



SOCIAL MEDIA

Description
Underlying Technology
Categories & Benefits
CI's Role
Dealing with CGM
BlogOSphere
Metrics
Future of SM
Q&A



**What
Is
Social
Media
???**

Please watch the following video:

<http://www.youtube.com/watch?v=MpIOCIX1jPE>

- Internet-based tools to share and discuss information among humans in various medium
- Integrates technology, social interaction, construction of words, pictures, video, and audio
- Core data is CGM around content, context, connection, community
- Content can be authored, searched, linked, tagged, recommended, and signaled

- Small or large audiences can be reached
- More democratic and harder to manipulate
- A huge bazaar with many dark alleys
- Easier to Spam
- Easier to game the system

- In contrast to traditional/mass media including newspapers, magazines, television
- Traditional media is expensive and requires significant financial capital
- Cost limits its usage to commercial / political purposes
- Mass media can manipulate large group of people through its outlets
- Mass media is regulated and policed



***Looking
Under
The
Hood***

- Built around Web 2.0
- If Web 1.0 was “read-only”, Web 2.0 is “read-write”
- 1.0: static, centrally managed, slow to change, unidirectional
- 2.0: dynamic, distributed, spread, bidirectional
- Enables proliferation and interconnectivity of web-based content

- Networked web is the new platform
- Humans interact through HTML, AJAX, Flash
- Machine interaction through web services: SOAP, REST
- XML, RSS, ATOM as data back bone transports



4 Categories

**Communicat
ion**

**Collaboratio
n**

Multimedia

**Entertainme
nt**

Communication

- Boards: Yahoo! Finance, Apple Discussions
- Blogs: Blogger, WordPress, LiveJournal
- Microblogs: Twitter
- Social networking: FaceBook, MySpace, LinkedIn
- Social network aggregation: FriendFeed
- Events: Upcoming

Collaboration

- Wikis: Wikipedia
- Social bookmarking: del.icio.us, StumbleUpon
- Social news sites: Digg, Reddit

Multimedia

- Photo sharing: Flickr
- Video sharing: YouTube
- Audio sharing: HypeMachine

Entertainment

- Virtual worlds: Second Life
- Online gaming: World of Warcraft





Benefits

Personal

Commercial

Political

Research

Personal

- New way of interaction
- Simply to socialize
- To express oneself using multi media
- To be part of a community
- To ask questions and provide answers to others

Commercial

- Using collective intelligence for various reasons
- Measure overall opinion/sentiment on products/brands
- Publicizing products/brands
- Free beta testing
- Risk management
- Huge targets for ad campaigns and PR agencies

Political

- Fund raising
- Reaching people and younger generation
- Polling
- Viral campaigns
- One person one vote, everyone can impact

Research

- Understanding information Propagation
- Trends
- Predictability Models
- Psychological, sociological, ethnographic studies
- Fundamental metrics



COLLECTIVE INTELLECT. UNLOCK THE POWER OF SOCIAL MEDIA

CI 's Role

Collection

**boards, blogs, news, micro,
video.....**

Analysis

**categorization, theme extraction,
ranking.....**

Presentation

MARKET MOVERS

Newsworthy

Determined based on increased frequency of appearance in news

STX	8.00	-0.11	-1.36%
LSI	3.83	-0.07	-1.79%
FFIV	19.78	+0.04	+0.20%
CVH	13.93	-14.56	-51.11%
VMW	20.06	+1.33	+7.10%

More Market Movers

- Most Active
- Most Price % Change
- Industries on the Move
- Unusual Trading Volume

COMMUNITY SENTIMENT

Top stocks creating buzz on Yahoo! Finance message boards

 **Bullish**
Apple Inc. (AAPL)
Yahoo! Inc. (YHOO)
SanDisk Corp. (SNDK)

 **Bearish**
Nektar Therapeutics (NKTR)
Harley-Davidson, Inc. (HOG)
Honeywell International Inc. (HON)

Powered by Collective Intellect, Inc.



Bear-Market Funds

Bear-market funds are designed to go up when the market goes down, and vice versa -- so they're popular lately. But it's not a one-size-fits-all option, and timing is everything...

- **How-to Guide:** Put Savings First With a Budget
- **Calculator:** Value of reducing or foregoing expenses

EXPERT OPINION



Mortgage Fright and Moral Quandaries

- Jack Guttentag

Fear in the current mortgage market is contagious. The only antidote to that is good information.

- Why I'm Still Buying - Ben Stein
 - 'We're All Hosed': A Wall Street Insider on the Economic Crisis - Laura Rowley
 - Vertigo on Wall Street - Mick Weinstein
- » [View all experts](#)

SPECIAL FEATURES

Many Banks Haven't Learned Their Lesson

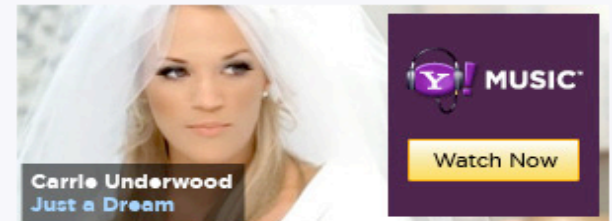
- NYT
Some financial institutions are continuing the credit free-for-all, even going so far as to target struggling borrowers whose finances have melted down amidst the economic turmoil.

More Features

- 4 Financial Crisis-Related Scams to Beware Of
- 6 Simple Ways to Save Thousands a Year
- Advice on Selecting Benefits During Open Enrollment Time
- ETFs vs. Stocks - How to Profit With Sector ETFs

6 month CD	3.07%	3.12%
1 year CD	3.60%	3.59%
3 year CD	3.71%	3.76%
MMA	2.47%	2.39%
\$10K MMA	2.81%	2.73%
\$25K MMA	3.13%	3.05%

» [View rates in your area](#)



Carrie Underwood
Just a Dream

Y! MUSIC

Watch Now

YAHOO! FINANCE

- Banking & Budgeting
- Currency
- Calculators
- ETFs
- Experts
- Investing
- Insurance
- Market Stats
- Message Boards
- Mutual Funds
- Personal Finance
- What's New

ALSO ON YAHOO!

