

# Searching the Social Network: Future of Internet Search?

---

*alessio.signorini@oneries.com*

# Alessio Signorini – Who am I?

I was born in Pisa, Italy and played professional soccer until a few years ago. No coffee or cigarettes for me.

In the past 3 years I was Director of Technology at Ask.com. Now I take care of Search at OneRiot.

I am also a PhD Candidate at University of Iowa. Published papers on Search Engines, Cryptography and Artificial Intelligence. Now working on Computational Epidemiology.

# Topics of Tonight

The exponential growth of the Internet made **search engines incredibly important**, but nowadays their job is much more complicated than just returning web pages.

In the last years **blogs and social networks boomed**. People rely on them for lots of reasons: find old friends, share pictures or videos, tell their lives, ...

**Combining search and social networks** can be extremely powerful if done correctly without overwhelming the users.

# Social Networks

If you have a job, a family, some friends or have recently been on a plane, **you are part of a social network.**

When I say “social network” people think to Facebook. In fact a social network is just **a structure of nodes** tied by some type of interdependency.

Each node represents a person or a company and **links are relationship**: family, friendship, business, conflict, ideas, ...

# Uses of Social Networks

In today's digital world, every email, phone call, Instant-Message, SMS, or even loading the same webpage creates a **connection between two users**.

Observing and studying the **evolution of those networks** is a very fascinating subject.

An important part of US Intelligence is based on **studying interactions graphs** of this kind (e.g. EELD - Evidence Extraction and Link Discovery).

# Social Network Services

**Social Networking Services** are built to facilitate the connection between people who share the same interests and/or activities.



LIVEJOURNAL



# Facebook Was Not the First!

**Usenet** (1980) and **BBS** (Bulletin Board Systems) were probably the first social network services. They encouraged discussions and chat around specific topics.

Other notable mentions go to:

- **The WELL** (1985)
- **Theglobe.com** (1994)
- **Geocities** (1994)
- **Tripod** (1995)

# Tripod - 1998



- About
- Help
- Search
- Sitemap
- Feedback

Member Services Map

Zone Map

What's Your Interest?

## It's January 21, 1998.

**TRIPOD: make it your home base on the Web. Check out the zones (chock full of content and resources), and join the pods (communities of interest).**

So far there are:

- **894,292** Tripod members
- over **1,400,000** member created pages
- **33** Pod communities

Image

**An Outbreak of Boss-itis**

It's not easy being boss. But how do you deal with one that's a royal pain? (1/19)



**Check out the newest pod!**

Be a businessperson, or just look like one.

Related Pods

Vote For Tripod - We Want t

We've been the Web -

## FEATURED ON TRIPOD

[Tripod Premium Membership](#): It's the perks, baby.

[Letters From Tripod](#): What it's like to be... us.

[Media Kit](#): Everything an advertiser needs to know.

[Tripod: The Book](#): We want YOU to participate.

[Tripod Member Panel](#): Tell us how to do our jobs!

Media

web/tech

[Bollywood: Cinema from India](#)  
[A Guy Confesses: I'm a Soap Fan](#)  
more

[Gif Girl on HTML Text](#)  
[Hands on DHTML](#)  
- more

Go to:

ITIA FREE ITIA F



# Geocities - 1999

**GeoCities**  
HOME PAGES & BEYOND

**ImGis.com**  
Your IS Resource

*your free home page is waiting*

[members' area](#) >>

Great ideas for building your page  
**BUILD NOW**

There's a person behind every page  
**MEET PEOPLE**

Save time and shop online  
**SHOP**

**GeoCities Web**  
 **Find**  
[Yellow Pages](#) [White Pages](#)  
[Domains](#) [Stocks](#) [Maps](#)  
**SEARCH**

**TODAY'S TOPIC** Monday, 25-Jan-99 02:44:08 PST



## DILBERT

Before finishing that interdepartmental status report, ask: "How would [Dilbert](#) handle this?"

**Your Daily Dose**  
[Kennedy](#) had the first presidential TV news conference on this date in 1961.

### Explore GeoAvenues, Home Pages and More

<a href="#">Arts &amp; Literature</a>	<a href="#">Entertainment</a>	<a href="#">Shopping Center</a>
<a href="#">Autos</a>	<a href="#">Family</a>	<a href="#">Society</a>
<a href="#">Business &amp; Money</a>	<a href="#">Health</a>	<a href="#">Sports &amp; Recreation</a>
<a href="#">Campus Life</a>	<a href="#">Home &amp; Living</a>	<a href="#">Travel</a>
<a href="#">Computers &amp; Technology</a>	<a href="#">People &amp; Chat</a>	<a href="#">Women</a>

Fly into [CapeCanaveral](#), one of our 41 [Neighborhoods](#).

