Social Sites: Opportunities and Challenges
(from a researcher's, a developer's, and a user's perspective)

Participants: **Amr El Abbadi** (University of California Santa Barbara, USA)

Sihem Amer-Yahia (Qatar Computing Research Institute – QCRI, Qatar)

Laks Lakshmanan (University of British Columbia, Canada)

Raghu Ramakrishnan (Yahoo! Research, USA)

Moderator: **Georgia Koutrika** (IBM Almaden Research Center, USA)

● **What is the most promising/useful/exciting use of online user and social info?**

- Information discovery/serving:
    - Improving quality of search;   ☺
    - Improving quality of recommendations as well as of what comes under the purview of RecSys.
    - Viral Marketing.
    - Targeted advertizing.

- Risks of abuse?
    - Stalking -- no privacy!   ☹
    - Witchhunting -- As an employer, I get to profile your tastes, the kinds of issues you comment on, what you said about them.

● **What can be done with all (past and current) social knowledge accumulated?**

- In a sense, social content will be "archived" for ever! (Try preaching Hippocratic storage doctrine to the masses or to SN owners!)
- User generated content (UGC) is time sensitive
  -- often comes with a timestamp.
  Adds additional dimensionality to RecSys.
  In Foursquare, UGC comes enriched with time and location dimensions. This can only help improve the quality of search and recs.

● **What would be the most promising/exciting thing to have people contribute to and how to leverage that?**

- Wrong question! :-)
  Users will say/do (aka "contribute") whatever they want!
- We can ask are there good *wikipedia* grade ideas we
  could encourage/incentivize users to make contributions to?
  E.g., the uptake rate on *wikimapia* doesn't seem to be as great.
- "Social breadcrumbs" left behind in foreign lands can be valuable
  for visitors who don't speak the language (well enough), for
  purposes of getting around and getting things done.
- *Groupon* and *TripAdvisor* -- great examples of phenomena
  w/ tremendous user traction. This question really amounts to
  asking what are some cool startup ideas that could benefit from UGC.

● **What would be the next exciting startup?**

- Short and honest answer: I don't know!
- Something in the viral marketing space.
- Something in the space that combines search and recs, going beyond simply recommending music and movies, which is boring. :)
  E.g., with foursquare kind of data, you can do really interesting things in the way recs are generated, leveraging multiple dimensions and user preferences.

● **What is the biggest challenge/problem that researchers (will soon) face?**

- availability of data.
    - Much better than a decade ago.
    - There are still major limitations.
- Dealing with scale and dynamics:
    - event detection & tracking.

● **Are all/any challenges data-management related?**

- CS (leave alone DMM) cannot solve all problems here.
- need to collaborate with psychologists, social scientists and economists.

● **How do we balance privacy and openness after all?**

- Nobody knows. There are obvious extremes.
- The exact balance may not be known unless and until some
  solutions are tried out.
- Privacy on the internet is almost a myth?
- A better question to ask is what infrastructure should be put in
  place in order to make it "sufficiently difficult" for attackers/
  hackers to get at private data. This is an ongoing dynamic balance.
- We are talking application security as much as we are data
  security here.
- Studies show most users don't even follow simple secure
  practices like checking/updating browser version etc.

● **Who owns Social Data (and what can we do about it)?**

- blunt reality: nobody and everybody.
- Who owns/owned the Vancouver Stanley Cup riots videos and London riots videos data? Who should have owned it?
- What about uprisings in the middle east?

● **What would the (social) future look in X years from now?**

- Most events organized organically using social media.
- Mostly e-books or online docs.
- Matchmaking, finding study buddies, finding project partners for a course being taught online across a widespread area.
- getting solutions for local engineering problems from afar.
- high tech crime enabled by technology.
- high tech crimefighting, … .