- 360
- Autos
- Finance
- Games
- Groups
- Health
- Hot Jobs
- Mail
- Maps
- Movies
- Music
- My Yahoo!
- News
- Shopping
- Sports
- Tech
- Travel
- TV

» [All Y! Services](#)

THINGS TO DO

- **NEW!** Read Our Blog
- Send Feedback
- Check Stock Quotes
- Search Homes for Sale
- Check Home Values
- Find a New Car
- Search Jobs Across the Web



***Consumer
Generated
Media
Analysis***

- Combing through data to discover information, patterns, knowledge, etc
- Routine monitoring of social media sources
- Analysis of social media content to identify leading issues and sentiment or tone
- Charts and visualizations
- Trend analysis
- Identification and profiles of influential bloggers/communities
- Buzz analysis
- Web-based dashboard for drill-down analysis and report generation
- Social Search can be used to drive traffic, conversions, and increase ROI by monitoring conversations happening online



Campaigns

Listen

Engage

Measure

Listen to map the landscape

- ✓ Comprehensive indexing
- ✓ Semantic analysis
- ✓ Sentiment analysis
- ✓ Theme extraction
- ✓ Demographic analysis
- ✓ Influencers analysis

Engage to participate and ignite the conversation

- ✓ Interact with influencers
- ✓ Comment on related posts
- ✓ Release videos
- ✓ Create high quality content on a dedicated blog
- ✓ Interact with social media (Facebook, Digg, ...)

Measure and optimize based on performance

- ✓ Measure influencers impact
- ✓ Track comments and feedbacks
- ✓ Track videos, page views, comments,
- ✓ Measure dedicated blog traction
- ✓ Analyze social media votes, comments, application usage



CGM
Case
Studies

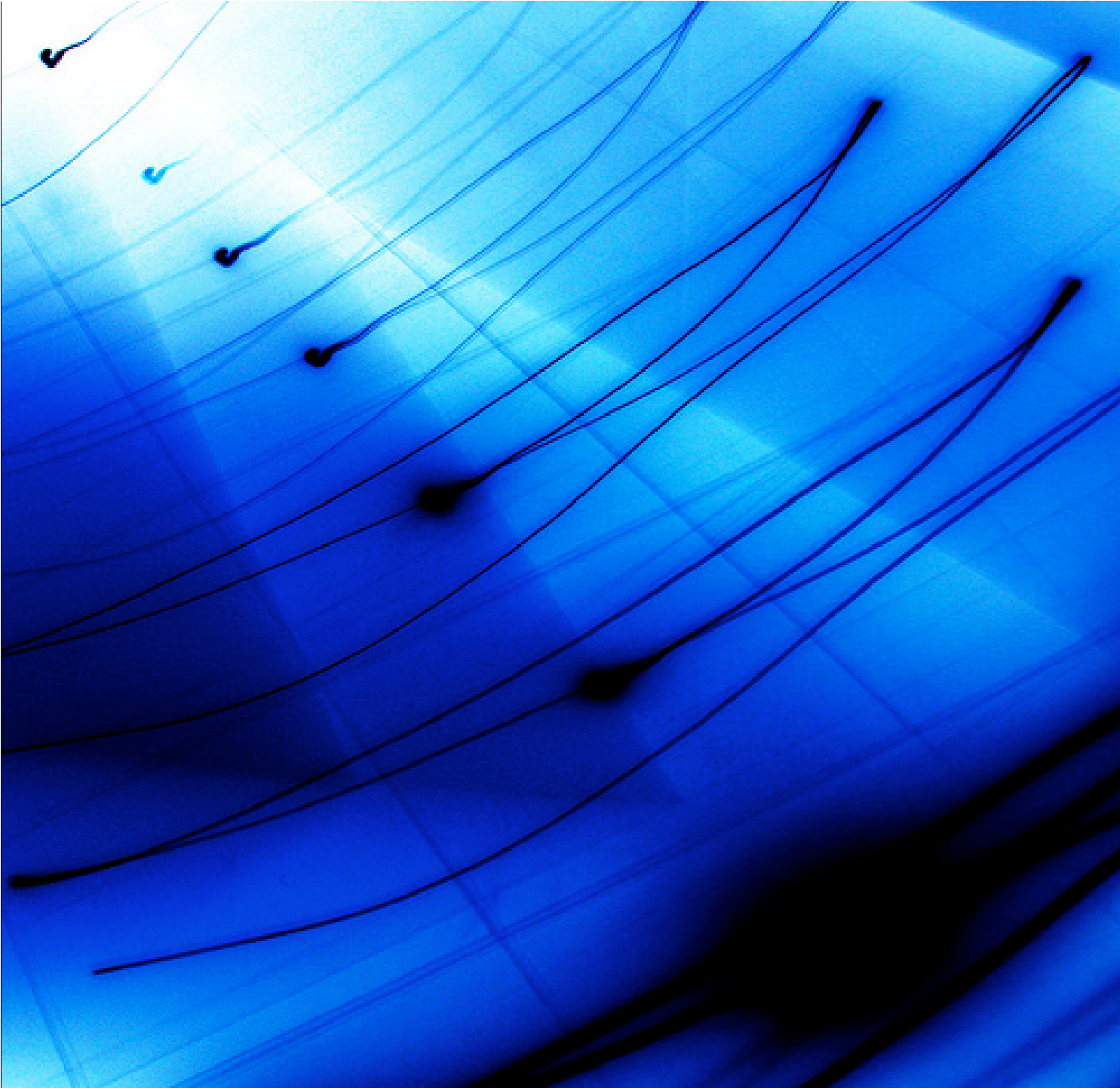
Cisco

- ❑ 3/2008: Cisco launched ASR 1000 Router Series via SM solely
- ❑ Their 5th campaign
- ❑ “Virtually, Virally, Visually”
- ❑ Routers aren’t sexy
- ❑ All about building an experience
- ❑ Created a fun micro site to create and spread the buzz
- ❑ Cross posted videos from micro site to YouTube to extend reach
- ❑ Established Second Life presence with a countdown calculator