### Get the Most Out of GeoCities

**Chat Events** — Looking for something to chat about? Check out our Chat Calendar for everything from skin care to war movie chats.

**Free Image Library** — Make your GeoCities home page sparkle with graphic pizzazz. Choose from more than 55,000 images!

**Plan your career**. Browse more than 150,000 jobs, meet people in your field and submit your resume to the GeoCities Career Planner.

### ABCNEWS.com

[Lewinsky Talks](#)  
[Olympic House-cleaning](#)  
[Jasper Trial Begins](#)  
[Papal Mass Draws 1 Million](#)  
[Privacy Groups to Boycott I...](#)

### Special Offers

Books  
**amazon.com**  
Computers  
**egghead AUCTION**  
Music/Videos  
**CDNOW**  
Credit Cards  
**9.99% APR Platinum**  
**Find your Dream Home**  
Compare homes for sale at the Home Advisor™ Buying Guide.  
**River Belle Casino**  
We've just paid out \$82,789! Click here to get lucky and win!

# Theglobe.com - 2000



theglobe.com

welcome

[register now](#)

02/29/2000

**Join or Start a globeclub** 

▶ Connect with friends and family

▶ Meet people with similar interests

▶ E-mail lists

**Build your  homepage**

 [shop.theglobe.com](#)

 **ELECTRONICS**

**Explore these clubs** 

Meet people who share your interests.  
Debate, flirt, ask questions, get support, speak your mind.

[Arts & Books](#)

[Auto](#)

[Business & Finance](#)

[Careers](#)

[Computers & Internet](#)

[Entertainment](#)

[Games](#)

[Health & Fitness](#)

[Hobbies & Pets](#)

[Humanities & Thought](#)

[News & Government](#)

[People & Lifestyles](#)

[Real Estate](#)

[Regions](#)

[Science & Tech](#)

[Society & Culture](#)

[Sports & Recreation](#)

[Travel](#)

[search clubs](#)

[services](#)

[build a homepage](#)

[forums](#)

[chat](#)

[check your email](#)

[help](#)

[marketplace](#)

[shopping](#)

[auctions](#)

[special offers](#)

[lowest airfares](#)

[hot gooey chat](#)

[get a dot com](#)

[sweepstakes!](#)

[make Free calls](#)

[Great Deals](#)

Search the Web 

 [quotes & indices](#)

sponsored by:

# The Bloggers Revolution

Before Twitter and Facebook, **blogs were growing exponentially**. Your very own online diary where to express your opinions, feelings, and make your life public.

Some used blogs to keep track of their lives, nice pictures, or to connect with each other. Others, **transformed blogs into newspaper**.

Some blogs became very popular (e.g. SlashDot) and some **bloggers are now celebrities** (e.g. Perez Hilton).

