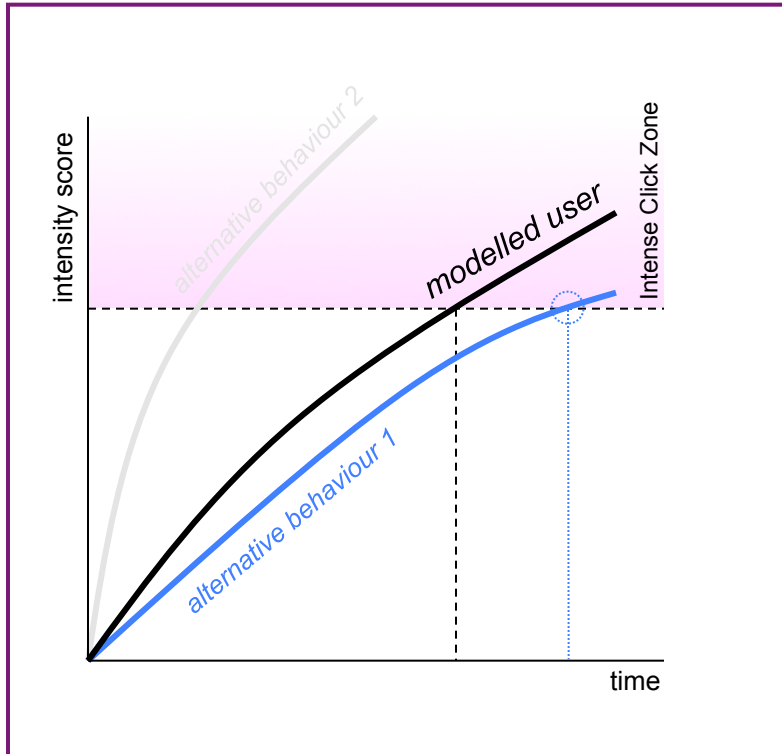


...and with feeling



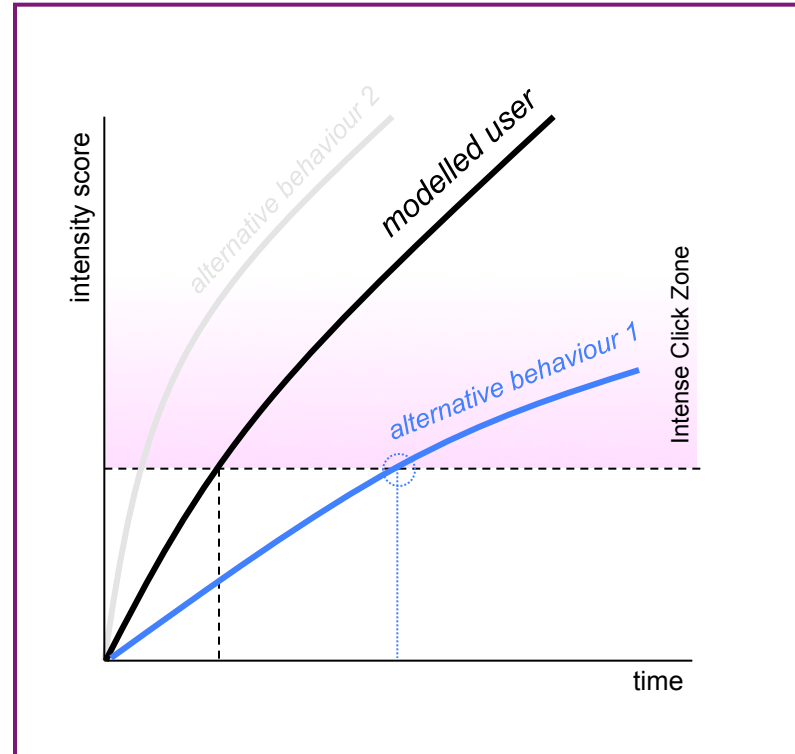
Differentiation | Category specific modelling