Cisco

- ❑ Cisco Created a Facebook group for non Second Life users
- ❑ Created an interactive 3D game - to attract the large gaming audience
- ❑ Created a widget holding videos, documents and images being shared
- ❑ Blogged about it on the Cisco blog to spread the buzz
- ❑ Created a social media release to reach out to bloggers
- ❑ Product was launched via video on Second Life cross posted FB and YT
- ❑ Created 'Ask the Expert' - a forum between customers and engineers

- ❑ Ron Paul in one day, raised over \$6 million in an online-only effort.
- ❑ Raised approximately \$20 million 4th quarter 2007 primarily online.
- ❑ The “Ron Paul Army” supporters overwhelmed blog comments sections which mentioned him, wrote thousands of independent blog posts, and posted well over 100,000 videos on YouTube.
- ❑ Obama offers the most current and in-depth social media content, including widgets, iPhone apps, pod casts, and Twitter.
- ❑ His content is up-to-date, full of informative, relevant content, keeping young supporters involved at a familiar level for them.



***Automation
Of
CGM
Analysis
And
Its
Challenges***

Duplicate posts

- ❖ People tend to copy and paste textual information
- ❖ High quality posts change hands frequently and rapidly
- ❖ May represent buzz

Spam blogs

- ❖ A huge percentage of blogosphere is spam
- ❖ Link farms
- ❖ Link laundering (bait / target)
- ❖ Software morphing
- ❖ Ping server abuse

Data categorization

- ❖ Posts discussing multiple topics
- ❖ Required training
- ❖ Short text
- ❖ Grammar rules

No unique resource ID

- ❖ Various URLs to identify a single site

System Gaming

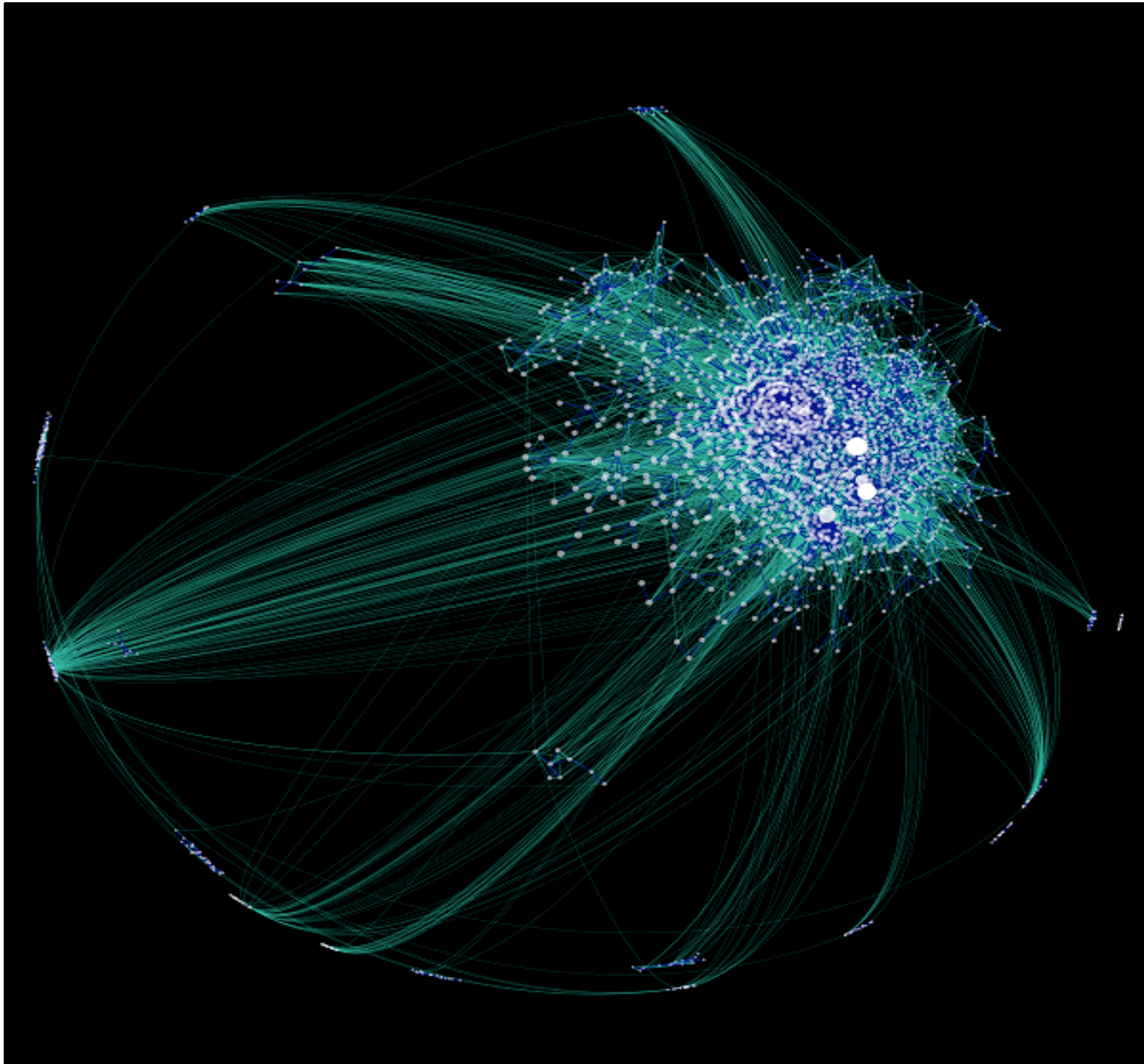
- ❖ Google bombing
- ❖ Aggregators

Real timeliness

- ❖ Various time zones
- ❖ Inaccurate publish dates
- ❖ Futuristic dates
- ❖ Scalability issues

Ads

- ❖ Miss categorization
- ❖ Identification of ads on a page
- ❖ Dynamic ads
- ❖ RSS/ATOM feed ads



