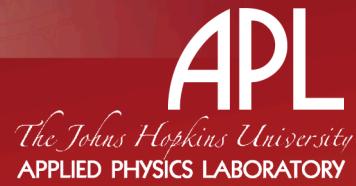


# **Social Relationship Identification: An Example of Social Query**

***Chris Diehl, Jaime Montemayor,  
Mike Pekala***

***August 31, 2009***

***Symposium on Social Intelligence and  
Networking 2009***



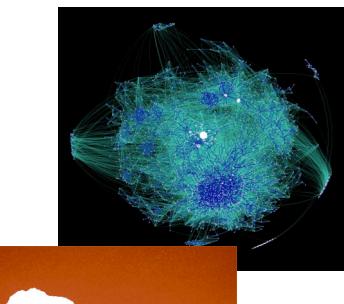
## Satisfying Social Needs, Expanding Possibilities



APL

In recent years, we have all witnessed the explosion of social media online. As the trends make clear, this is not a fad restricted to certain demographics. Instead it is a reflection of our basic human desire to be social. Social media has opened up the range of options we have to connect with others, both near and far, in some fundamental ways. It allows us to enhance our current relationships and form substantive new relationships with others online. In this talk and in our paper, we attempt to address the following question: do our currently available search tools support our desire to connect with others like ourselves? We believe currently available tools are inadequate. Furthermore, to address the way forward, we believe it is important to introduce the distinction between informational query and social query. Standard search tools available today address informational query. Social query is aimed at identifying individuals, relationships or groups that exhibit certain social attributes. Over the course of this talk, we will discuss the challenges and opportunities available to motivate the need to study and advance social query.

## The Rise of Participatory Media



APL

With the rise of the Internet, social media and mobile communication technologies, we've seen a significant transformation in how we communicate and connect with others. No longer is the ability to share content with large groups limited to the few. Anyone can share content with a global audience that is discoverable in near-real-time. With the availability of these new communication channels, the cost of communication among groups has collapsed. With the ability to share content asynchronously with anyone worldwide, the potential to form groups and sustain groups has exploded. No longer are the barriers of time and distance nearly as constraining.

## Digital Personas: Permanent, Public, Searchable



*"What's a nice guy like you  
doing in a persona like that?"*

APL

With content production no longer restricted to the few, we are increasingly embracing the concept of our digital persona. The digital social artifacts that we create, which reflect aspects of our relations with others, are permanent, public, searchable reflections of ourselves. The permanent and public aspects of our digital persona are characteristics we are increasingly aware of. Boundaries are explored, and crossed, in our online presentation. Meanwhile feelings evolve about where the public-private divide lies and how strict that divide should be.

But what about search? In this global, online community, how do I find the tribes in which I belong? How do I find other people around the world that share my passions, my curiosities, my feelings? Once a group of people is aware of one another, the tools are available to support group communication. But how do we find one another to start?

## Where's My Tribe Online?



APL

To make this more concrete, let me provide a personal example. I enjoy rock climbing and mountaineering. I love exploring the outdoors. I'm currently in the process of relocating to California and am eager to meet others that share my passions. As a blogger, I have had the experience of forming close relationships with others online. I'd like to discover other bloggers that I find compelling and engaging. How can I find them?

The screenshot shows the Google blog search interface. At the top, there are links for Web, Images, Videos, Maps, News, Shopping, Gmail, and more. The search bar contains the query "rock climber san francisco". Below the search bar, there are buttons for "Search Blogs" and "Search the Web", along with "Advanced Blog Search Preferences".

**Blog results**

**Browse Top Stories** New!

**Published**

- Last hour
- Last 12 hours
- Last day
- Past week
- Past month
- Anytime
- Choose Dates

**Subscribe:**

- Blogs Alerts
- [Atom | RSS](#)

**Related Blogs:** climber.me - rock climbing blog - <http://www.climber.me/>  
 UrbanBaby - Activities, Childcare, Preschools - San Francisco - <http://blogs.urbannbaby.com/sanfrancisco/>

**The Honor System: Mystery Product Revealed | BikeHacks**  
 1 hour ago by Matt  
 For the next few months I used a clunky \$60 rock climbing belay device to loop my helmet to my U-Lock. During those months, I started noticing how many people were walking the streets of San Francisco with a bike helmet under their arm. ...  
[BikeHacks - http://bikehacks.com/](http://BikeHacks - http://bikehacks.com/)