Example 1: Category Automotive



Alt Behaviour 1: 5 pages, 2 search keywords, 1 search click, 1 ad click

Example 2: Category Travel/Last Minute

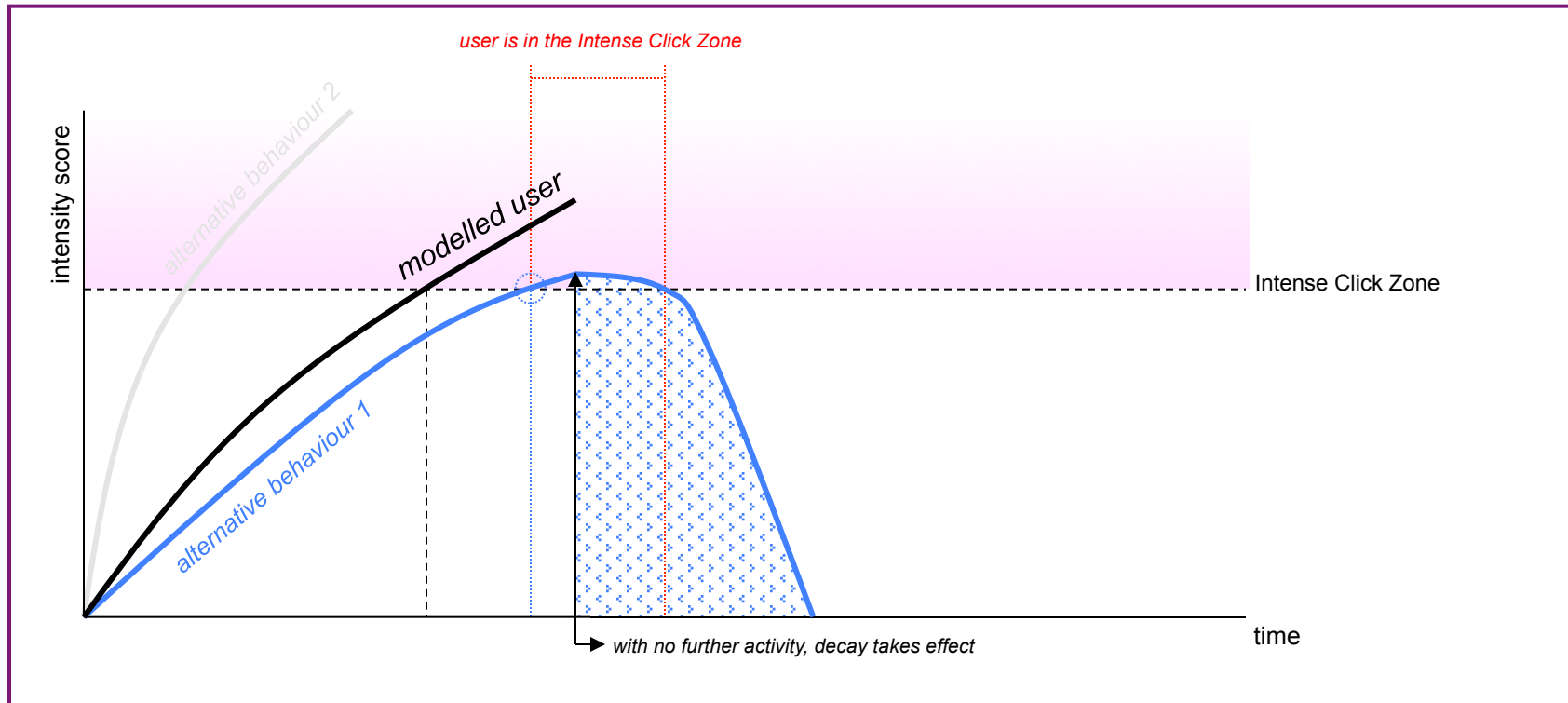


Alt Behaviour 1: 5 pages, 2 search keywords, 1 search click, 1 ad click

Different models allow us to weight and determine intensity and recency

Differentiation | Category specific modelling

Example 1: Category Automotive



Alt Behaviour 1: 5 pages, 2 search keywords, 1 search click, 1 ad click

Different models allow us to weight and determine intensity and recency

Automobile Purchase Intender Example

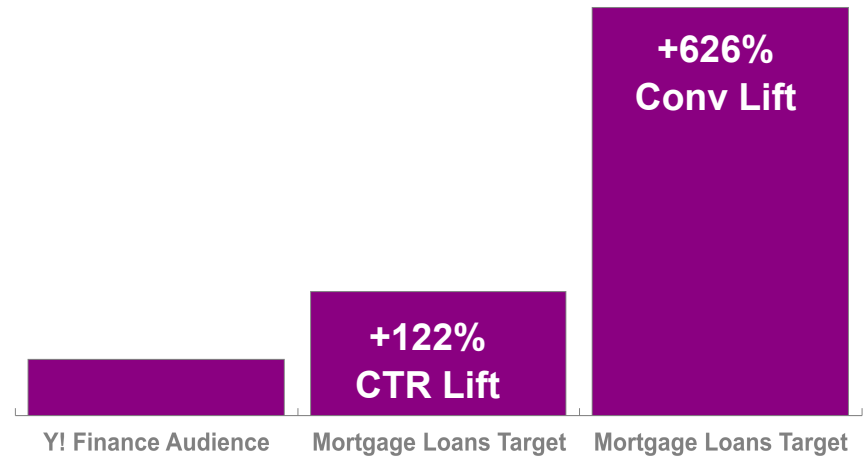
- A test ad-campaign with a major Euro automobile manufacturer
 - Designed a test that served the same ad creative to test and control groups on Yahoo
 - Success metric: performing specific actions on Jaguar website
- Test results: 900% conversion lift vs. control group
 - Purchase Intenders were 9 times more likely to configure a vehicle, request a price quote or locate a dealer than consumers in the control group
 - ~3x higher click through rates vs. control group

Mortgage Intender Example

Example: Mortgages

We found:
1,900,000 people looking
for mortgage loans.

Results from a client campaign on Yahoo!
Network



Example search terms qualified for this target:

Mortgages Home Loans Refinancing Ditech

Example Yahoo! Pages visited:

Financing section in Real Estate

Mortgage Loans area in Finance

Real Estate section in Yellow Pages

Experience summary at Yahoo!

- Dealing with the largest data source in the world (25 Terabyte per day)
- BT business was grown from \$20M to about \$500M in 3 years of investment!
- Building the largest database systems:
 - World's largest Oracle data warehouse
 - World's largest single DB
 - Over 300 data mart data
 - Analytics with thousands of KPI's
 - Over 5000 users of reports
 - Largest targeting system in the world
- Big demands for grid computing (Hadoop)

Social Network

Social Graph Analysis (no time)

Social Network Marketing

Understanding Context for Ads

Case Study: TWITTER Social Marketing?

Viacom's VH1 Twitter campaign on ANVIL (the movie)

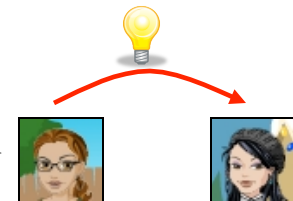
(week of May 14th , 2009 – see AdAge article)

Marketing ANVIL movie

Very niche audience, how do you reach them?

Diffusion

Give 20 free DVD's to major related artists/groups, ask them
To notify Twitter groups – reached over 2M people



Social Identity: power of word of mouth...

What is the Cost to VH-1?

Compare with traditional approach: TV commercials to promote a documentary film?



The Display Ads Challenge Today

The screenshot shows a news article from Associated Press titled "Body parts delivered to Michigan home". The article is dated "1 hour, 23 minutes ago". The main headline is circled in blue. The article text describes two packages containing human body parts (a liver and part of a head) that were mistakenly delivered to a home in Cascade Township, Michigan, instead of a medical research lab. A blue box with the text "What Ad would you place here?" is overlaid on the right side of the article. The article includes quotes from a resident and a sheriff's deputy. On the left side, there are sections for "ELSEWHERE ON THE WEB" with links to CNN.COM, ABC NEWS, and THE CHRISTIAN SCIENCE MONITOR, and "Offbeat News" and "Offbeat Digest" sections from CNN.COM and USATODAY.COM.

Search: All News Advanced

Body parts delivered to Michigan home

AP Associated Press

1 hour, 23 minutes ago

CASCADE TOWNSHIP, Mich. - Two packages containing human body parts — including a liver and part of a head — meant for a medical research lab instead were delivered to a home.

ADVERTISEMENT

The body parts, sent from China, were mistakenly dropped off Thursday at Franck and Ludivine Larmande's home by a DHL express driver who believed the bubble-wrapped items were pieces to a table.

"My husband started to unwrap one and said, 'This is strange, it looks like a liver,'" Ludivine Larmande said. "He started the second one, but stopped as soon as we saw the ear."

"Something wasn't right. It was scary, and I'm glad I didn't open them."

The couple called Kent County sheriff's deputies, who determined the preserved body parts were for medical research, Lt. Roger Parent said.

Authorities believe 28 more bubble-wrapped human organs and body parts could be dispersed across the country, The Grand Rapids Press reported. Two of five packages headed to the northern Michigan lab broke open, scattering their contents.

"There will definitely be a shock to people if they see these things, but there is no hazard to health," Parent said.

What Ad would you place here?

ELSEWHERE ON THE WEB

CNN.COM
Roadside bombs target Iraqi police, officials say

ABC NEWS
In Wealthy Santa Barbara, Some Call a Parking Lot 'Home'

THE CHRISTIAN SCIENCE MONITOR
Need barbed wire? Try the concierge.

CNN.COM
Offbeat News
Get the latest news, photos and video of the funky, and weird.
» More at CNN.com

USATODAY.COM
Offbeat Digest
Latest news on the strange, quirky and just plain odd.
» More at USATODAY.com

The Display Ads Challenge Today

Damaging to Brand?

The screenshot shows a news article titled "Body parts delivered to Michigan home" from Associated Press. The article text includes: "1 hour, 23 minutes ago", "CASCADE TOWNSHIP, Mich. - Two packages containing human body parts — including a liver and part of a head — meant for a medical research lab instead were delivered to a home.", "The body parts, sent from China, were mistakenly dropped off Thursday at Franck and Ludivine Larmande's home by a DHL express driver who believed the bubble-wrapped items were pieces to a table.", "My husband started to unwrap one and said, 'This is strange, it looks like a liver,'" Ludivine Larmande said. "He started the second one, but stopped as soon as we saw the ear.", "Something wasn't right. It was scary, and I'm glad I didn't open them.", "The couple called Kent County sheriff's deputies, who determined the preserved body parts were for medical research, Lt. Roger Parent said.", "Authorities believe 28 more bubble-wrapped human organs and body parts could be dispersed across the country, The Grand Rapids Press reported. Two of five packages headed to the northern Michigan lab broke open, scattering their contents.", "There will definitely be a shock to people if they see these things, but there is no hazard to health," Parent said.