Blog

O

Sphere

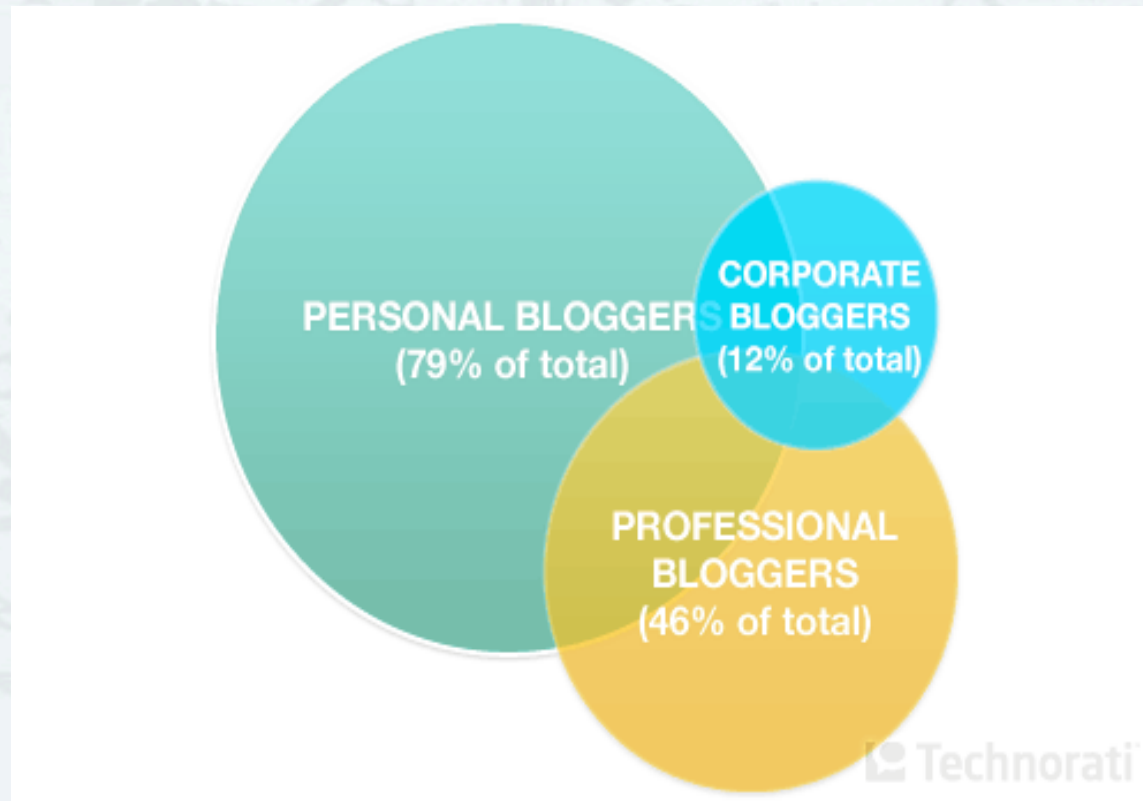
- comScore MediaMetrix (August 2008)
 - o Blogs: 77.7 million unique visitors in the US
 - o Facebook: 41.0 million | MySpace 75.1 million
 - o Total internet audience 188.9 million

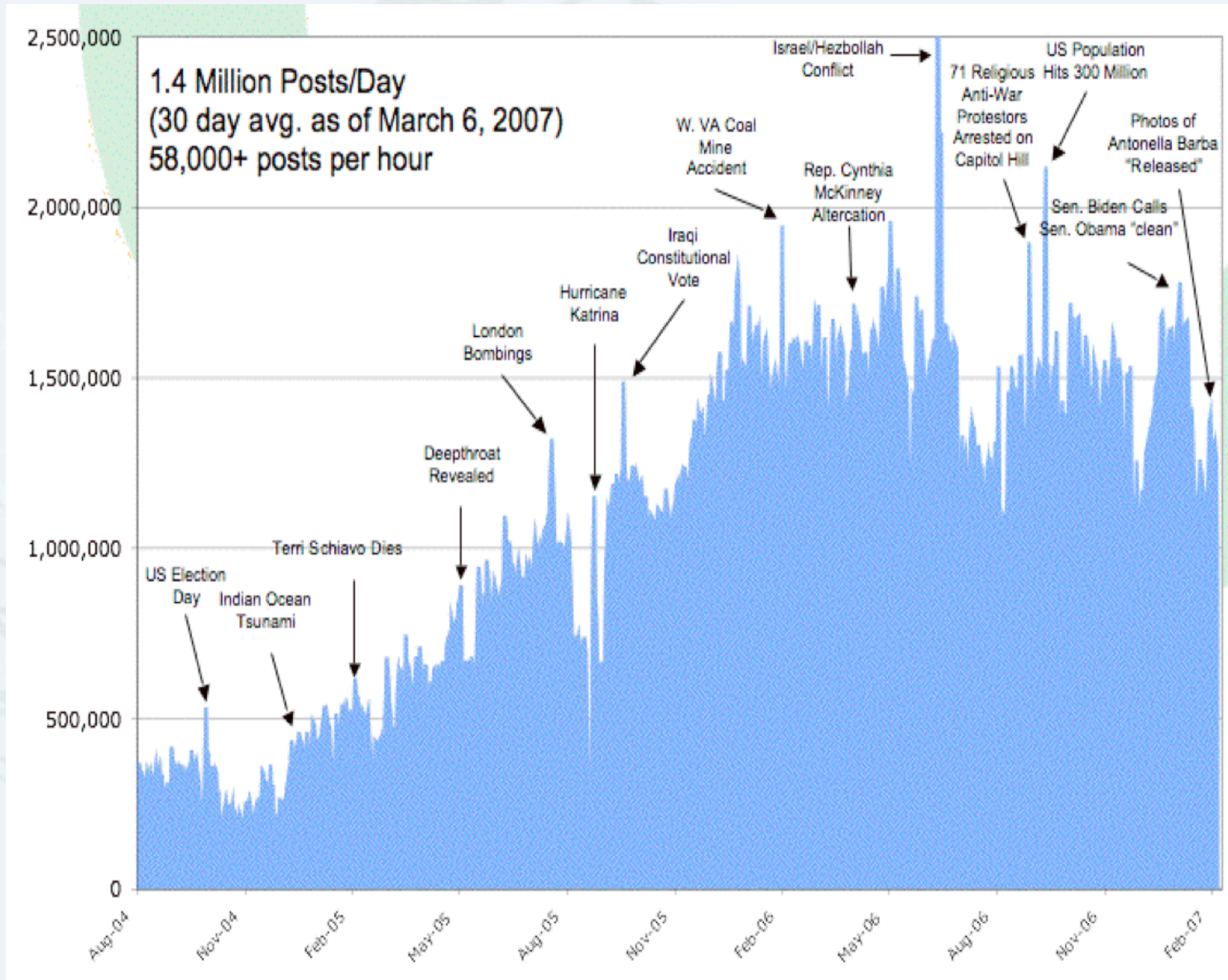
- Universal McCann (March 2008)
 - o 184 million WW have started a blog | 26.4 US
 - o 346 million WW read blogs | 60.3 US
 - o 77% of active Internet users read blogs