**UrbanBaby San Francisco - Tame Tantrums At the Climbing Gym**  
 21 Jul 2009 by Joyce Slaton  
 Scaling the heights at an indoor rock-climbing gym is an excellent way to exhaust those little arms. Every Bay Area gym owned by Touchstone Climbing (and there are Touchstone gyms in San Francisco, Berkeley, and Oakland, amongst other ...  
[UrbanBaby San Francisco - http://blogs.urbannbaby.com/sanfrancisco/](http://UrbanBaby San Francisco - http://blogs.urbannbaby.com/sanfrancisco/)  
 [ More results from UrbanBaby San Francisco ]

**Looking for a Rock Climbing partner in Brussels - Expat Forum ...**  
 15 Jul 2009  
 Hi, my name is Nicole and I just moved to Brussels from San Francisco. I used to rock climb several days a week but it has been ages! I would like to get back into it but need a climbing partner here. Level doesn't really matter but I ...  
[Easy Expat Forums - http://www.easyexpat.com/forums/](http://Easy Expat Forums - http://www.easyexpat.com/forums/)

**Michael Bialas: Folk U. (Day 3): 10 Best of the Fest Moments**  
 20 Aug 2009 by Michael Bialas  
 "San Francisco," "The One That Loves You The Most" and "She's Mine" were feel-good dance-in-the-aisle numbers that would work just as well in a Caribbean calypso bar. Then his encore presentation of "World Keeps Turning" was ... Much was made of Saturday's rock climbing accident involving a 28-year-old woman who fell 15 to 20 feet. Dennen, who happened to point out a few hikers on his way into the Planet Bluegrass venue Sunday, was unaware of the incident until he was ...  
[Entertainment on HuffingtonPost.com - http://www.huffingtonpost.com/feeds/verticals/entertainment\\_](http://Entertainment on HuffingtonPost.com - http://www.huffingtonpost.com/feeds/verticals/entertainment_)

APL

I started with Google blog search. These are the results returned for the query “rock climber san francisco.” I get a post that talks about a particular climbing product, a post about bringing kids to a climbing gym, a post from a woman who used to live in SF but is now in Brussels looking for a climbing partner, etc. What I want is a list of personal blogs written by climbers that I find engaging! The results are nowhere near what I’m after. Even if Google blog search returned a list of personal blogs, I would still have to read through them to find the ones that I find engaging. The tool is not helping me.

**TweepSearch**  
In ur Tweeps, Crawlin' their Bioz!

**TweepSearch** **BETA** rock climber san francisco

Pro tip: Search the location field by using: location:"Seattle, WA"

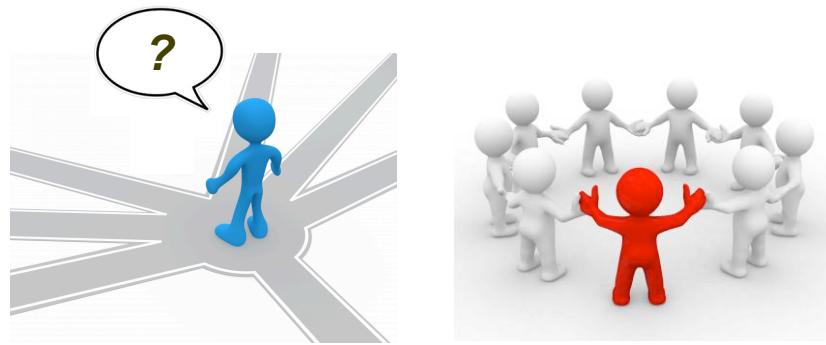
Results: 1 - 20 of 20 for rock climber san francisco.