Overlaid on the article is a hand-drawn advertisement in a dark brown frame. It features the UPS logo in the top left corner. The text inside the frame reads: "It's UPS early morning delivery." and "« Rollover for video." The word "ADVERTISEMENT" is written in small capital letters above the text.


The Display Ads Challenge Today

News World news Greece

Violence continues in Greece as rioters firebomb buildings

Protesters in Athens torch offices and cars amid clashes with police after memorial for teenager

Ant Dawar
guardian.co.uk, Sunday 21 December 2008 17:05 GMT
Article history



A youth assaults a police officer in Athens during a week of riots after the shooting of a teenager. Photograph: Bela Scandelszky/AP

World news
Greece

More news:

What Ad would you place here?


The Display Ads Challenge Today

News World news Greece

Violence continues in Greece as rioters firebomb buildings

Protesters in Athens torch offices and cars amid clashes with police after memorial for teenager

And Dawar
guardian.co.uk, Sunday 21 December 2008 17:05 GMT
Article history



World news
Greece

More news:

Win a mini-break in gorgeous Greece...

Irrelevant and Damaging to Brand

Completely Irrelevant

Related Topics: [make it snacktastic](#) [working it](#) [summer living](#) [summer travel](#) [summer beauty](#)

What happens to your body after you drink a soda every day, for a long time

by Bliss Tree, on Tue Jun 7, 2011 4:58am PDT
627 Comments | [Post a Comment](#) | [Read More from This Author »](#) | [Report Abuse](#)

[Share](#) [retweet](#) 187 [Email](#) [Print](#)



Sugar rushes and caffeine highs followed by a depressing energy crash are **what happens to your body if you drink a soda right now**, but plenty of BlissTree readers actually seem to be okay with that. Some of you think it's alarmist to compare a caffeine and sugar rush to doing drugs, and some just don't really care about the slump they'll find themselves in after drinking 39 grams of sugar, but what makes us really worried about a soda-slurping

What is Your Credit Score as of 06/15/2011

TransUnion.	478
Experian	511
EQUIFAX	540

See Yours Instantly! [FreeScore.com](#)

Shine on facebook

NetSeer: Intent for Display

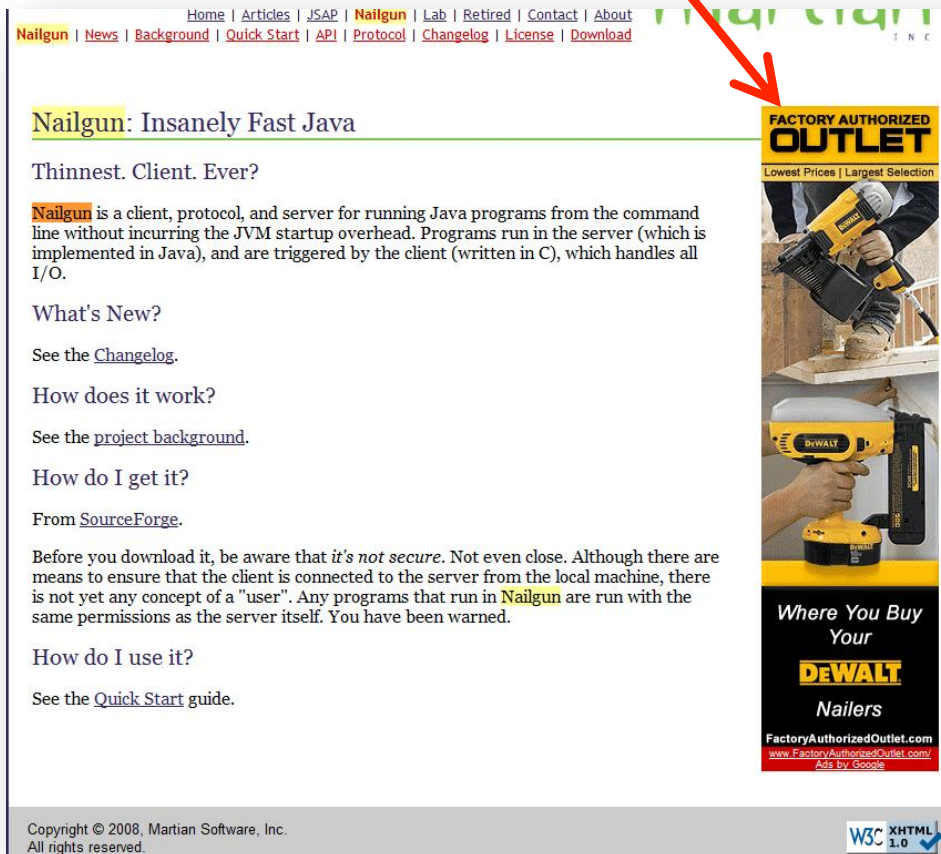
- Currently Processing 4 Billion Impressions per Day

The screenshot displays a website interface with a sidebar on the left for 'TRADE-S' containing navigation links like 'Home', 'FREE Display Quotes', 'Plan for Success', 'Booths', 'Promotions', and 'Help & Contacts'. The main content area features a profile for 'Rachael' with a photo and a weekly schedule. Below this is a section for 'cook with rach' with recipe links. The central article is titled 'Prevent a Stroke By Drinking Caffeine' and includes an image of a human head with a glowing brain. A blue highlight box surrounds a Starbucks advertisement for K-Cup Packs. To the right of the article is a table of caffeine content in various products.

RECENTLY ADDED	CAFFEINE
TEHO Energy Drink	160mg
Speed Energy Drink	186mg
Chameleon Cold Brew Coffee	2160mg
Ed Hardy King Dog Energy Mints	97mg
Cocoa Powder (Hershey's)	8.4mg
...	...

Problem: Hard to Understand User Intent

Contextual Ad served by Google



Home | Articles | JSAP | **Nailgun** | Lab | Retired | Contact | About
Nailgun | News | Background | Quick Start | API | Protocol | Changelog | License | Download

Nailgun: Insanely Fast Java

Thinnest. Client. Ever?

Nailgun is a client, protocol, and server for running Java programs from the command line without incurring the JVM startup overhead. Programs run in the server (which is implemented in Java), and are triggered by the client (written in C), which handles all I/O.

What's New?

See the [Changelog](#).

How does it work?

See the [project background](#).

How do I get it?

From [SourceForge](#).

Before you download it, be aware that *it's not secure*. Not even close. Although there are means to ensure that the client is connected to the server from the local machine, there is not yet any concept of a "user". Any programs that run in **Nailgun** are run with the same permissions as the server itself. You have been warned.


How do I use it?

See the [Quick Start](#) guide.

Copyright © 2008, Martian Software, Inc.
All rights reserved.

