

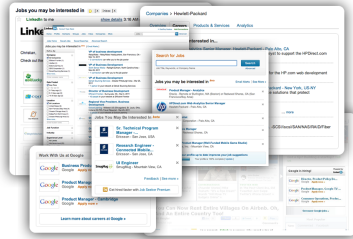


# Content, Connections, and Context

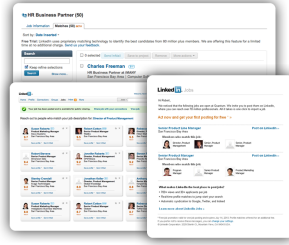
**Daniel Tunkelang**  
Principal Data Scientist at LinkedIn

# Recommendation Products at LinkedIn

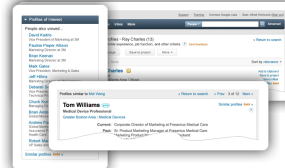
## Jobs You May Be Interested In



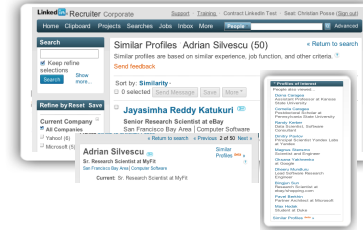
## Talent Match



## CAP

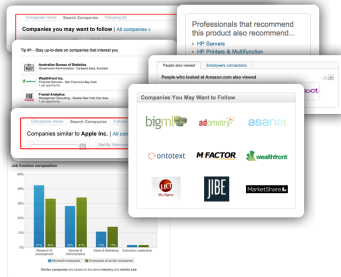


## Similar Profiles

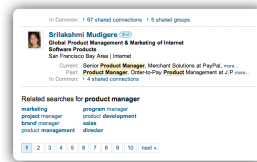


## Companies

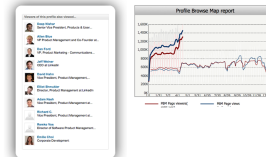
Recommendations, similar companies search, peer companies, and company browse maps, company products and services browse maps



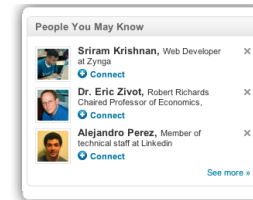
## Related search



## Profile browse maps

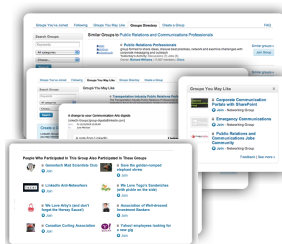


## Connections

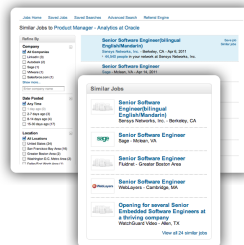


## Groups

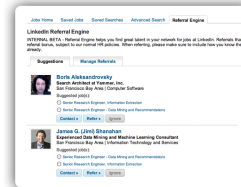
Recommendations, similar groups search



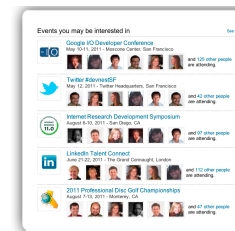
## Similar jobs



## Referral Engine



## Events You May Be Interested In



## Ad matching engine

$$pCTR = f(\text{member, creative, advertiser, context, inventory, CTR})$$

## News

LinkedIn Today: See all Top Headlines for You

Meine Machine: SOPA Edition

"No" is the New "Yes": Four Practices to Reprimize Your

Yahoo Co-Founder Jerry Yang Resigns



More than 50%




# Recommendations drive:

> 50% of connections

> 50% of job applications



> 50% of group joins

**People You May Know** *updated*

-  **Sophie Perisic**, French Teacher at Crystal Springs Uplands School  
[Connect](#)
-  **Lissa Juan**, Senior Manager, Product Management - Mobile &  
[Connect](#)
-  **Jeff King**, Sr. Director X.commerce Platform Partnerships at eBay  
[Connect](#)




[See more »](#)

**Jobs You May Be Interested In** *beta*

-  **Director of Engineering**  
Pandora - Oakland, CA
-  **Director of Engineering for Core...**  
CyberCoders - Santa Clara, CA
-  **Director of Engineering**  
HotelTonight - San Francisco Bay...