|   |  | Sort by: last update ↑   screen name ↓   followers ↓   friends ↓  |
|---|--|---|
| This could be you! (@dacort on Twitter) - Seattle, WA<br>Interested in having your profile here? Contact dacort [at] tweepsearch.com for more details in order to help TweepSearch grow. Reserve a keyword for \$10 for 30 days, or share with five others for \$5. last recorded update about 1 awesome-second ago |  |   |
|    | @taylorclaire / Claire Taylor<br>San Francisco   | Brit living in SF who likes all things tech, <b>rock climbing</b> , photography and green tea. last recorded update 3 days ago  |
|    | @vanrenen / Gabriel Vanrenen<br>San Francisco<br><a href="http://www.skimble.com">www.skimble.com</a>      | rock climber, tech entrepreneur, dartmouth 03 last recorded update 3 days ago   |
|    | @ashley_martin / ashley_martin<br>San Francisco, CA  | Geek, vegetarian, runner, t-shirt deconstructionist, <b>rock climber</b> , live music lover, dog mommy, and hopeful romantic last recorded update 3 days ago  |
|    | @skimbl / skimble.com<br>San Francisco, California<br><a href="http://www.skimble.com">www.skimble.com</a> | Plan and track your active life! From <b>rock climbing</b> , hiking, running, cycling to yoga and gym fitness...see progress on skimble! co-creators @marlaly @vanrenen last recorded update 2 days ago |
|    | @jroualdes / jroualdes<br>San Francisco  | Waggener Edstrom communications consultant, scuba diver, photographer, <b>rock climber</b> and wanna be chef/mixologist last recorded update 14 days ago  |
|    | @kevinrose / Kevin Rose<br>San Francisco, CA<br><a href="http://kevinrose.com">kevinrose.com</a>           | Founder of social media site digg.com. Random ideas, entrepreneur, <b>climber of rocks</b> , video blogger, & tea drinker. last recorded update about 18 hours ago                                      |
|    | @wawonabrewchik / Kelly Hymes<br>San Francisco, CA   | Home Brewer, Bicycle Commuter, Music Lover, <b>Rock Climber</b> , Mushroom Hunter, Nature Lover last recorded update 26 days ago  |

APL

Next I tried TweepSearch for Twitter which searches Twitter user profiles. The results for the query “rock climber san francisco” are slightly better. Here I have explicit users listed on the left and profile information displayed. Yet once again I have a lot of work ahead of me to find the climbers I might connect with. I have to read through the tweets of each user and make decisions about who I find interesting. The burden of search is still on me.

## Shifting the Burden of Search



APL

So at this point, tools for informational query are not aiding me in my desire to connect. I'm left to search through retrieved digital social artifacts on my own to find bloggers that are compelling. How can we change this equation? How can I shift the burden of search to the machine?

## Social Signals

***How do you like the ice cream?***

Person A: I'd have to say that I like the ice cream.

Person B: The experience of eating the ice cream  
is quite satisfactory.

Person C: Yummy. Good stuff!

Person A: Could you check into this for me?

Person B: Get back to me on this ASAP.

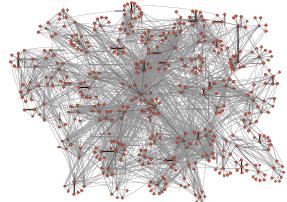
---

APL

Within the digital social artifacts that we all are creating online are social signals that reflect aspects of our personality and our relationships with others. Consider for example the three responses to the question "How do you like the ice cream?" All three individuals are expressing that they like the ice cream. Yet each person has a distinct way of expressing that feeling. In the second example, the two individuals have the same goal in mind when speaking with the other person. Yet their approaches are quite different, signaling distinct differences in the existing relationships.

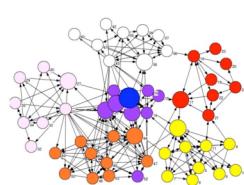
## Leveraging Social Signals: Social Relationship Identification

Communication Graph

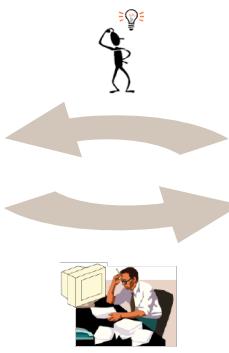


Nodes: Network References  
Edges: Communication Events

Social Network



Nodes: Entities  
Edges: Social Relationships

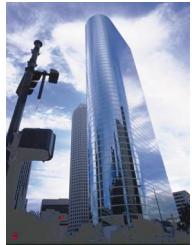


14

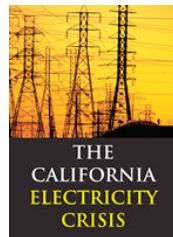
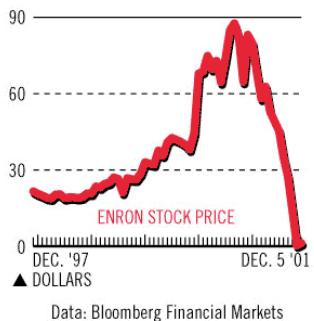
APL

Our initial insights into online social signals came from studying the problem of social relationship identification within the context of organizational email. The task of social relationship identification involves identifying communication relationships that exhibit social relations of interest. Within organizational email, we may be interested in friendship, trust and manager-subordinate relations to name a few. How can we identify these types of social relations from observations of the communications? What form does this type of social query take?