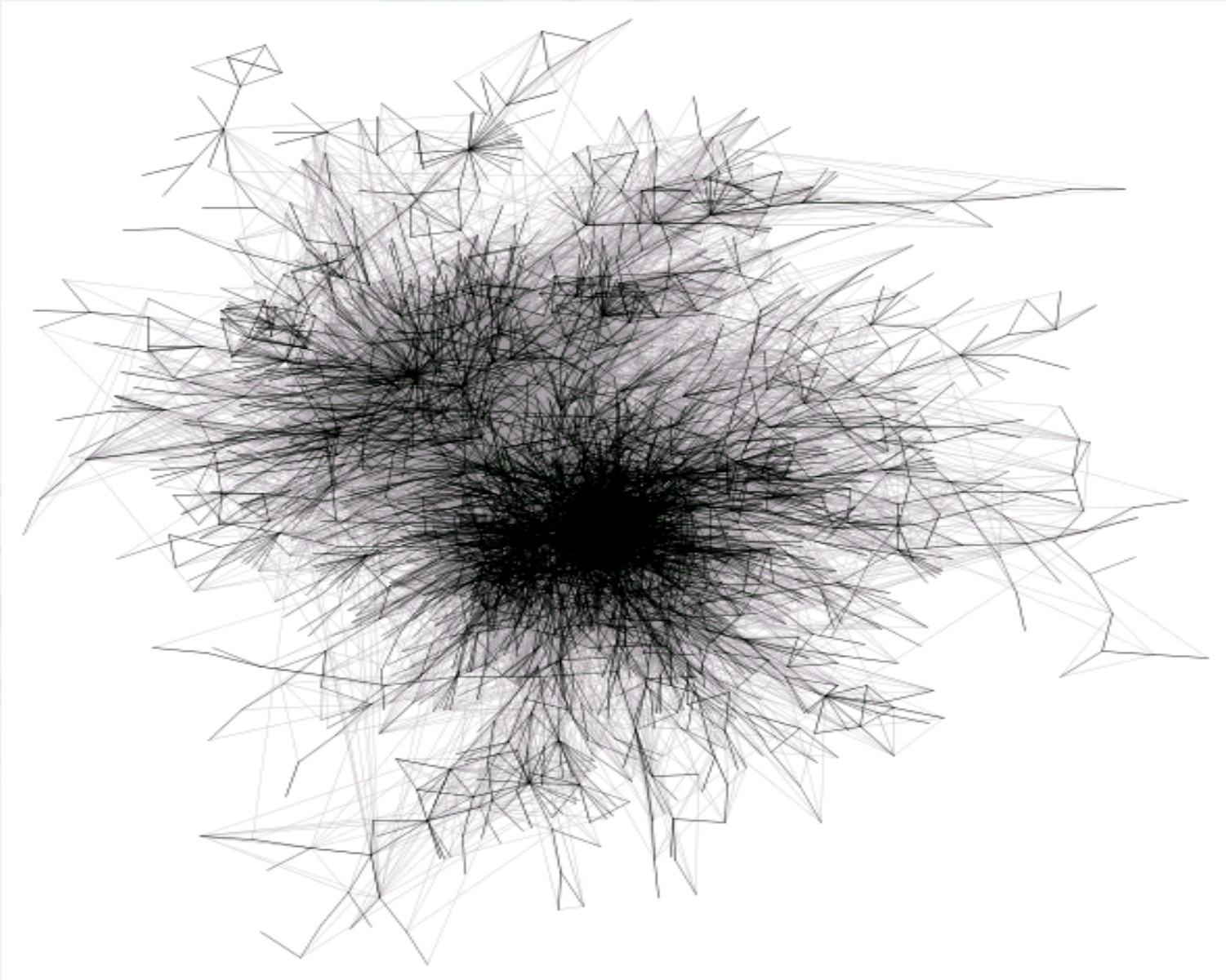
- o “Web log” or a blog is a Web site maintained usually by a single individual
- o Author publishes posts on various topics including commentaries, events, diaries
- o Information takes variety of forms such as text, video, audio, and/or image
- o Entries displayed in reverse-chronological order
- o Readers can comment on articles, subscribe to its feed, create track backs