[Feedback](#) | [See more »](#)

**Groups You May Like**

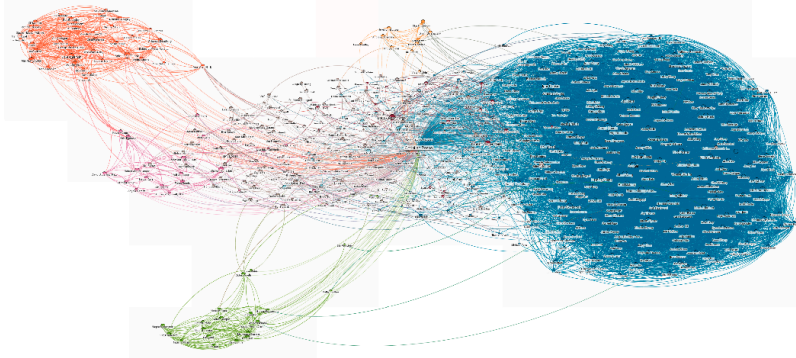
-  **LinkedIn Groups Product Forum**  
[Join - Other...](#)
-  **RDataMining**  
[Join - Professional Group](#)
-  **Data Mining, Statistics, and Data Visualization**  
[Join - Professional Group](#)

[Feedback](#) | [See more »](#)

# Inputs for Recommender Systems

## Social Graph

## Content



## Behavior



**Applied Researcher / Data Scientist: Data Analysis, Data Mining and Machine Learning**  
 LinkedIn - Mountain View, CA (San Francisco Bay Area)

**Job Description**  
 The engineering culture at LinkedIn is based on building and integrating cutting-edge technologies while encouraging creativity, innovation and exploration. Our engineers constantly raise the bar for excellence, motivating each other to tackle challenges and take intelligent risks. The industry is moving fast and our engineers are right there with it.  
<http://engineering.linkedin.com/>

**About Data Analysis, Data Mining and Machine Learning**  
 We're looking for an Applied Researcher / Scientist to work on our massive semi-structured text and graph datasets. This role is ideal for a senior researcher who has both domain experience in data mining, information retrieval, or machine learning and a strong systems orientation. It offers the unique opportunity to conduct applied research and have a huge end-to-end impact on key LinkedIn product initiatives, including search relevance, ad targeting, information extraction, and recommendations.

- Desired Skills & Experience**
- You are a computer science drump with strong coding ability and familiarity with Java, C++ or any other OOP language.
  - You have worked with data mining toolkits like Weka, Mahout, H, NTK, etc.
  - You have experience with Hadoop or other MapReduce paradigms.
  - You're familiar with information retrieval libraries like Lucene / SOLR.
  - You're interested in how data connections influence ad relevance for individual evolution. Ad relevance requirements:
  - You have put
    - SS, MS, or Ph.D.
    - You thrive in
    - You enjoy it

**Visual Analytics**  
 Discussions Members Promotions Jobs Search Manage More...

**Company Desc**  
 Founded in 2003, LinkedIn is successful with our company. LinkedIn is identified business bring solutions. Has

**Additional Info**  
 Post: **Most Popular Discussions**  
 Experience: **Here is a nice visualization of timeline events - an impressive interactive design.**  
 Functions: **Arab spring: an interactive timeline of Middle East protests**  
 Industries: **Ever since a riot in Tunisia burst into a push in December 2010 in protest of the treatment of public, an upsurge in violence has engulfed across the Middle East. Our interactive timeline traces key...**  
 Job ID: **George Ingram (1) · Jeffrey, great presentation, we are now in 2012. Is there a company you can draw us to start with the Arab culture desk?**

**LINKEDIN TODAY JAN 18**  
 Defend Our Freedom to Share (or Why SOPA is a Bad Idea)

**Principal Data Scientist and Product Manager**  
 LinkedIn (7)  
 Public Company | 1001-5000 employees | LNKD | July 2008 - Present (2 years 7 months) | Me  
 Design and develop from ground up data infrastructure to the user experience, wor

**1. Highly scalable real-time content-base built on top of Hadoop, Lucene and Viole**  
 - Personalized recommendations: Jobs 'Companies You May Want to Follow', 'Jobs For You' Web Ads, 'Events You May Be Interested In'  
 - Similar recommendations for jobs, groups, companies, industries, profiles and events  
 - Matching solutions for our recruiting and marketing products: 'Identification', 'Relevant Engineer', 'Job Applicant Scoring', 'Ad Targeting Engine'  
 - Crowd based recommendations: 'Related Searches', 'Viewers of This Profile Also Viewed', 'People Who Viewed This Job Also Viewed', 'Professional That Recommend This Product Also Recommend', 'People Who Viewed This Company Also Viewed', 'People Also Explored [this group]'

**LinkedIn Today**  
 In the press: <http://latu.com/vmmpaa>, <http://latu.com/qyq2>, <http://latu.com/bfvy>

**Your Groups (48) Reorder**

- Visual Analytics**  
 48 members
- The R Project for Statistical Computing**  
 3 members
- Data Mining, Statistics, and Data Visualization**  
 2 members
- Online Experimentation Professionals**  
 1 member



# Take-Aways

Content is king.