## The Enron Scandal: What Happened?



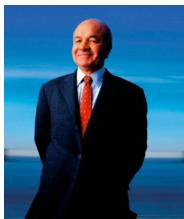
### STUNNING COLLAPSE



APL

In our investigation, we chose to focus on the Enron scandal. Enron was one of the world's leading energy companies with claimed revenues of \$111 billion in 2000. By late 2001, Enron was bankrupt. What happened? Clearly something highly suspect transpired.

## The Enron Scandal: Key Players

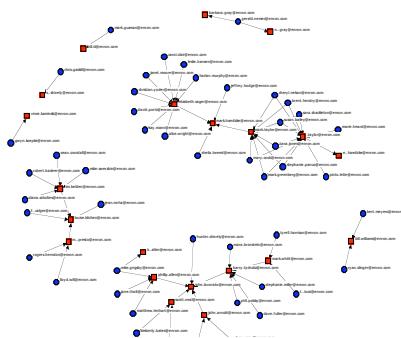


APL

What do we have to work with? In the case of Enron, we are certainly aware of some of the key actors at the top. In addition, there are volumes of documents and email that were obtained and made part of the public record. Our subject of interest was the gentleman in the upper right, Tim Belden, who was the mastermind behind Enron's many financial schemes. Some of the fundamental questions we wanted to address included (1) who did Tim Belden report to? (2) who reported to Belden? And (3) when were these relationships active?

## Specifying the Query

*Organizational Ground Truth  
January 2000 – November 2001*



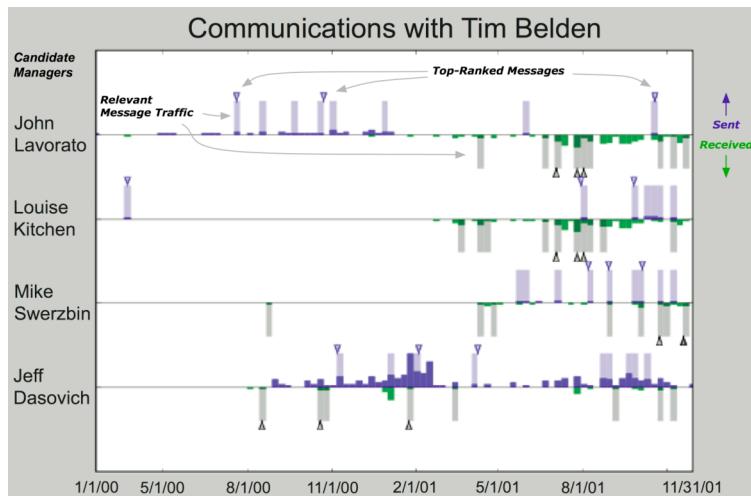
### ▪ Distinguishing Characteristics

- Implicit Specification of Social Relevance
- Explicit Specification of Relevant Time Periods

APL

So what is the nature of a query supporting social relationship identification? What form does it take? In the case of our example, we started with a series of labeled communication relationships that exhibited the manager-subordinate social relation over the specified time period. Each relevant communication relationship consists of the set of email messages exchanged between the two individuals in the relationship over the specified time period. By labeling the set of messages as relevant, we are not specifying specific examples of communication that are indicative of the social relation. Instead we are simply suggesting there is language embedded in this set of messages that signals the existence of the social relationship. Implicit specification is key for two reasons: (1) we don't expect that we necessarily are aware of all the available cues and (2) we don't want to assume the user will label any specific instances. Later on while exploring the data, the user may wish to make such specific assertions. For now, we examined the weakest form of relevance feedback. This type of implicit specification is in contrast to informational queries where we typically construct an explicit specification of relevance. The other distinguishing aspect of this social query from informational queries is the explicit specification of time periods. In the specification of the relevant relationships, we are indicating the time periods over which the social relations exist. We are also specifying the time periods over which we are interested in identifying more of these social relations.