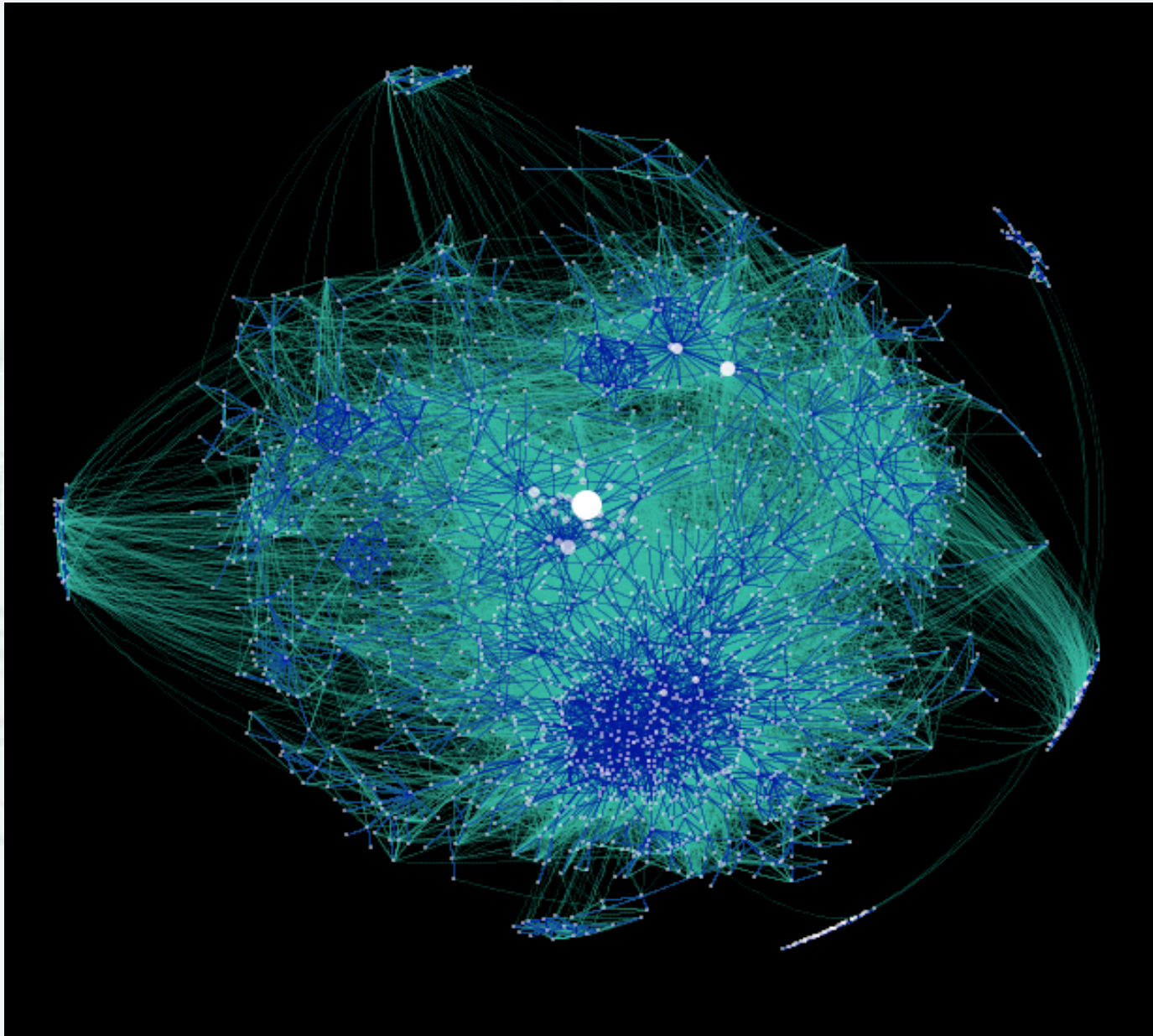
- o Blogs, inbound links, outbound links constitute the Blogosphere
- o Decent portion of Blogosphere consists of spam blogs or sblogs
- o Popularity tracked by Technorati, BlogPulse, Alexa, QuantCast
- o Search engines: Technorati, Google Blog Search, IceRocket
- o Blogging platforms: Blogger, WordPress, LiveJournal, Spaces

- o As of 8/2008, 133M blogs and growing
- o 1.4M blog posts per day









Source: Matthew Hurst

- o Blogosphere follows a power law
- o Other aliases: 80/20 and long tail
- o A/B/C breakdown





Metrics

Why

Usage

Examples

Challenges

Why

- To understand, monitor, and measure various aspects of SM
- Commercial / political spaces need metrics for their advantage
- Metrics are used to understand and target audience
- For Social Media Optimization (SMO v. SEO)

Usage

- Reach 1 to influence 1000 => Which one?
- Where to place product / brand ads
- How to propagate information
- Who to listen to
- Viral marketing
- Engagement: Interaction + Attention
- Reach
- Authority/Influence
- Loyalty

Examples

- Visitors (page views, visits)
- Time on Site
- Bounce Rate
- Inbound Links
- Comment Count
- Posting Frequency
- RSS Subscriber Count
- Votes

Challenges

- Lack of metric centrality
- Large data amount to monitor
- Unstructured underlying data
- Multi languages
- Unknown time zones
- Unknown / inaccurate demographics
- Unknown geolocations
- Dynamic environment (new v. dead resources and links)



Approaches

Sampling

Code Insertion

Link Analysis

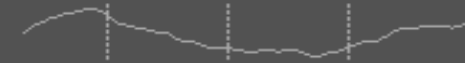
Aggregation

Semantic Analysis

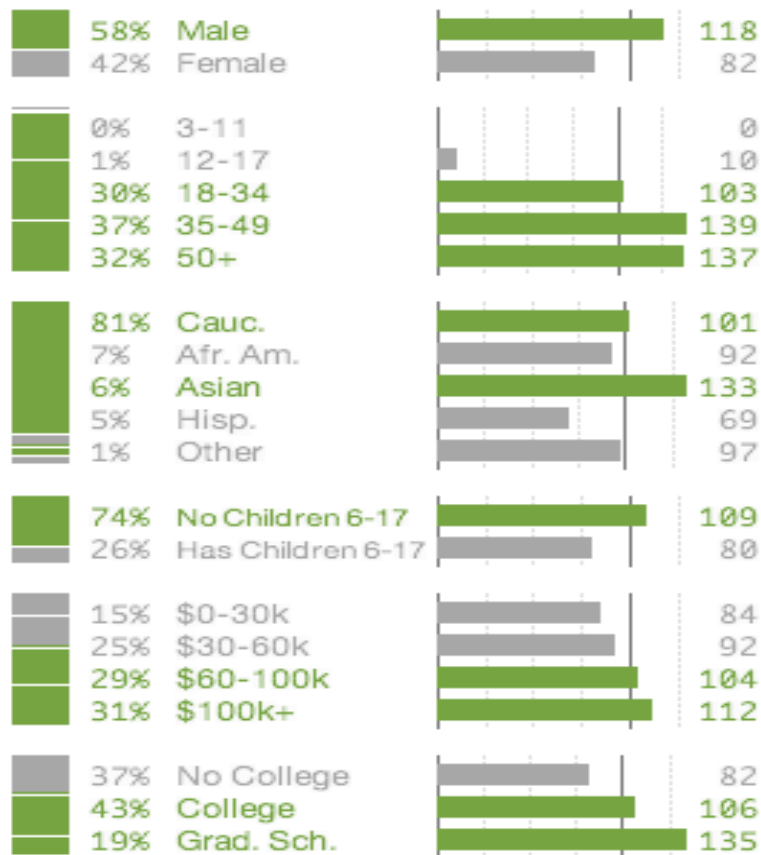
Sampling

Alexa, QuantCast

- ✓ Millions of subscribers
- ✓ Web surfing activities are monitored
- ✓ Gathered statistics are used to rank visited sites
- ✓ Data is correlated to rest of the web