Connections provide social dimension.

Context determines where and when a recommendation is appropriate.

# What is the goal of recommendations?

- Growth
- Engagement
- Revenue



# What is the goal of recommendations?

- Growth
- Engagement
- Revenue



# Users First!

Provide relevant content

and establish social connections

in appropriate context.

# Content is King





# Goal-Seeking is about Content



# Content (Ir)relevance

- No right answer, but many wrong answers.

%no

[Voorhees, 2004]

<http://bit.ly/percentno>

WTF!

@k

<http://bit.ly/wtfatk>

# Example: Related Searches

The screenshot shows the LinkedIn search interface. At the top, the LinkedIn logo and navigation menu are visible. The search bar contains the text "web developer". Below the search bar, there are two profile cards for "Sajid Momin" and "Monique Andrews". Below these profiles, a red-bordered box highlights the "Related searches for web developer" section, which lists several related job titles and technologies in two columns.

LinkedIn Go to LinkedIn Recruiter » Daniel Tunkelang ▾

Home Profile Contacts Groups Jobs Inbox Companies News More People ▾ web developer

Groups: MOO Crew  
▶ 1 shared connection · Similar

**Sajid Momin** (2nd)  
Sr. Consultant/Partner at WebltUp Solutions  
San Francisco Bay Area · Internet  
197 connections · 6 recommendations

Current: Sr. **Web Developer**/Partner at ... more  
Past: Sr. **Web Developer**/Consultant at ... more  
Groups: Web Content Management Systems · ... more  
Similar

**Monique Andrews** (2nd)  
**Web Developer** at Shopping.com, an eBay company  
Phoenix, Arizona Area · Internet  
332 connections · 5 recommendations

Current: **Web Developer** at Shopping.com, an ... more  
Past: Internet/Software **Developer** and ... more  
Groups: eBay Inc. Employees · USWeb Alumni · San ... more  
▶ 1 shared connection · Similar

**Related searches for web developer**

web designer	php
ruby on rails	developer
web development	graphic designer
software engineer	lamp

# Collaborative Filtering as Content Signal

- Use temporal locality within sessions.
- Find queries with clicks on similar results.
- Look for query overlap.
- Learn more at CIKM! [Reda et al, 2012]

# Example: Jobs You Might Be Interested In

**Daniel Tunkelang**



Principal Data Scientist at LinkedIn




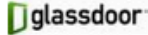
Mountain View, California | Internet

**Current** LinkedIn, New York CTO Club  
**Previous** Google, Endeca, Context Integration  
**Education** PhD, Computer Science at Carnegie Mellon University

- Information Retrieval
- Data Mining
- Recommender Systems
- Machine Learning
- Search
- Patents
- Algorithms
- Natural Language Processing
- Text Mining
- Information Extraction
- Text Classification
- Text Analytics
- Collaborative Filtering
- Analytics
- Information Visualization
- Web Mining
- Human-computer Interaction
- Predictive Analytics
- SEO
- Blogging