W3C XHTML 1.0

FACTORY AUTHORIZED OUTLET
Lowest Prices | Largest Selection



Where You Buy Your
DEWALT
Nailers

FactoryAuthorizedOutlet.com
www.FactoryAuthorizedOutlet.com/
Ads by Google

What NetSeer Sees:

URL:



- Programming in Java
- Java Development Kit
- Sun Java
- Java Programmers
- Java Virtual Machine
- Static IP Address
- DNS Servers
- Dynamic DNS
- Java Runtime Environment
- Java Software
- Java Programming Language
- Server Configuration
- Dynamic IP Addresses
- Free Java
- Code in Java
- Implementation of Java
- Local Machine
- Running Java Programs
- Java Language
- Java Platform
- Spring Framework
- Home Server
- Virtual Server

ContextLinks © by NetSeer www.netseer.com

Case Studies

nPario – Data Management Platform
ChoozOn – Big Data over Offers Universe

Example of a Big Data DMP



Scale

nPario builds an infinitely scalable data management platform (DMP) that allows advertisers and marketers to **manage**, **understand**, and **monetize** their data. Their technology has been proven at companies such as Yahoo and EA.

Applications

nPario applications include **segmentation** of audiences for increasing the value of advertising, **reporting/analytics** for examining performance, **attribution** to show which advertising works, and **experimentation** to test ideas.

Access

nPario emphasizes putting access to data in the hands of marketing, advertising and other business users

nPARIO™

Powerful Technology.

**540m Users, 8+ Petabytes,
& 16 Patents**



nPario has the only **commercially available** Big Data management technology built for one of the “Big Five”.

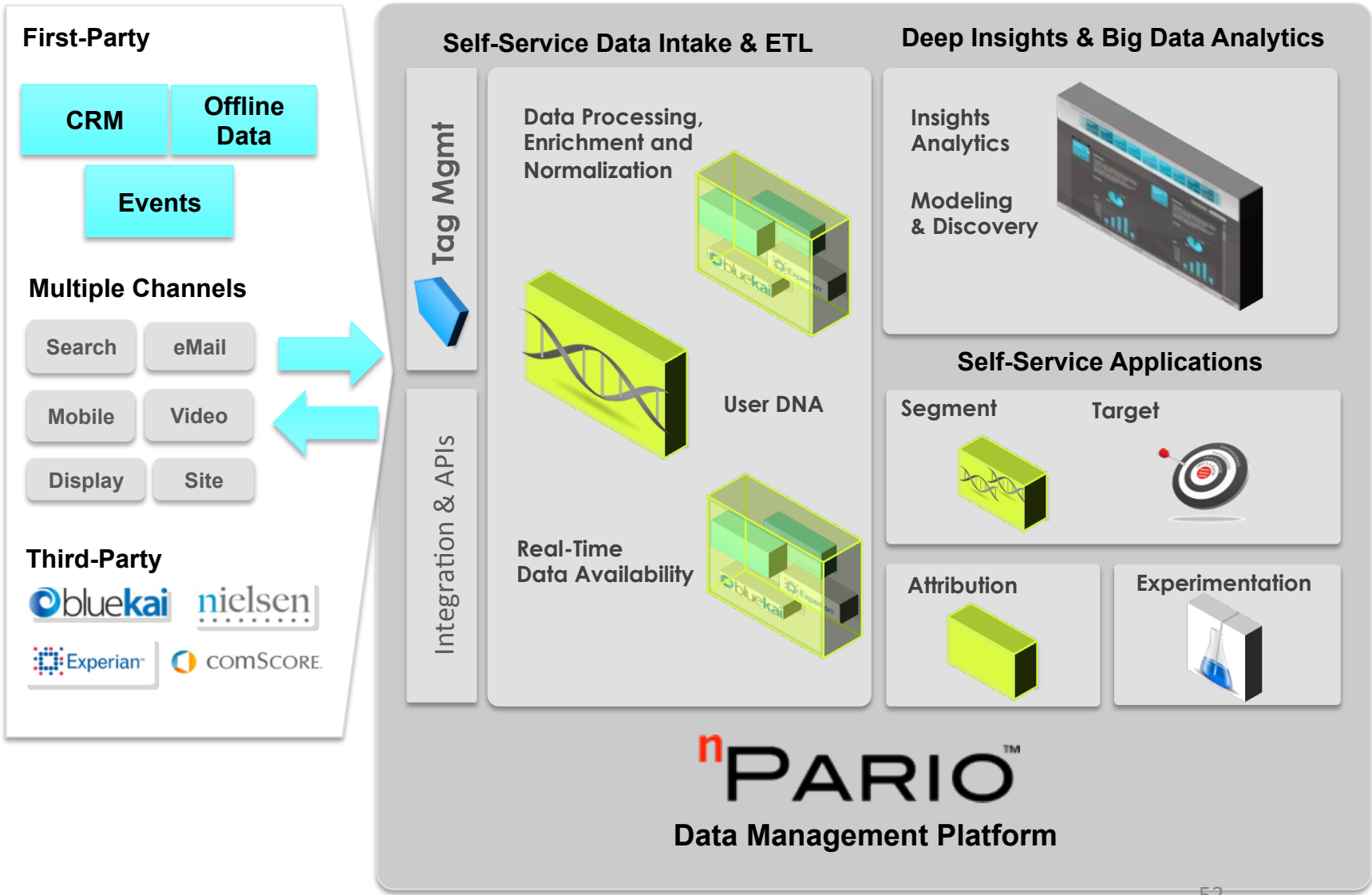
Significant Investment

nPario’s technology is the result of more than \$50m investment in development and 16 issued patents.

In Production at Yahoo

nPario technology manages the world’s largest data system. Used for Yahoo’s Marketing and Advertising business. Used across Yahoo’s platforms and 120 online properties. Used by hundreds of analysts

Data Sources



Marketing to 100+ million Gamers

nPario Cas

Challenge: Provide cross-platform campaign insights for advertisers and enable audience discovery across channels.

Result: Unified view of gamers across multiple cross platform data sources.

Pogo (online), sponsored content, Console game interaction and ad interaction (Xbox, PS3), Mobile, Playfish (facebook), external sources (Collective Media, Comscore, Omniture, Dynamic Logic)



“EA increased its worldwide audience reach by 30% this year [...]. Combining that major jump in reach with the launch of *EA Legend* puts us in perfect position to compete”

**Dave Madden, Senior Vice President of Global Media Solutions
Electronic Arts.**



Advertiser

Campaign

Media Channel

Date

Advertiser



All

June, 2011

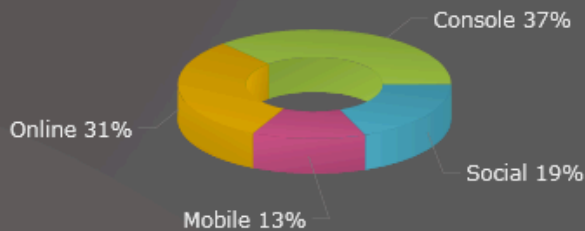
Summary

Save To Excel

Save To PDF

Total Reach

Unduplicated Uniques: 24.6 mm



Performance

1.17M Total DLC downloads

2.7M Social quest completions

21M Virtual goods claimed

.6% Top CTR from Pogo (July '11)

1.2% Top CTR from Mobile (June '11)