## Query Results



18

APL

For social queries, we need to reconsider how the query results are presented as well. The process we have defined for social relationship identification takes the set of communication relationships for an individual of interest and rank orders them according to their relative likelihood of expressing the social relation of interest. The ranker also rank orders messages within a given communication relationship and highlights the most compelling message traffic that supports the overall rank score for the communication relationship. The relationship timelines show the exchange of message traffic along with the time periods in which the most compelling messages were sent.

## Query Results

From: [tim.belden@enron.com](mailto:tim.belden@enron.com)

To: [john.lavorato@enron.com](mailto:john.lavorato@enron.com),  
[louise.kitchen@enron.com](mailto:louise.kitchen@enron.com)

Subject: Off-Site Travel Question

The e-mail that was sent out many weeks ago about the off-site indicated that it would run from Wednesday night to Saturday AM. It is now running Thursday until Sunday... I had already made arrangements to attend a wedding in Oregon on Saturday night. It is a good friend of mine and my wife's... While I have authority to place millions of dollars of the company's money at risk, I don't feel comfortable signing up for a \$7,000 extra flight without talking to you guys... Any thoughts would be greatly appreciated.

APL

Here is a particular example of a message highlighted by the ranker. Tim Belden routinely reported to both John Lavorato and Louise Kitchen, which are the top two candidates highlighted by the ranker. In this message, Belden is explaining how he and his wife had existing plans to attend a wedding. The dates for the company offsite were changed, causing a conflict with the wedding. The highlighted statement makes it quite clear that Belden is seeking approval from them to make travel plans to handle the change.

We recognize what makes this message compelling to us, but what is the machine focusing on when highlighting this message?

## Query Results

From: [tim.belden@enron.com](mailto:tim.belden@enron.com)

To: [john.lavorato@enron.com](mailto:john.lavorato@enron.com),  
[louise.kitchen@enron.com](mailto:louise.kitchen@enron.com)

Subject: Off-Site Travel Question

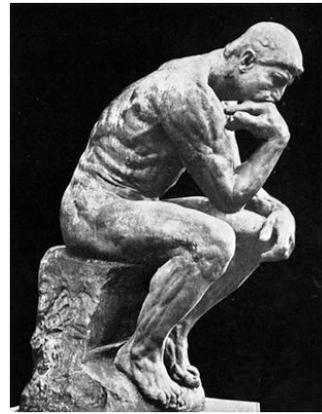
The e-mail that was sent out many weeks ago about the off-site indicated that it would run from Wednesday night **to** Saturday AM. It is now running Thursday until Sunday... I had already made arrangements **to** attend a wedding in Oregon **on** Saturday night. It is a good friend of mine and **my** wife's... While I have authority **to** place millions of dollars of the company's money at risk, I don't feel comfortable signing up for a \$7,000 extra flight without talking **to** you guys... Any thoughts would **be** greatly appreciated.

APL

Interestingly enough, the most compelling / discriminative terms are the highlighted function / stop words. In fact if you look at the list of the top 50 significant terms identified by the ranker, many of them are function words. This was particularly surprising to us. Yet to our colleagues in social psychology who study emotional, social and psychological indicators in text, this result is not surprising at all. Function word use determines the style of the communication and therefore plays a significant role in social signaling. The function words that we routinely remove in preprocessing text for information retrieval tasks in fact provide the signals we are looking for.

## Social Query: Rethinking the Fundamentals

- **Query Formulation**
- **Features**
- **Learning Tasks**
- **Visualizations**
- **Relevance Feedback**



APL

21

By examining the social relationship identification task, we hope to have made clear that for social query, we need to rethink all aspects of the query process. Social query presents new twists at every stage that require investigation. A rich landscape of research lies in wait.

## Take Aways

- **Digital Personas,  
Social Signals**
- **Inadequacy of  
Informational Query**
- **New Paradigm Needed  
for Social Query**



***Help me find my tribe!***

APL

22

In our paper and through this talk, we aimed to highlight the tremendous opportunity to leverage digital personas online to help people find others they truly resonate with. We believe digital social artifacts contain informative social signals that can be analyzed to aid in the discovery process. Current tools designed for informational query do not offer substantive support for social query. In fact, they leave the user with the burden of search. We presented the task of social relationship identification as an example of social query to clarify many of the distinguishing aspects. We believe that social query requires us to revisit the fundamentals of the query process in order to make it a reality. We look forward to working with others to define and realize that reality.

*Thank You!*  
*Questions?*



*APL*