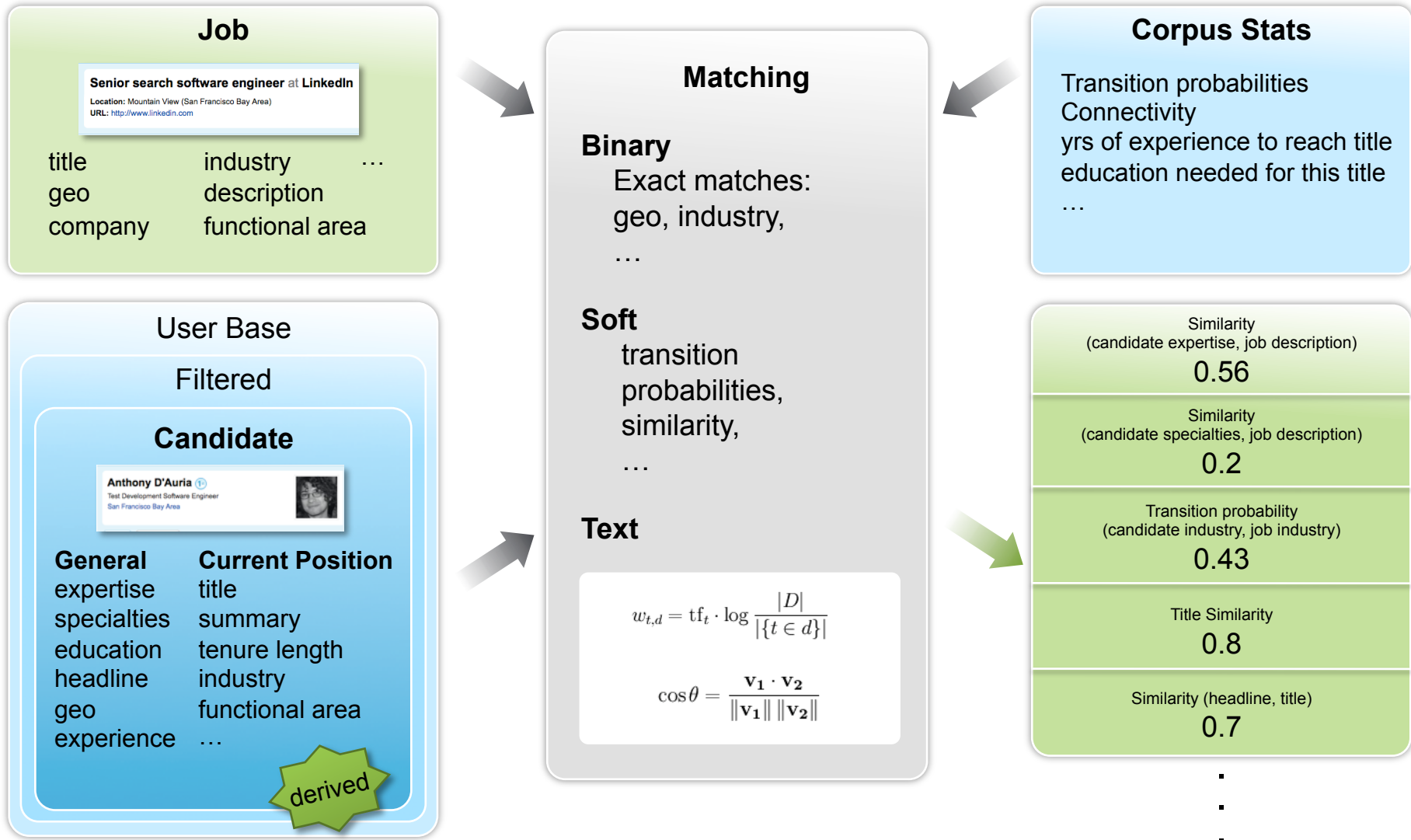
## JOB YOU MAY BE INTERESTED IN

-  **Business Intelligence Manager** Sponsored  
Eliason Corporation - ...
-  **Senior Data Scientist** ×  
SCEA - San Francisco, US...
-  **Senior Data Scientist (Product...)** ×  
Netflix - Los Gatos, CA
-  **Director of Data Science & Analytics** ×  
Glassdoor.com - San Fran...

[Feedback](#) | [See more](#) »