Brand Impact

+5% Increase in total US sales

+25% Growth in positive social conversations

+15% Awareness lift

+22% Brand Favorability lift

+31% Purchase intent lift

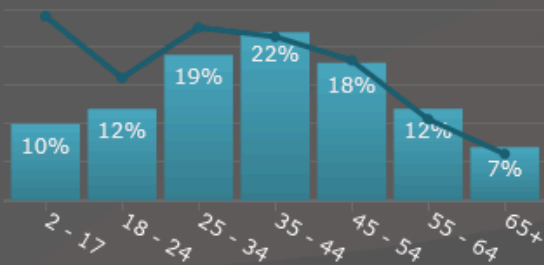
Audience

Gender

37% MALE

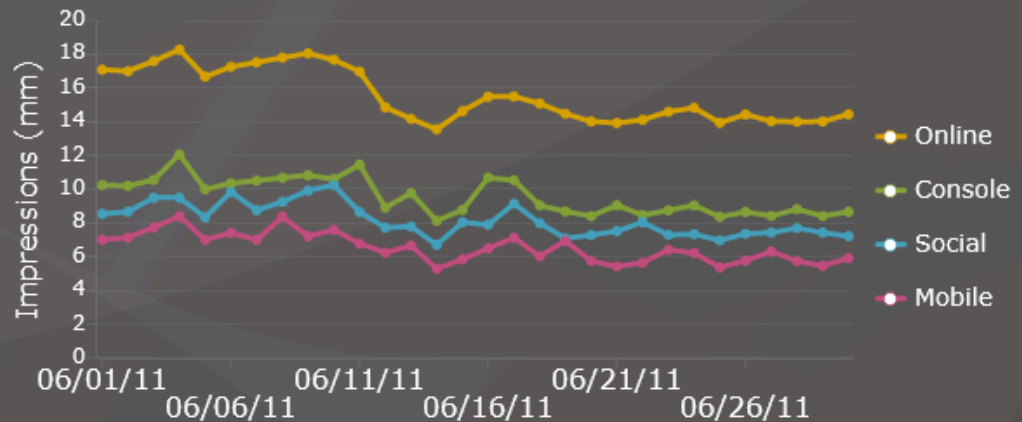
63% FEMALE

Age



■ Campaign ● General Internet Pop.

Ad Delivery



Campaigns

Audience

Admin

powered by
PARIO



Specialized Search through Big Data
Analytics over the Offers Universe

Chaos for Consumers

RESTAURANT.COM
EAT. DRINK. SAVE MONEY.

6PM.COM
It's time to get your brand hot!

COUPON CABIN

Coupon Mountain

RetailMeNot

savings.com

UNITED

VISA

AMERICAN EXPRESS

DISCOVER

bradsdeals
handpicked savings

EBATES
WHICH IT PAYS TO SHOP ONLINE

upromise

MyPoints

chozOn

ONE KINGS LANE

GILT

RueLaLa

ideeli

livingsocial presents
your daily deal

GROUPON
Collective Buying Power

eversave
SHARE the SAVE

Google

bing

YAHOO!