# Blogs: Some Numbers

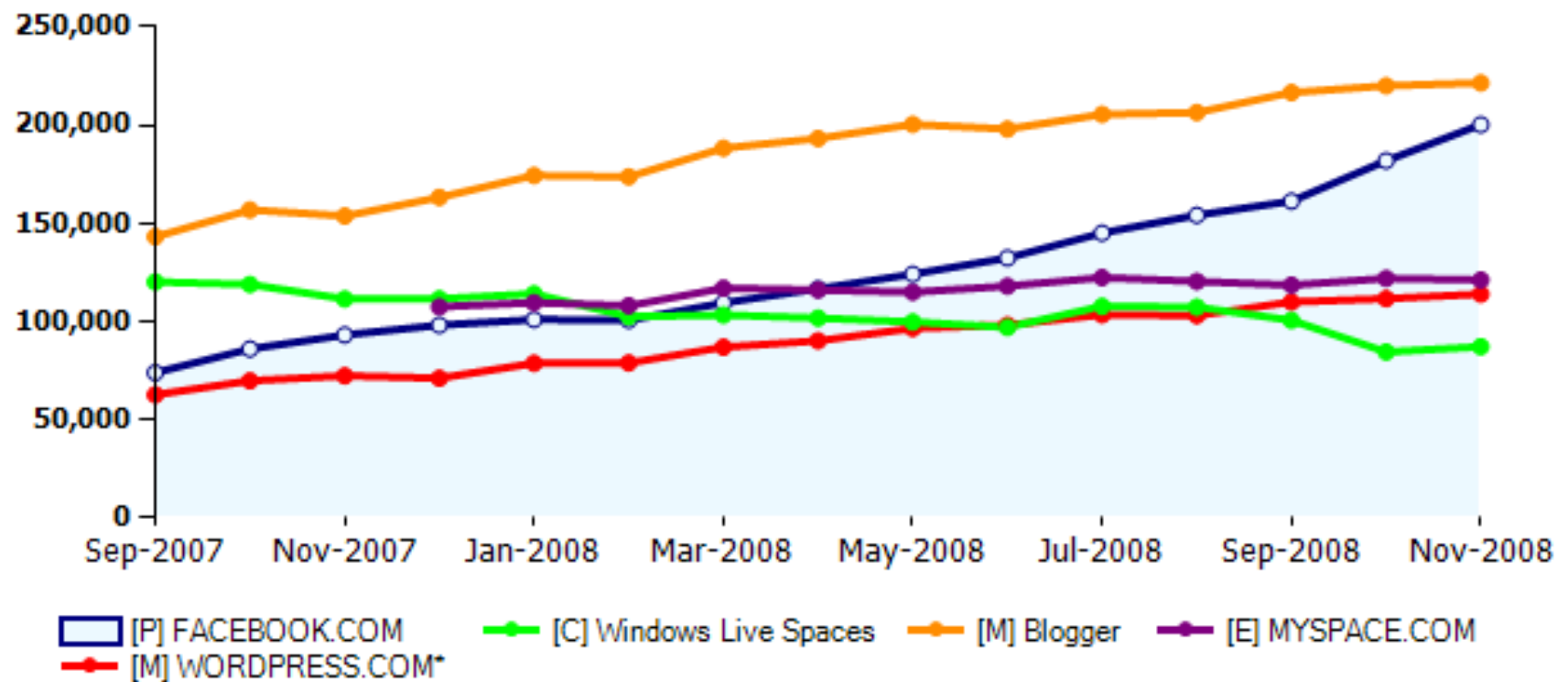
According to some recent statistics [thefuturebuzz.com] more than 900,000 blog posts are published daily.

A [comScore] study of March 2008 estimated more than 348 Million blog readers.

Quantcast reports more than 5.6 Million daily US visitors for Blogspot and 6.4 Million worldwide visitors for WordPress.

# Growth of Social Networks

Total Unique Visitors (000)



# Friendster



Friendster was founded in 2002 and launched in 2003. In a few months 3 Million users registered, today they are 90 Million.

Friendster was founded to create a safer, more effective environment for meeting new people by browsing user profiles and connecting to friends and friends of friends.

In 2003 declined a \$30 Million buyout offer from Google.  
Traffic declining: 7M visitors in Jan 2009, 5.4M in April.

# Friendster - 2003



## Friendster Tour

[Next](#) | [Sign Up](#) | [Testimonials](#) | [More Info](#)

Quickly create your profile

Welcome Jonathan!




[\[Edit Photos\]](#)

### My Profile: [\[Edit Profile\]](#)

Gender:	Male
Interested In:	Women
Status:	Single
Age:	30
Occupation:	Stunt Man
Location:	Mountain View, CA
Hometown:	Los Angeles
Interests:	<a href="#">Extreme sports</a> , <a href="#">golf</a> , <a href="#">wine</a> , <a href="#">travel</a> , <a href="#">books</a> , <a href="#">painting</a> , <a href="#">guitar</a> , <a href="#">sleeping</a>
Favorite Music:	<a href="#">Eric Clapton</a> , <a href="#">Madonna</a> , <a href="#">Weezer</a> , <a href="#">Miles Davis</a> , <a href="#">Metallica</a> , <a href="#">Jack Johnson</a>
Favorite Books:	<a href="#">About a Boy</a> , <a href="#">A Heartbreaking Work of Staggering Genius</a> , <a href="#">Beggars in Spain</a> , <a href="#">Stranger in a Strange Land</a> , <a href="#">The Last Guru</a>
Favorite TV Shows:	<a href="#">Ed</a> , <a href="#">Alias</a> , <a href="#">Law &amp; Order</a> , <a href="#">Undeclared</a> , <a href="#">Seinfeld</a> , <a href="#">The Simpsons</a> , <a href="#">Daria</a>

# Friendster - 2009



Help make ☺friendster. even better! [click here](#)

Please take a moment to complete this survey...