# Content Signals Dominate Social Signals



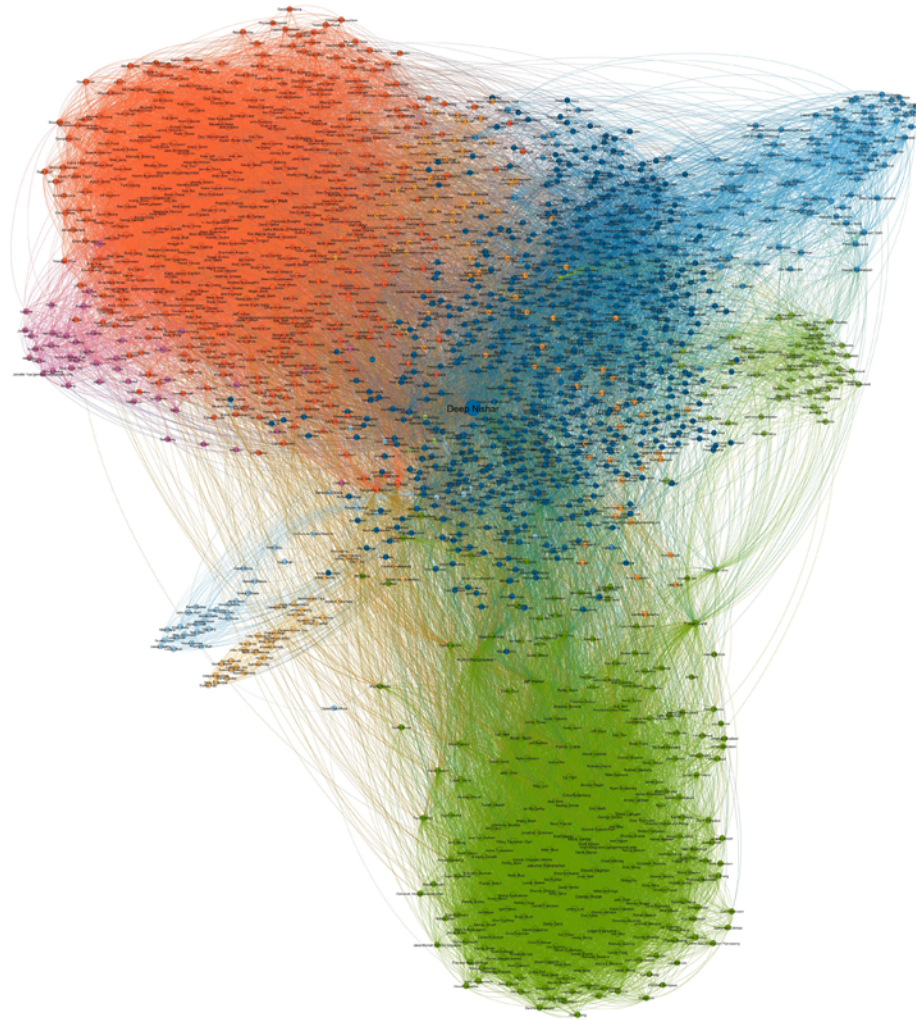
# Summary

**Provide relevant content**

**and establish social connections**









**in appropriate context.**

# Connections



# Connections are Social Dimensions

**A**











-  **Abrahams, Jessi** 130  
Senior Software Engineer - Endeca
-  **Ackley, Keith**  
Senior Director of Worldwide Customer Education - Endeca
-  **Adams, Richard**  
Director of Sales - Connotate, Inc.
-  **Adler, Joseph**  
Sr. Data Scientist - LinkedIn
-  **Adornato, Patrick**  
Business Analytics Account Executive - SAP
-  **Agarwal, Arun**  
Co-founder, VP Product & Marketing - Silver Lining Systems
-  **Agarwal, Mehul** 247  
Software Engineer - Google
-  **Aggarwal, Rahul** 227  
Principal Software Engineer

**Current** Principal Data Scientist at LinkedIn  
Member at New York CTO Club

**Past** Tech Lead at Google  
Chief Scientist at Endeca  
Consultant at Context Integration  
see all ▾

**Education** Carnegie Mellon University  
Massachusetts Institute of Technology  
St. Hilda's & St. Hugh's

Viewers of this profile also viewed...

-  **Monica Rogati**  
Senior Data Scientist @ LinkedIn
-  **Manu Sharma**  
Global Data Scientist at LinkedIn
-  **Skomoroch**  
Global Data Scientist at LinkedIn
-  **[unreadable]**  
Data Scientist in Residence at Greylock
-  **[unreadable]**  
Data Scientist at LinkedIn
-  **Wang**  
Data Scientist at LinkedIn
-  **[unreadable]**  
Data Scientist at LinkedIn
-  **Ian Posse**  
Global Data Scientist and Product...
-  **Scott Nicholson**  
I'm a data scientist & economist who...
-  **Heyning Cheng**  
Data Scientist and Product Innovator

# Example: People You May Know

**People You May Know** beta See people from different parts of your professional life

Google

Stanford University

LinkedIn

University of California, Berkeley


YAHOO!

Massachu... Institute of Technology


Carnegie Mellon University


Harvard University

+





**Cameran Evans** 2nd  
Manager, Business Intelligence at VMware  
San Francisco Bay Area

[+ Connect](#) 3 shared connections





**Miroslav Dudik** 2nd  
Researcher at Microsoft  
Greater New York City Area

[+ Connect](#) 13 shared connections




**Ledio Ago** 2nd  
Engineering Lead, Manager at Splunk  
San Francisco Bay Area

[+ Connect](#) 1 shared connection




**Rachael Carbone**  
Intern: Speech Language Pathologist Assistant at Spaulding Rehabilitation Hospital  
Greater Boston Area