US Demographics Sep 2008 100 index is internet average



Income represents total household income.
Index represents how a site's audience compares to the online internet population as a whole. An index of 100 indicates a site's audience is at parity with the total internet population.

[See Demographics](#)

Estimated Monthly Traffic

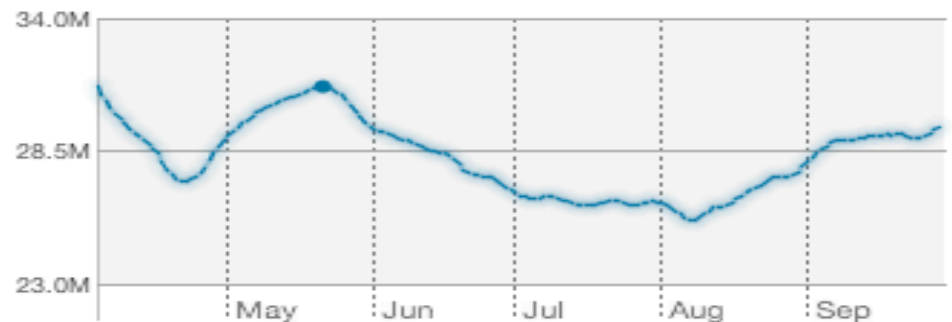
29.5M
Estimated US People

Estimated Monthly U.S. People

[Compare](#)

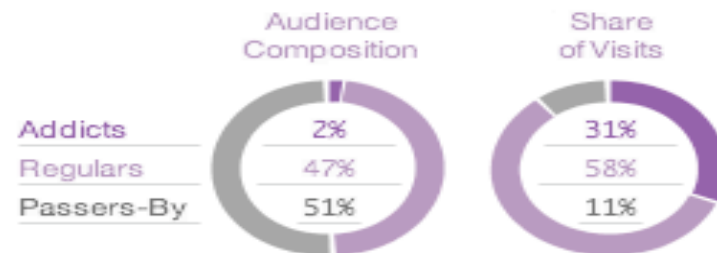
Monthly U.S. People
cnn.com

Panel Estimate 09/30/08



●US: 29.5M Max: 31.2M

Traffic Frequency



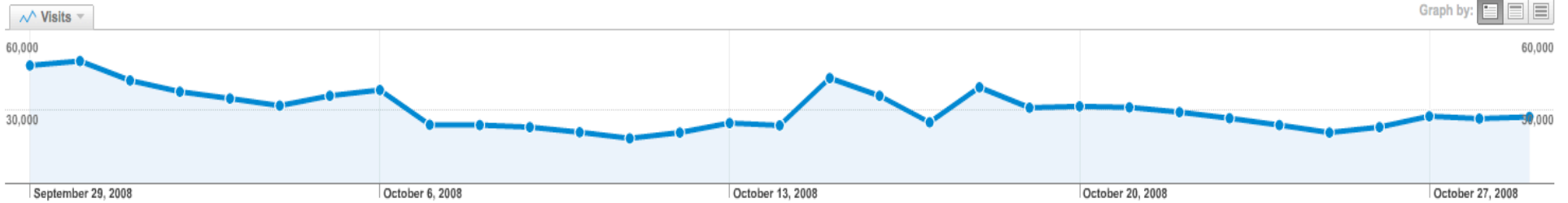
Code Insertion

Site Meter, Google Analytics

- ✓ Scripts inserted into page's HTML code
- ✓ Each page visit triggers script's execution
- ✓ Gathers stats: visit time, incoming IP, time on page....
- ✓ Gathered data sent to main server and stored

Dashboard

Sep 29, 2008 - Oct 29, 2008



Site Usage

939,255 Visits

8,770,443 Pageviews

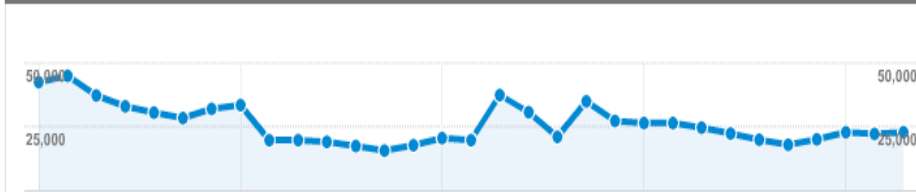
9.34 Pages/Visit

25.68% Bounce Rate

00:05:31 Avg. Time on Site

68.72% % New Visits

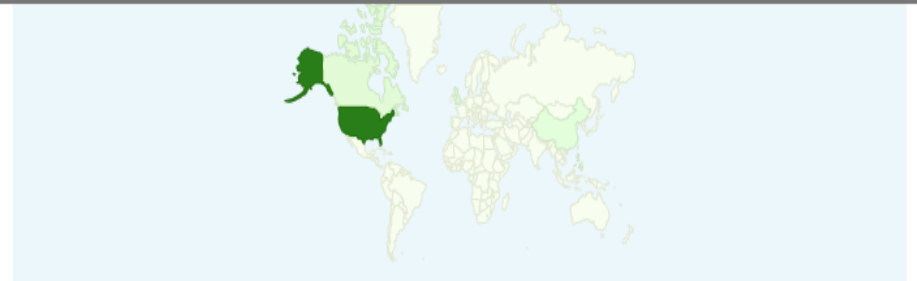
Visitors Overview



708,290 Visitors

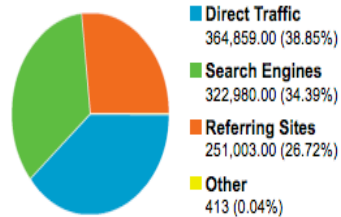
[view report](#)

