

The Search and Social Media Workshop at SIGIR 2009

Eugene Agichtein
Emory University
eugene@mathcs.emory.edu

Marti A. Hearst
UC Berkeley
hearst@ischool.berkeley.edu

Ian Soboroff
NIST
ian.soboroff@nist.gov

1 Introduction

Social applications are the fastest growing segment of the web. They establish new forums for content creation, allow people to connect to each other and share information, and permit novel applications at the intersection of people and information. However, to date, social media has been primarily popular for connecting people, not for finding information. While there has been progress on searching particular kinds of social media, such as blogs, search in others (e.g., Facebook, Myspace, or flickr) are not as well understood.

To address these questions, the second workshop on Search and Social Media (SSM 2009) was held at SIGIR 2009 in Boston, MA in July 2009. The main workshop website is available at <http://ir.mathcs.emory.edu/SSM2009/>. SSM 2009 followed on the highly successful SSM 2008 workshop held at CIKM 2008 in Napa, CA. As in the previous year, the workshop had nearly 50 attendees from academia and industry.

The purpose of this workshop was to bring together information retrieval and social media researchers to consider the following questions: How should we search in social media? What are the needs of users, and models of those needs, specific to social media search? What models make the most sense? How does search interact with existing uses of social media? How can social media search complement traditional web search? What new search paradigms for information finding can be facilitated by social media?

2 Workshop Format

Search in social media is still a new field, so submissions were solicited in the form of late-breaking and novel research results, and position and vision papers discussing the role of search in social media. Eight contributed papers were accepted for presentation and a number of additional presenters were invited to participate in panel discussions.

For two blocks of time, a focus topic was chosen and led off by a keynote speaker. Then other speakers made brief presentations, and then they all joined a discussion panel including participation from the audience. The other two sessions were panels consisting of a combination of submitted presentations and invited speakers. The workshop organizers attempted to bring the feel of social media to the workshop by projecting the results of a live Twitter feed (using the hash key #SSM09) next to the main projector screen. We report selected tweets in the corresponding sections, with the full twitter stream also available¹. The workshop program, listing the contributed and invited presentations, is available on the workshop website².

3 First Session: Online Communities and Recommender Systems

Prof. Joseph Konstan of University of Minnesota led off the workshop with a talk entitled “Help Is Out There: Online Community, Community Artifacts, And A New Way Of Harnessing Knowledge.”

The talk discussed studies that explore how online communities create and organize information and how we can draw lessons on how social psychology, economics, and other sciences of human behavior can be harnessed to understand and then design effective online communities. In particular, Konstan looked at cases where computation and machine learning have the potential to improve the functioning of such communities and how to derive insights into the future of searching in social media. Some tweets made about this talk are reported in Figure 1.

After this talk, Konstan joined a panel discussion on the topic of recommender systems with Manish Agrawal, Scott Golder, and David Carmel. Agrawal from UIUC spoke about crowdsourcing local newsletters with a Facebook application, Golder of Cornell spoke about preliminary work in analyzing social networks for finding people, and David Carmel from IBM explained the SAND project for content/people recommendation in large companies.