Deep Discount Sites

Coupon Sites

Loyalty Programs

Daily Deals Sites

Consumer

Online Loyalty Program

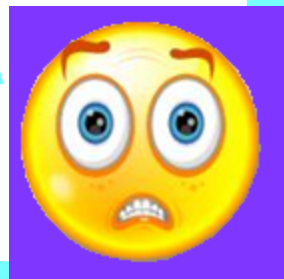
Search Engines

Social Network

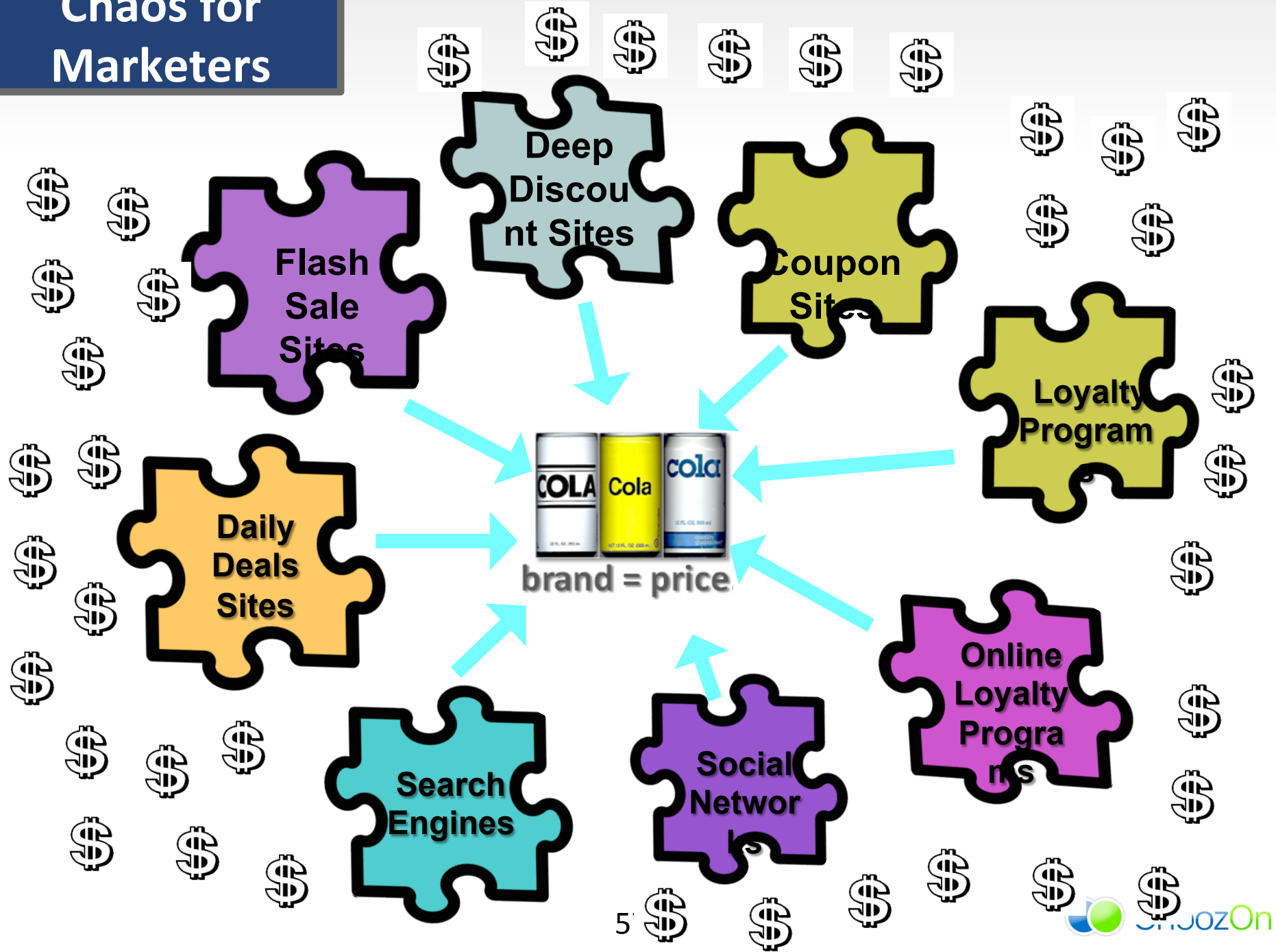
facebook

LinkedIn

twitter



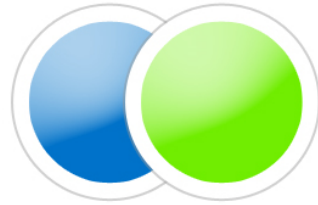
Chaos for Marketers



What Consumers & Marketers Want

Consumers

- Value from brands they love
- Tame the deal chaos



ChoozOn

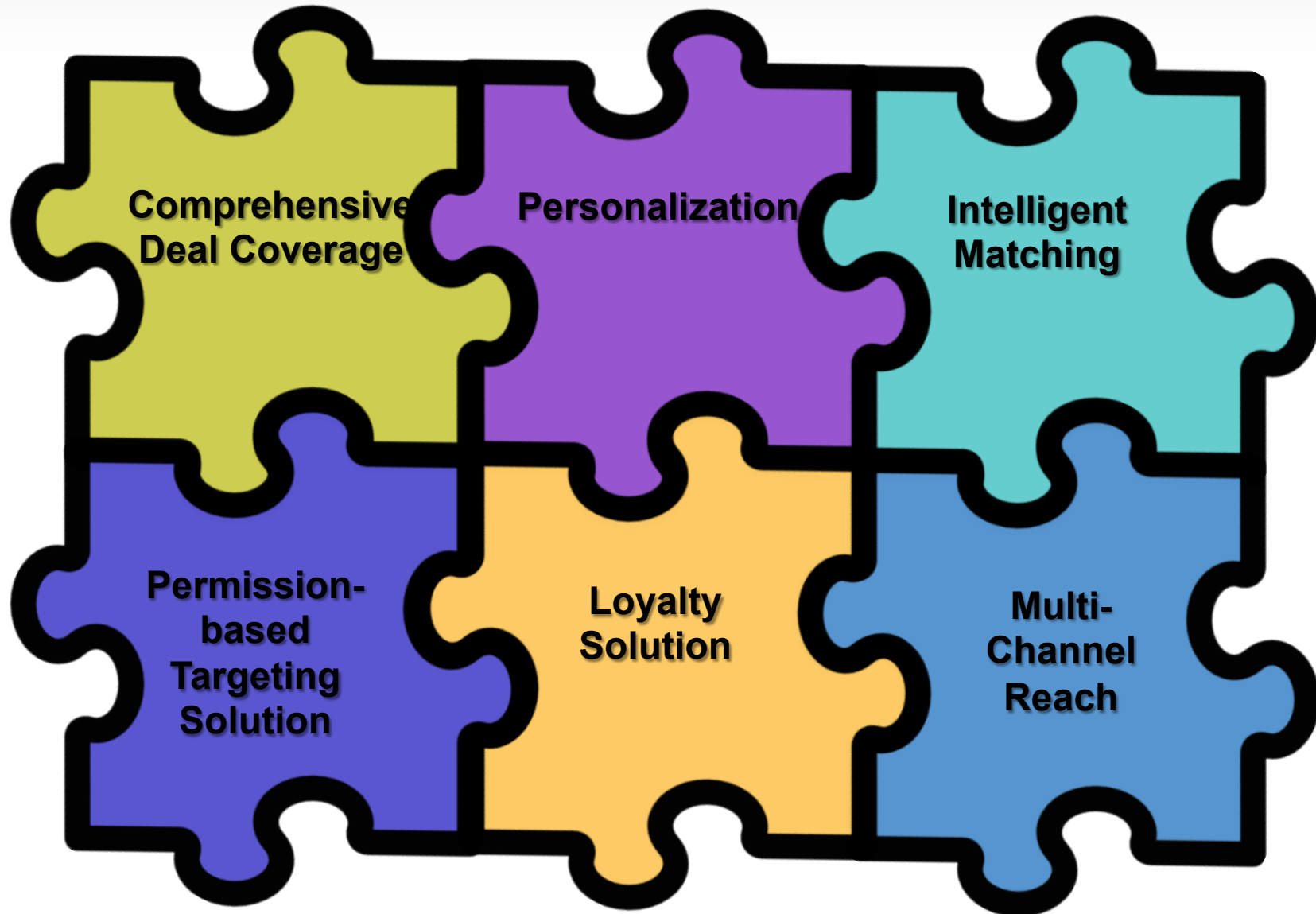
Your Brands, Your Savings. Your Way.™



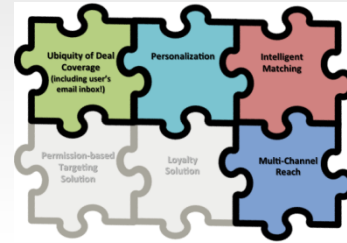
Marketers

- Reach targeted consumers
- Build loyalty
- Create brand evangelists

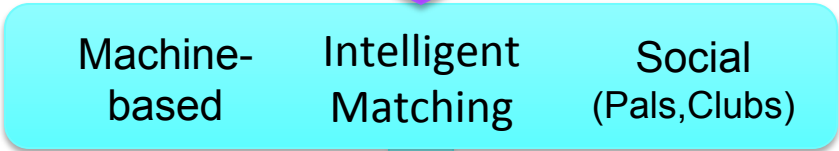
Keys to Being THE Consumer Network