Search:  Web Friendster


powered by Google

Select Language: English

[Messages](#) | [Settings](#) | [Help](#) | [Log In](#)

[Home](#) [Profile](#) [Apps](#) [Connections](#) [Explore](#) [Search](#) [Classifieds](#) [Find Friends](#)

**josefa intac**





441W (0214).JPG

[Add as Friend](#) [Forward to Friend](#)

[Add Bookmark](#)

**Location:** Winnipeg, MB  
**Hometown:** Batangas, Philippines  
**Last Login:** 24 hours  
**josefa's URL:**  
<http://profiles.friendster.com/joahsark>

 **Private Profile**  
[Restricted to josefa's friends]

 **Messaging Off**  
[Restricted to josefa's friends]

**How you're connected:** [view all](#)

You → josefa is in your extended network → josefa



# MySpace



eUniverse saw the potential of Friendster and decided to launch its own. MySpace was **launched in August 2003**.

MySpace registered **100 Million accounts in August 2006**. Updated statistics are not available but in April 2009 it received 124 Million visitors.

In August 2006, Google signed a **\$900 Million advertisement deal** with MySpace.

# MySpace - 2004




MySpace lets you share photos, journals, and interests with your growing network of mutual friends. **Sign up for FREE!**

Take the MySpace Tour!

[Next](#) | [Sign Up](#)

## Step 1: Create your profile

dream



"Butt like a butterfly, sting like a bee"

32 years old female

Jersey City, NJ, United States

[View more pics](#)

dream is in your personal network

[You <> dream](#)

dream's Latest Journal Entry

Hey people ( [Click here to read more](#) )

[View All Journal Entries](#)

dream's Blurbs


About me:  
you'll find out when your good and ready

Who I'd like to meet:  
ummmmmmmmm...do we HAVE to meet?

dream's Friend Space


dream has 12 Friends.

Phil




[Online Now!](#)


Ruc




darlisa



Tom

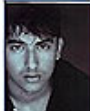


alfie




[Online Now!](#)


Benit



randy



HOY NEW YORKERS



[View All of dream's Friends](#)

dream's Details

Making Space for:	Dating, Serious Relationships, Friends, Activities Partners, Penpals, Hanging Out
Status:	Single

Contacting dream

<a href="#">send an email</a>	<a href="#">send a smile</a>	<a href="#">forward to friend</a>
<a href="#">add to friends</a>	<a href="#">add to favorites</a>	<a href="#">request review</a>

## Step 1: Create your **FREE** profile

Creating a profile in MySpace is fast, fun, and easy. You can disclose as much or as little information about yourself as you want. MySpace will always keep your personal information private and confidential. We will only show what you choose to share. You can even post pictures of yourself to add to the fun - ALL FREE!

[Next](#) | [Sign Up](#)

[About MySpace](#) | [Tour](#) | [News!](#) | [FAQ](#) | [Terms](#) | [Privacy](#) | [Contact MySpace](#)


# MySpace - 2009

[MyProfile.com](#)  
**Extras.com**  
ATER GRAPHICS  
FACE GRAPHICS  
FACE COMMENTS  
FACE LAYOUTS  
FACE ICONS  
FACE ANIMATIONS  
FACE BANNERS  
FACE GENERATORS

**Sponsored Links**  
**[Pimp My Profile ® \(Free\)](#)**  
Get Layouts, Backgrounds & Glitter: Download Now to Pimp Your Profile!  
[www.webfetti.com](#)  
**[STP Official Site](#)**  
Need Help With Basic Car Maintenance? Find Tips at STP.  
[www.STP.com](#)  
**[Get Showtime® & \\$25 Back](#)**  
The Best Series, Movies and Sports. Order Showtime & Get \$25 Cash Back!  
[OrderShowtime.com](#)


  
a place for friends.  
Home Browse People Find Friends Local Music Video More Andrea Log In



**Andrea**  
Female  
24 years old  
Altoona,  
Pennsylvania  
United States  
 Online Now!  
Last Login:  
5/11/2009  
**Mood:** on a mission.... 😊  
View My: [Pics](#) | [Videos](#) | [Playlists](#)