[+ Connect](#) 2 shared connections




# Relationships are Social





**Werner Geyer** 1st  
Research Manager, Persuasive & Social Technologies  
Greater Boston Area | Information Technology and Services


### Werner's Connections


All (231) | **Shared Connections (11)**

- 

**Shlomo Berkovsky**  
Senior Researcher at NICTA, Australia  
Sydney Area, Australia  
10 connections in common
- 

**Peter Brusilovsky**  
Professor of Information Science and Intelligent Systems at University of Pittsburgh  
Greater Pittsburgh Area  
74 connections in common
- 

**Ed Chi**  
Research Scientist at Google  
San Francisco Bay Area  
82 connections in common
- 

**Ido Guy**  
Manager and Technical Lead at IBM  
Israel  
21 connections in common
- 

**Tessa Lau**  
Research Staff Member at IBM  
Argentina  
20 connections in common

## Mark Granovetter

Professor at Stanford University  
San Francisco Bay Area | Higher Education

Current	Professor of Sociology at Stanford University
Previous	Northwestern University
Education	Ph.D., Sociology at Harvard University

### How you're connected to Mark


You  
↓  
Mark Hull  
Duncan Watts  
↓  
2nd Mark Granovetter


# Shared Connections as a Signal

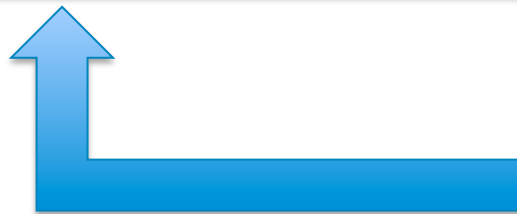
People You May Know

Stanford University    **LinkedIn**    University of California, Berkeley    Google    Mass Institute of Technology


[All Suggestions](#) / Google (40)

 **Linus Upson** 2nd  
VP, Engineering at Google  
San Francisco Bay Area

[+ Connect](#)     8 shared connections



**Linus Upson** 2nd

 [See all »](#)

# The Power of Social Proof

InMail™ What is Inmail?

**Compose your message**

To: Marissa Mayer  
From: Daniel Tunkelang  
 Include my contact information

---

**Enter the contact information you would like to share**

Email:

Phone:

---

Category:

Subject:

---

Your message to Marissa:

Marissa is interested in:

**Tips** to increase your likelihood of a response by 50%.

- 1 Mention people you know in common
  - Adam Nash
  - Craig Nevill-Manning
  - Maria Stone
- 2 Limit your message to 100 words - make it direct and to the point

# Beyond Triadic Closure

- Triads suggest and affect relationships.  
[Simmel, 1908], [Granovetter, 1973]
- Triangle closing is a Big Data problem.  
[Shah, 2011]
- Use machine learning to rank candidates.



# Summary

Provide relevant content

**and establish social connections**

in appropriate context.



# Context

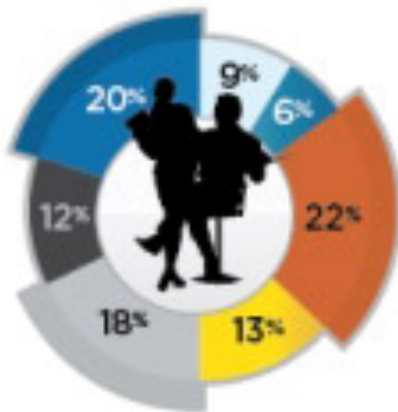


"to every thing there is a season"  
[Ecclesiastes 3:1]

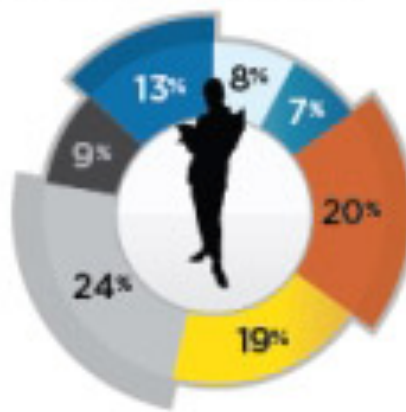
# One Platform, Many Users, Many Needs

What do you use LinkedIn for the most?