Map Overlay



[view report](#)

Traffic Sources Overview



[view report](#)

Content Overview

Pages	Pageviews	% Pageviews
/	881,258	10.05%
/products/women/	425,361	4.85%
/c-4-Footwear.aspx?reqid=4&reqProdTypeld=4&subsectionname=f	228,316	2.60%
/products/men/	205,124	2.34%
/outlet.aspx	186,525	2.13%

[view report](#)

Link Analysis

Google BlogSearch, Technorati

- ✓ Rank sites according to inbound links count
- ✓ Can only consider number of incoming links
- ✓ Can also consider source of incoming links
- ✓ Options are incoming links to posts v. sites

“....Technorati Authority is the number of blogs linking to a website in the last six months. The higher the number, the more Technorati Authority the blog has.

It is important to note that we measure the number of blogs, rather than the number of links. So, if a blog links to your blog many times, it still only count as +1 toward your authority.

Of course, new links mean the +1 will last another 180 days....”

Semantic Analysis

Microsoft's Live Search (Neural Net)

- ✓ Designed to be trained with a small set of “good” results
- ✓ The training is used as a model to do ranking in general
- ✓ Supervised training is used to build ranking function
- ✓ 569 different properties to predict document relevancy
- ✓ Considers the patterns of human searches into account



Where

Do

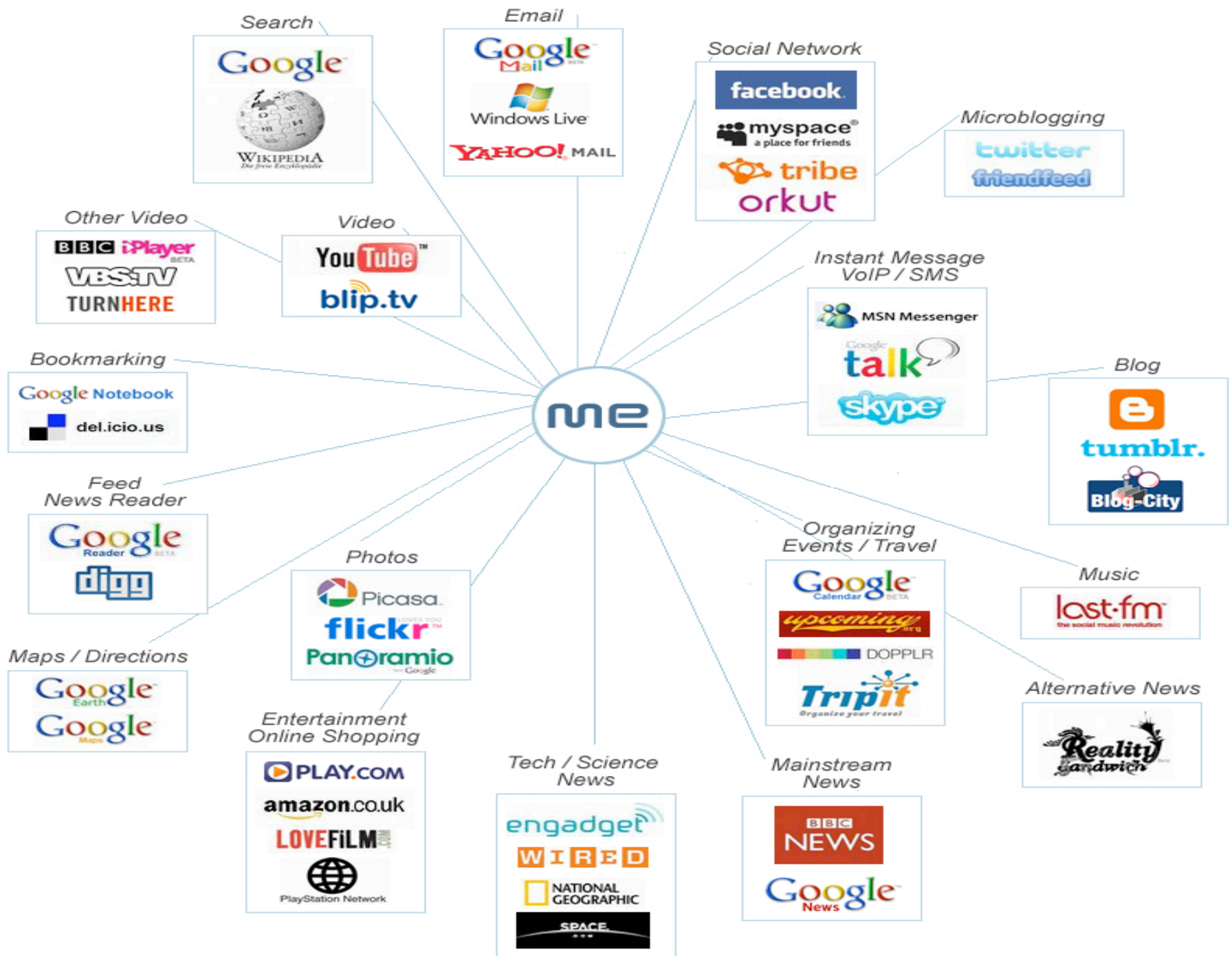
We

Go

From

Here

???



❖ Demise of print media:

“....In 2009, the Monitor will become the first nationally circulated newspaper to replace its daily print edition with its website; the 100 year-old news organization will also offer subscribers weekly print and daily e-mail editions....” (csmonitor.com)

❖ Rise of blogs:

“....4 of top 10 entertainment sites are blogs: OMG, TMZ, Asylum, PerezHilton...” (technorati.com)

❖ Dawn of Web 3.0:

- i. OpenID
- ii. Semantic Web
- iii. Open applications
- iv. Ubiquitous connectivity
- v. Network computing
- vi. Distributed databases