[Write me](#)  [Share me](#)  
[Like me](#) [Love me](#)  
[Talk to me](#) [Hate me](#)

[view more](#)

  
[www.pimp-my-profile.com](#)

**Andrea's Latest Blog Entry** [[Subscribe to this Blog](#)]  
Animated Skull ([view more](#))  
[[View All Blog Entries](#)]

**Andrea's Blurbs**  
**About me:**  
I am a single 24 year old mother of 4. Angelina Jaden 6, Alexis Mar Xavier 4, Alivia Glana-Annabelle 2 months. Life is crazy, life is hard kids. I love tattoos, music, stand-up comedies, crazy movies, ect...



# Facebook



Mark Zuckerberg created **Facemash** in **October 2003**, but was shut down by Harvard administration.

Facebook was launched in February 2004. They now have more than **200 Million users registered**, half of which log in daily.

Facebook now offers an **open login platform**, an internal **advertising network** and a **public API**.

# Specialized Social Networks Services



LinkedIn is a social network created to share ideas and business opportunities. It is a Facebook for grown ups.

Epernicus is a social network for scientists build to efficiently connect researches and researches.



Flixster is a social network for movie lovers. Rating movies you like connects you with similar users and suggests new movies.

# Specialized Social Networks Services

**Ning** Create Your Own Social Network

[Sign Up](#) [Sign In](#)



**Ning** lets you discover and create new social networks for your interests and passions.

[Sign Up](#) to create a new social network

...or [Sign In](#)

Discover New Social Networks



[English \(U.S.\)](#)

Copyright © 2009 Ning, Inc. [About Ning](#) [Legal](#) [Resources](#)

# Search Engines = Personal Concierges?

Do you remember **Altavista**? It was one of the first web search engines. It relied purely on **keyword matching**.

Today, **we expect more than pages** from Search Engines:

- **Maps and Directions** (e.g. “Boulder, CO”)
- **Money Conversion** (e.g. “3USD in EUR”)
- **Weather or Time Information** (e.g. “Weather in Paris”)
- **Flights Information** (e.g. “US 1264”)
- **Translation Services** (e.g. “dog in french”)

# Question Answering

Lots of people believe that search engines are actually **super-computers which know everything**, and they ask them questions (sometimes even politely)!

Among **my favorites queries** there are:

- Could you please help me find a good book?
- What is the mpg of the car in my garage?
- Am I pregnant?
- Why my iPod does not turn on anymore?



# Your Answer might Already be Out There

Nowadays, more than 83% of US population has Internet access. Most of them write or read a blog, login into a social network, uses email or chat.

If you have any question or problem, there is an high probability that some other Internet user knows the answer and published it somewhere on the web.

Many companies believed in this model and started building services around it.

# Google Questions & Answers

In August 2001 Google launched a services which allowed people to ask Google's employees to do a search for them and provide the results through email (for \$3).

They had to close one day later. A proper service was launched in May 2003 and each answer would cost between \$2 and \$200.

They were using the free time of their editors to do searches for users. The service closed in 2006.

# Yahoo! Answers

The **most popular service is Yahoo! Answers**. The folks at Yahoo! tapped into their massive community of users and created a very polished and easy to use service.

Yahoo! Answers has more than **90 Million users worldwide** (21M in US) and receives more than 1.8 Million visitors per day.

It is available in many languages and features experts in various categories. People are **ranked based on their answers** which are voted by the community.

# Yahoo! Answers


[Yahoo!](#) | [My Yahoo!](#) | [Mail](#) | [More](#) ▾

Make!My Homepage

[New User? Sign Up](#) | [Sign In](#) | [Help](#)


**YAHOO! ANSWERS**


[WEB SEARCH](#)

**ask.** 


Can't find it with search? **Ask**

[Post Question](#)


**answer.** 

**Share knowledge  
Help others  
Earn points**

**What people think of Answers  
How does it work?**

**discover.** 

**Stop burning your  
biscuits:  
food & drink**

  
[Explore category](#)

Search for questions:

[Search](#) [Advanced](#)

[My Profile](#)

**Ready to Participate?  
Get Started!**

**Categories**

- [Best of Answers](#)
- [Arts & Humanities](#)
- [Beauty & Style](#)
- [Business & Finance](#)
- [Cars & Transportation](#)
- [Computers & Internet](#)

**Yahoo! Answers Blog**    

**Rating Reminder**

We all love being able to ask questions and get answers about any topic and share knowledge and experiences with a simple click. Whether you're answering questions or just browsing, there are many ways to contribute to improving the content on Answers.