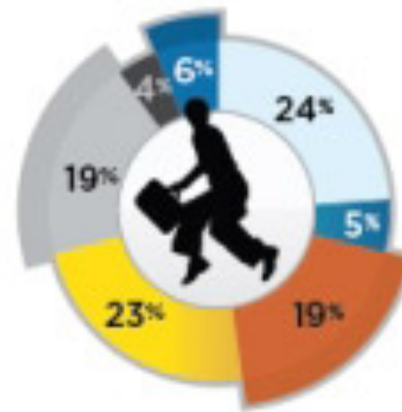
Top-level Executives



Middle Management



Entry-level



Job search

Networking w/ former employers

Industry networking

Co-worker networking

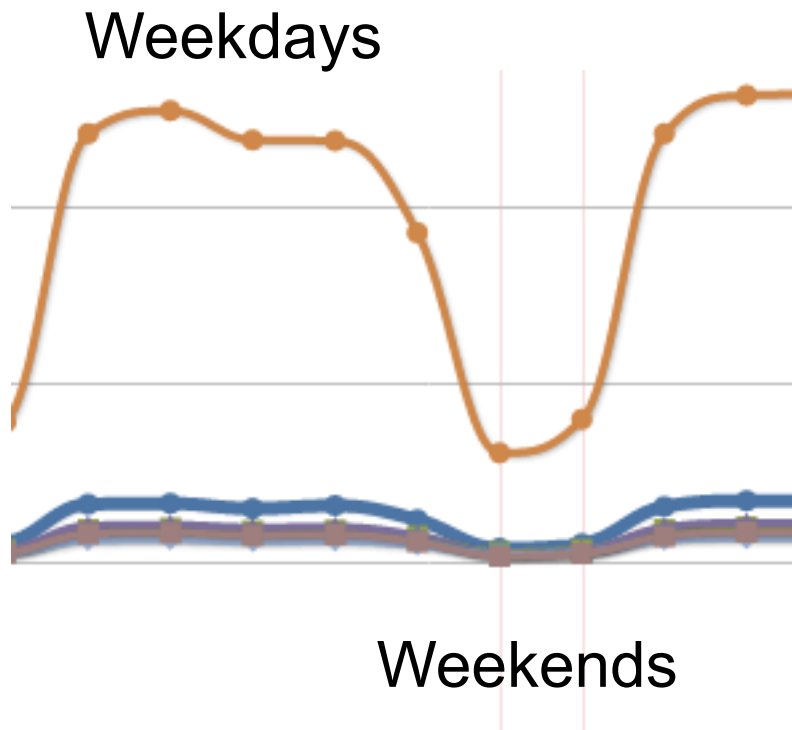
Keeping in touch

Hiring

Promoting my business

<http://blog.lab42.com/the-linkedin-profile>

# No Weekend Meetings!



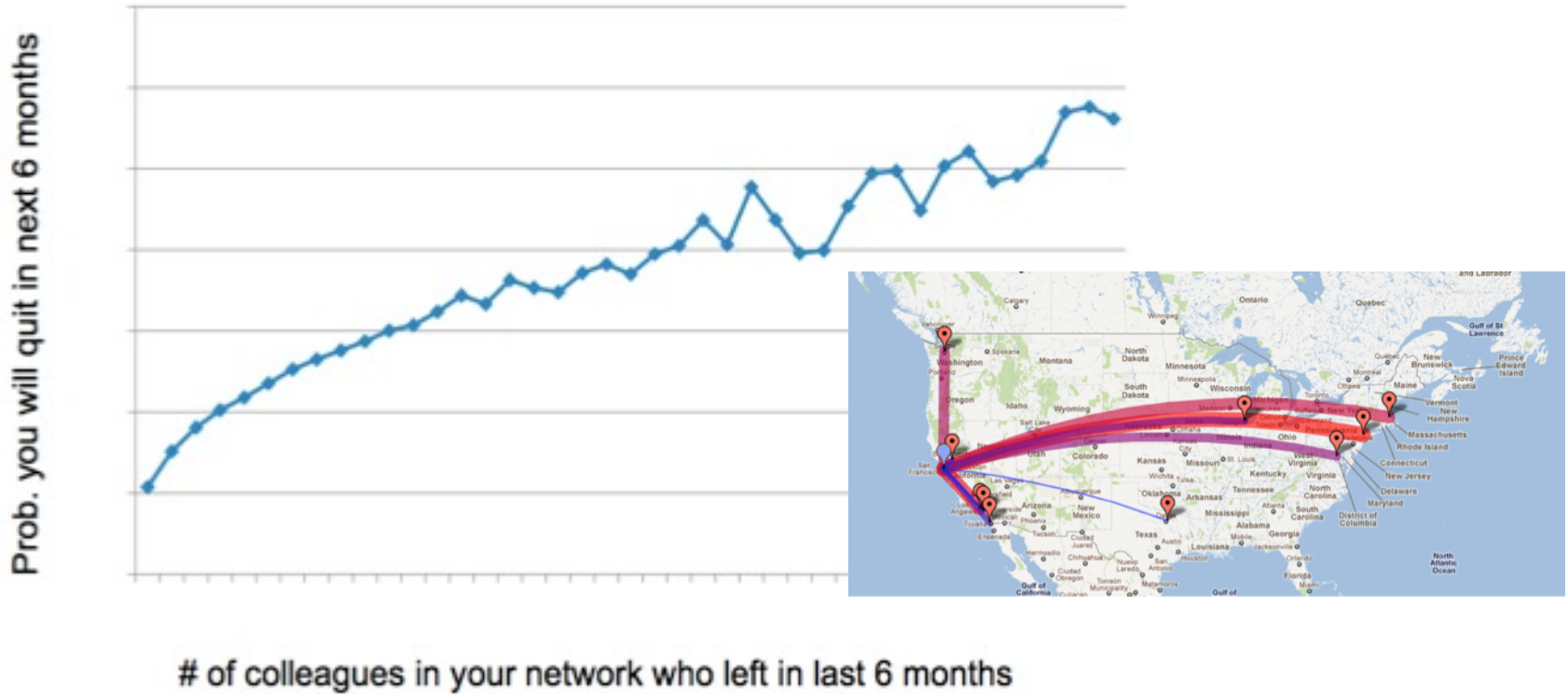
navigational queries

exploratory queries

# Different Devices for Different Needs



# Time and Place for a Career Change



# Active vs. Passive Job Seekers



**>10x** more likely  
to reply if active

**Mario Rodriguez**  
Analytics Engineer at LinkedIn  
San Francisco Bay Area | Internet

Previous UC Santa Cruz, Suncoast Schools FCU, Magenic Technologies  
Education MS, Computer Science at University of California, Santa Cruz

Improve your profile Edit 304 connections

<http://www.linkedin.com/in/mechanistician> Contact Info

## Summary

- Interests include analytics, data mining, visualization, data management systems, web technologies, software architecture.