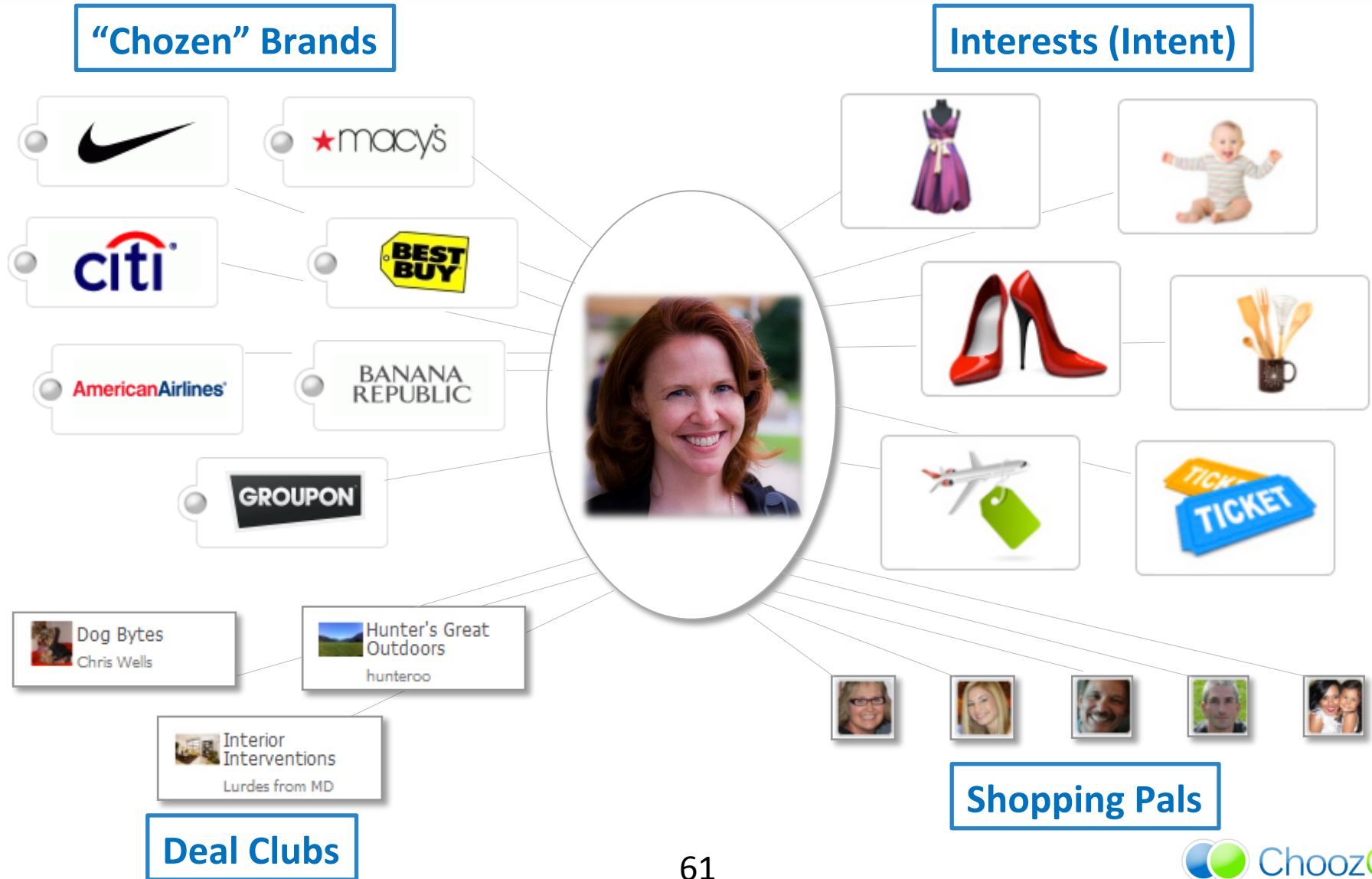
Solution for Consumers



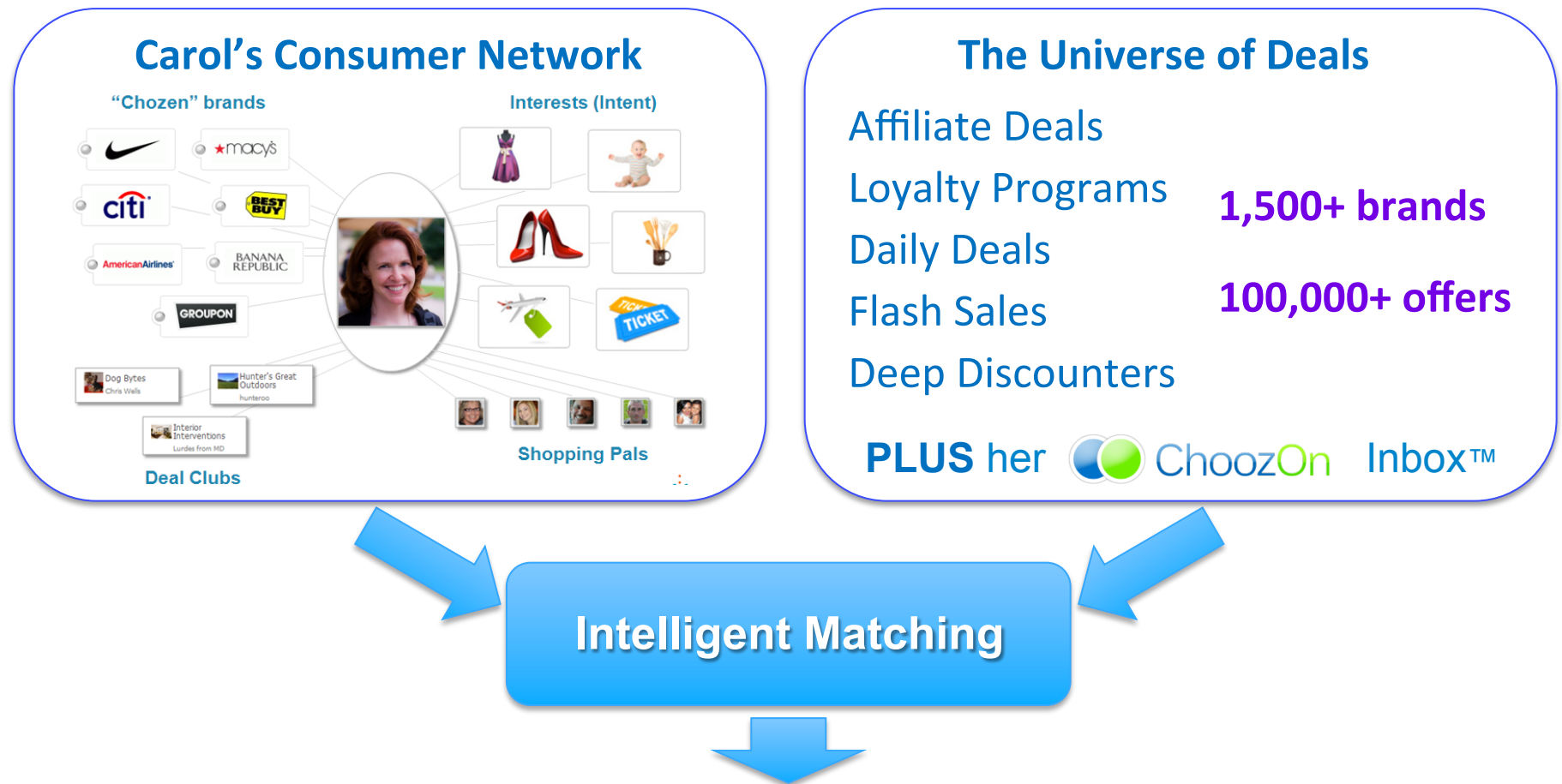
Choozer Interests & Preferences



Carol's Consumer Network



What Carol Gets



A Personal Shopper for Deals

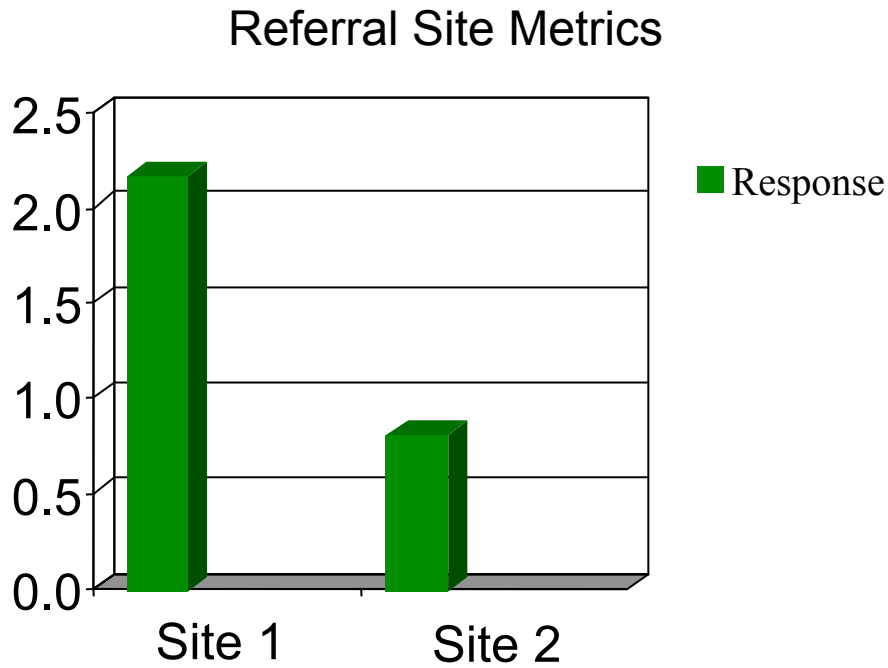
Big Picture on Big Data Analytics

Key points

Don't Forget The Basics

- Metrics and Scorecards are the first steps to awareness
- Plays a huge role in deploying predictive models and monitoring and proving their effectiveness
- Often scorecards require
 - Going through huge amounts of data to produce the required metrics
 - Ability to get to the metrics in low latency
 - Ability to modify metrics and update quickly
 - Integration with data warehouse