[Read the Answers Blog >](#)

[«](#) [||](#) [»](#)

**Answer Questions**

[Recent](#) [Popular](#)

# Other Question Answering Services

 Answers.com was the first question answering site. Receives daily about 3 Million of users worldwide.

AllExperts is a directory of experts in various fields to whom you can ask questions.



ChaCha searches the Internet for the answer you need and provides it to you. It is targeted to mobile phones.

# Social Network + Questions = Answers!

Question Answering Services can work great if combined with the appropriate Social Network.

- **Would you like a pizza?** Ask to your friends or their Italian friends for the best pie in the area.
- **Problems installing Linux?** Ask to your geek friends or some other expert.
- **Building an Aquarium?** Ask to the local aquarium community for advices and tips.

# Artificial Intelligence to Avoid Overload

But what if your phone **keeps ringing at every message**?  
You would be leaving the service very soon.

To avoid that, use some artificial intelligence. For example, **classify the questions** (e.g. sports, restaurant, linux, ...) and direct them only to whom is likely to answer.

You could also **learn when** people are more likely to answer (e.g. calling me at 2am is not a good idea).

# Implicit Social Networks

Unfortunately, for some questions **you might not have the right friends** on Facebook. What to do then?

Looking at your searches and your browser history it is possible to “classify” you and **find other similar users**.

For a movie to watch tonight, what people similar to you liked could be a good start. At the same time, Linux users out there could **easily answer your tech questions**.



# Concerned about your Privacy?

Studying your searches and your browsing history **might seem evil**, but it is not that far from what search engines are already doing today.

Using your IP address is easy to **track down your location** and the type of connection that you use.

When you log into Gmail, a cookies is set on your browser and **your visits are tracked** on every page which contains Google Ads.

# Ranking URLs

Traditionally, search results ranking is done **mixing relevance** to the query and **link popularity** of the URL.

In simple terms, a page which is **linked by many other pages** (especially if they are authoritative) is more popular and probably more important.

This is the **foundation of PageRank** (Google's algorithm) and also of practically all the other ranking algorithms.

# Ranking URLs using Social Networks

If you are able to identify similar users, you could build a search engine which ranks the URLs considering the **popularity among the group**.

Then you could match the user (or its query) to the appropriate group, find out relevant pages, and **return just the group's favorites**.

**Ambiguous queries are difficult:** if I type “apache” do I mean the population, the helicopter or the web server?

# Ranking Shared URLs

Lots of Social Networks (e.g. Facebook, Twitter, Digg) offer the possibility of **sharing and/or voting URLs**.

When people think a video is cool or a page important, they **publish it** on one of those services.

Every day **Digg** receives 9,000 submissions and 200K diggs. **Twitter** receives 10M tweets and about 1.5M URLs. On **Facebook** are shared about 1M URLs per day.

# Resolving & Downloading

Obtaining the URLs instead of **having to discover them** (through crawling) is very nice, but lots of work still needs to be done.

For example, many of those **URLs are shortened** (e.g. <http://bit.ly/AhYi>) and require to be expanded.

The expanded **URL needs to be crawled** and the system needs to avoid to re-crawl the same page multiple times to avoid to waste resources.

# De-chroming

By John Zarrella  
CNN

**KENNEDY SPACE CENTER, Florida (CNN)** -- With space shuttle Atlantis on its way to upgrade the Hubble Space Telescope, that leaves one shuttle, Endeavour, at the ready on the other launch pad here.



Space shuttle Endeavour, in background, sits poised to launch in case Atlantis, in foreground, meets trouble.

And that's where everybody at NASA wants it to stay.

"We have high confidence that we're just having that thing over on pad B to make it look nice," said mission flight director Tony Ceccaci.

Endeavour is far more than a postcard picture -- it's on standby in case something goes seriously wrong with the Atlantis mission. For example, NASA has estimated there's a 1-in-221 chance the shuttle could be struck and crippled by orbiting [space debris](#).

If such an incident were to thrust the seven [Atlantis](#) astronauts into danger, Endeavour would blast off within days in a last-ditch attempt to save them.

Endeavour Cmdr. Christopher Ferguson believes

his crew is up to the task.

"I feel as confident about our ability to pull this off, if need be, as I would any other mission," he said.