- Broad development life-cycle experience (domain analysis, applied research, system design & development, product & project management).
- Proficient with Java, C, C#, Python, R, MATLAB, SQL, Pig, Hadoop, Mahout, Weka, Lucene, Spring, .NET, MySQL, Voldemort, SQL Server Analysis Services, Pentaho.

## Experience

### Analytics Engineer

LinkedIn

Public Company; 1001-5000 employees; LNKD; Internet Industry  
2011 – Present (1 year)

Working on enhancing LinkedIn's state-of-the-art recommendation engine and models.

[Rodriguez et al, 2012]



# Summary

**Provide relevant content**

**and establish social connections**

**in appropriate context.**



# Recap

Content is king.

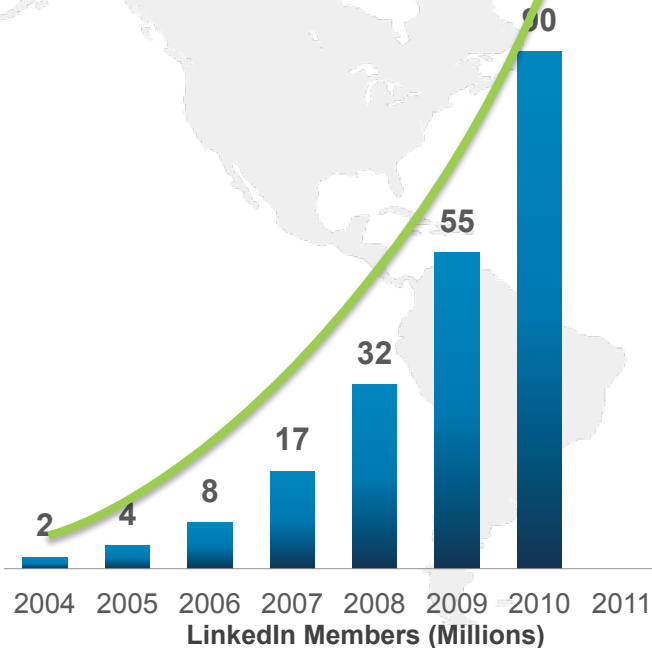
Connections provide social dimension.

Context determines where and when a recommendation is appropriate.

# Thank You!

**175M+**

62% non U.S.



**We're Hiring!**



**2/sec**



**25th**

Most visit website worldwide  
(Comscore 6-12)



**>2M**

Company pages



**85%**

Fortune 500 Companies use  
LinkedIn to hire

Learn more at <http://data.linkedin.com/>