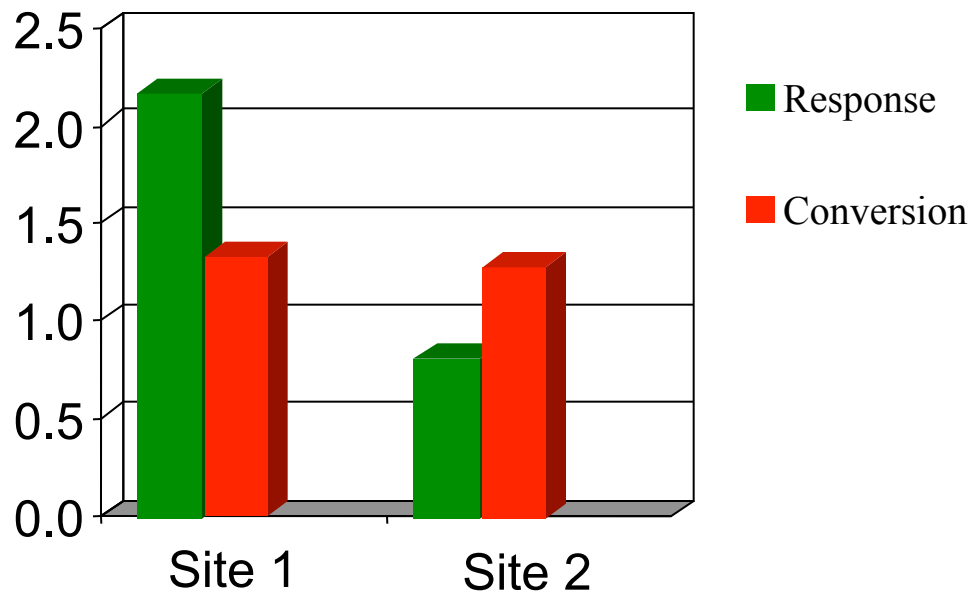
Focus On The Right Measure



- Total traffic not a good performance measure
- High-traffic referral sites often produce poorer quality click throughs
- Ads best response not most effective
- Target the message

Focus On The Right Measure

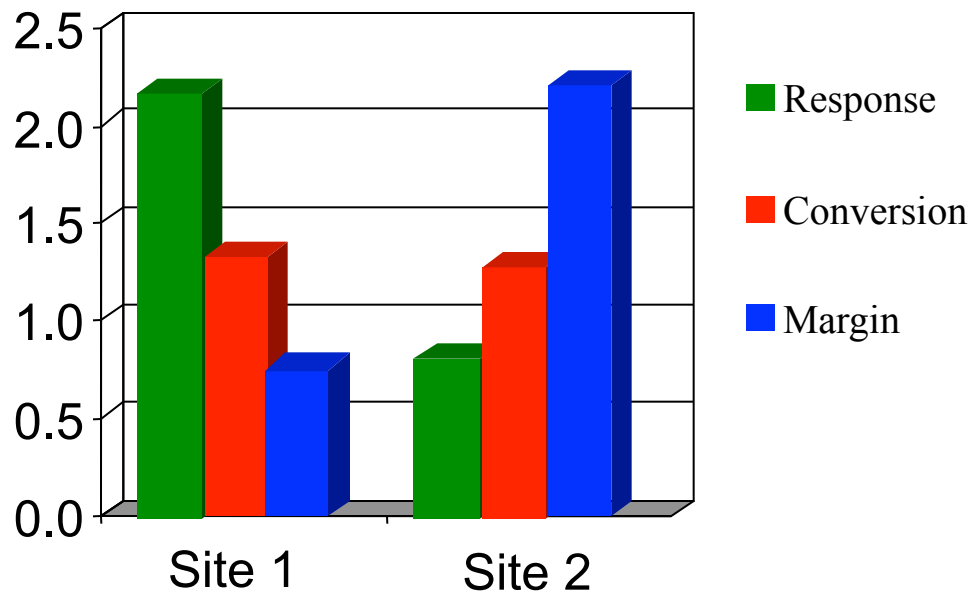
Referral Site Metrics



- Total traffic not a good performance measure
- High-traffic referral sites often produce poorer quality click throughs
- Ads best response not most effective
- Target the message

Focus On The Right Measure

Referral Site Metrics



- Total traffic not a good performance measure
- High-traffic referral sites often produce poorer quality click throughs
- Ads best response not most effective
- Target the message

Sometimes,
Simple is Very Powerful!

Retaining New Yahoo! Mail Registrants

Integrating Mail and News

- Data showed that users often check their mail and news in the same session
 - But no easy way to navigate to Y! News from Y! Mail
- Mail users who also visit Y! News are 3X more active on Yahoo
 - Higher retention, repeat visits and time-spent on Yahoo

“In the news” Module on Mail Welcome Page

- Increased retention on Mail for light users by 40%!
 - Est. Incremental revenue of **\$16m a year** on Y! Mail alone

Free Camera Phone
Free Shipping

Folders [Add - Edit]

- Inbox (172)
- Draft
- Sent
- Bulk (44) [Empty]
- Trash [Empty]

My Folders [Hide]

- Friends
- Jobs
- Shopping

678 is US average credit score. See yours.

\$150K loan for only \$550/month!

2% of 100.0MB

ADVERTISEMENT

Visit your local Audi dealer for extraordinary Audi offers. Audi

2004 TT
Click for details

Never follow bayareaaudidealers.com

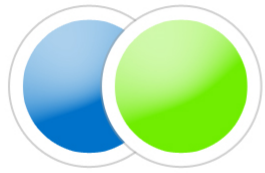
In the News: Top Stories [Hide]

TOP STORIES

- Najaf sees worst fighting since 2003
- Arrests reportedly damage al-Qaida network
- Reservist: Intelligence officials led Iraq abuse
- Judge lifts Letourneau no-contact order
- Markets: Dow ↓1.5% Nasdaq ↓2.5%

Benefits of Advanced Analytics

- Advanced Analytics brings out the real value of data
- The business begins to understand the true value and role of data in moving the big needles
- Focus on useful requirements from data, rather than “data acrobatics”
- Value creation from data leads to proper investment scoping
 - Many are realizing predictive analytics and data mining are much more useful than reporting
 - Integration of analytics story with data storage very critical
- Big data makes analytics even more essential and more useful
 - Avoiding the challenges of separating analytics from big data are increasingly important



ChoozOn

Your Brands. Your Savings. Your Way.™

Thank You! & Questions?

Twitter: @usamaf

Usama@choozOn.net

www.ChoozOn.com