The idea of a standby rescue spacecraft grew out of the 2003 Columbia disaster, in which that shuttle disintegrated over Texas during re-entry into the Earth's atmosphere.

But until now, there's been no need. All the previous missions since Columbia have gone to the International Space Station. The crew of a damaged shuttle could take refuge there for months if need be until another shuttle was readied to retrieve them.

But Atlantis is going to Hubble, whose orbit isn't anywhere near the space station. If their spacecraft malfunctioned, the seven-member Atlantis crew would have to survive aboard -- probably for days -- until they could be rescued.

"In this case it's a much shorter fuse," Ferguson said. "Consumables like food and oxygen would run out quickly. So the reason we've gotten this crew trained and spooled up and ready to go on literally a moment's notice is because they [Atlantis's crew] have no place comfortable to go stay for a long period of time."

TEXT SIZE - +

## Most Popular on CNN

### STORIES

Most Viewed

Most Emailed

Top Searches

- 1 [Michael Landon's son found dead](#)
- 2 [Beauty queen's fate up in air](#)
- 3 [Priest admits 2-year affair](#)
- 4 [Possible space rescue readied](#)
- 5 [5 killed in Baghdad](#)
- 6 [Jordan and Peter Andre separate](#)
- 7 [Demjanjuk deported to Germany](#)
- 8 [Million child prostitutes in India](#)
- 9 [What release means for Iran, U.S.](#)
- 10 [Airline defends pilot's training](#)

[more most popular >](#)

### VIDEOS

### TOPICS

from **cnet**

Today's Featured Product:



Samsung Tint  
(MetroPCS)

[Check Prices](#)  
[Read Review](#)  
6.0 out of 10

# Context Analysis

## Commander's Ouster Is Tied to Shift in Afghan War

By ELISABETH BUMILLER and THOM SHANKER  
Published: May 11, 2009

WASHINGTON — The top American commander in [Afghanistan](#), Gen. [David D. McKiernan](#), was forced out Monday in an abrupt shake-up intended to bring a more aggressive and innovative approach to a worsening seven-year war.

[Enlarge This Image](#)



Haraz N. Ghanbari/AP, left; Helene C. Stikkel/Dept. of Defense, via Reuters  
Gen. David McKiernan, left, spoke at a news conference in Washington in October. Lt. Gen. Stanley A. McChrystal, right, in Washington in 2003.

### Related

Times Topics: [Stanley A. McChrystal](#) | [Afghanistan](#)

[Enlarge This Image](#)



Defense Secretary [Robert M. Gates](#)

[announced the decision](#) in terse

comments at the Pentagon, saying that “fresh eyes were needed” and that “a new approach was probably in our best interest.” When asked if the dismissal ended the general’s military career, Mr. Gates replied, “Probably.”

The move reflects a belief that the war in Afghanistan, waged against an increasingly strong [Taliban](#) and its supporters across a rugged, sprawling country, is growing ever more complex. Defense Department officials said General McKiernan, a respected career armor officer, had been removed primarily because he had brought too conventional an approach to the challenge.

He is to be replaced by Lt. Gen. [Stanley A. McChrystal](#), a former commander of the Joint [Special Operations Command](#). He served in Afghanistan as chief of staff of military operations in 2001 and 2002 and recently ran all

COMMENTS  
(144)

[SIGN IN TO E-MAIL](#)

PRINT

REPRINTS

SHARE

[Next Article in World \(1 of 23\) »](#)

### Politics E-Mail



Keep up with the latest news from Washington with the daily Politics e-mail newsletter.

[See Sample](#) | [Privacy Policy](#)

[Sign Up](#)

[Advertise on NYTimes.com](#)

### MOST POPULAR

[E-MAILED](#) [BLOGGED](#) [SEARCHED](#)

1. [A Cautionary Video About America's 'Stuff'](#)
2. [Backlash: Women Bullying Women at Work](#)
3. [Frugal Portland](#)
4. [Paul Krugman: Harry, Louise and Barack](#)
5. [For Many Teachers, a Famously Fertile Market Dries Up Overnight](#)
6. [Prague Journal: Hard Times Give New Life to Prague's Golem](#)
7. [China Outpaces U.S. in Cleaner Coal-Fired Plants](#)
8. [Cyberwar: Cadets Trade the Trenches for Firewalls](#)
9. [Europe Imagines Its Suburbs Without the Car](#)
10. [State of the Art: Wi-Fi to Go, No Cafe Needed](#)

[Go to Complete List »](#)

### Living with Asthma

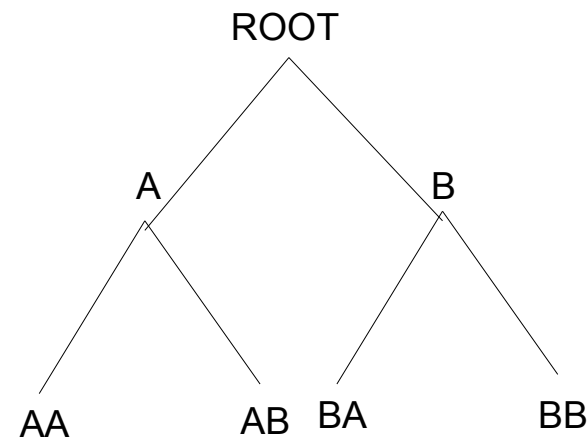
[ALSO IN VIDEO »](#)

# Indexing

During indexing, the system creates an **inverted map between words and documents** in which they appear.

Scanning a list of millions of words is **not efficient**. For this reason indexes are organized as trees.

Still, constantly modifying structures and writing on disk is **very expensive**.





# Handling Signals in Real-Time

Pages also need to be **classified**, images and videos **extracted**, spam and duplicates **removed**, ...

In addition, the **real-time signals** (e.g. number of Diggs or tweets) cannot easily be written on a database since the heavy load would make it collapse very quickly.

And while you are crawling, parsing, indexing, and classifying, you also need to **provide search capabilities** to your users.

# Many Real-Time Engines are just Filters

Many new “Real-Time Search Engines” are in fact just **filters** and often limit their matching to the tweet or title of the page to avoid the heavy lifting.

You can **think about them as big pipes**: information flows into the pipe and if matches your keywords gets redirected to your browser.

<http://obama.collecta.com/>

<http://www.scoopler.com/>

# Search the Real-Time Web at OneRiot



Search the realtime web

Web Video

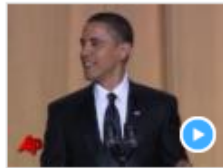
obama

Search

**Trending Topics:** Star Trek, RealWizkhalifa, Goodnight, Canucks, Denver, Dallas, Nuggets, Dirk, Gossip Girl

Realtime results: **Tuesday, May 12 at 12:16am**

Sort by: **Realtime** | Pulse



**YouTube - Obama Pokes Fun at Washington, Himself**

2:08 mins  
YouTube

Last shared: 1 minute ago | 2185 shares (expand) ▼ | Tweet it

First shared on Digg by bluehard

**Barack Obama and Miss California Share More than Over-Hype | Comedy Central I...**

So, what do President Barack **Obama** and Miss USA runner-up Carrie Prejean have in common with one another? (Okay,... Disclaimer The opinions...

<http://ccinsider.comedycentral.com/2009/05/05/barack-obama-and-miss-california-share-m...>

Last shared: 2 minutes ago | 4 shares (expand) ▼ | Tweet it

First shared on Digg by novenator

**Obama tax proposal would hit securities dealers, life insurance firms, big es...**

The **Obama** administration is digging deep into the tax code to find ways to help pay for its ambitious plans to cut taxes for middle-income families and...

[http://www.latimes.com/business/la-fi-taxes12-2009may12\\_0,6632793.story](http://www.latimes.com/business/la-fi-taxes12-2009may12_0,6632793.story)

**Most shared today**

**Zachary Quinto's fingers  
glued together for S...**  
[news.com.au](http://news.com.au)



First shared by BuzzEdition

**Could There Be a Better NBA  
Finals Matchup Th...**  
[coedmagazine.com](http://coedmagazine.com)



First shared by TheKilimanjaro

**Christians dominate in  
'American Idol'**  
[worldnetdaily.com](http://worldnetdaily.com)



First shared by MightyRighty

**Farrah Fawcett tells of  
cancer battle, blasts...**

# How to Be Successful in Search

Inspired you to build a **new social search engine**? Here is a recap of what you should keep in mind:

- Find as many good pages as you can
- Be fast in crawling, parsing and indexing
- Study your users and identify groups
- Classify queries and assign to the right group
- Provide instantaneous and gratifying feedback
- Create an API to interact with your platform