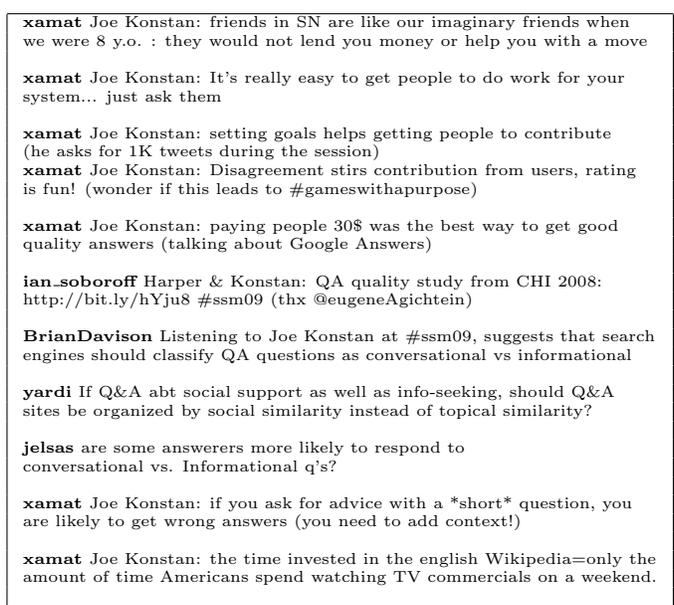


Figure 1: Selected tweets during Joe Konstan’s keynote presentation

4 Second Session: Collaborative Search

¹SSM 2009 Twitter archive available at <http://twapperkeeper.com/ssm09/>

²Available at <http://ir.mathcs.emory.edu/SSM2009/program.html>

The second session was on the relatively new topic of collaborative search. Collaborative search is a different angle on social media; it is social in the sense that multiple people are searching together aided by collaborative software tools. The panelists were Jaime Teevan of Microsoft Research, Ivana Marenzi of Hannover University, and Jeremy Pickens of FXPAL. The speakers then answered questions from the audience as well as those submitted via Twitter. The consensus was that some search tasks are inherently collaborative, and new search interfaces and modalities are just starting to emerge to support such tasks. Selected tweets posted during this panel are reported in Figure 2.

BrianDavison Jaime Teevan giving tutorial of who/what/when/where/why collaborative search

ian_soboroff popular collab search tasks: travel planning, shopping, technical information/literature

jerepick @ApolloGeek I agree; domains for collaborative search definitely have more in common with non-web srch than with web-ish known-item srch

ApolloGeek We have seen some benefits from collaboration for recall-oriented search.

ApolloGeek Search jamming is a wonderful metaphor.

ian_soboroff @xamat my view is indeed, search is secondary in SM, but I wonder how much of that is we don't have tools or know the tasks yet

ApolloGeek Search using social media to support collaboration seems very interesting, but is it what we mean by "searching social media"

Figure 2: Selected tweets during the session on collaborative search

5 Second Keynote: Twitter Search

Abdur Chowdury, head of search for Twitter, gave the second keynote, dropping provocative questions for the audience to consider.³ He focused the discussion in Twitter Trends, which try to capture important topics that people around the worlds are discussing. He noted that hackers band together and try and game the system to influence the trends. He also discussed efforts to group trends by geolocation and by category. He also noted that one of the big uses of Twitter is shared events, as evidenced by its use for the search and social media workshop itself. Selected tweets are summarized in Figure 3.

jeffd is watching the twitter address at the SIGIR #SIGIR09 SSM workshop. I love the audience participation!

redlog Blown away by @abdur's animated visualization of Super Bowl. Fine-grained temporal data FTW!

redlog Q: when trends cover the world, every subculture/community gets washed out, right? How to prevent building hegemony into the system?

BrianDavison @abdur: "algorithms have very few ethics"

jteevan Just to repeat myself, I think searching social media is particularly cool bc yr SN can instantiate what you're seeking if not there.

xamat @abdur giving all sorts of examples of location based twitter search to identify trends

craig_macdonald research on twitter is of interest, but will not be sufficiently attractive without available static data instead of the feed?

Figure 3: Selected tweets during Abdur Chowdury's keynote presentation

6 Concluding Sessions: Tags and Search in Social Media

The last two sessions of the workshop considered user-generated *tags*, a key characteristic of social media, and how tags and other features of social media can be exploited for search. The first presentation of paper by Harvey et al. described using tag clouds for content indexing. Joshi described

³The summary of this talk was aided by a blog post at Jeff Dalton's Search Engine Caffe blog, <http://www.searchenginecaffe.com/2009/07/sigir-social-media-workshop-abdur.html>

using community information to improve image search. Heymann analyzed tags to determine potential usefulness for search, and Yeung discussed identifying expertise in collaborative tagging systems.

After a short break, Seo continued the topic of expert finding as applied to online forums. The next paper, by Ganjisaffar et al., examined exploiting user reviews to improve search in wikipedia, followed by Seki who described automatically identifying spam blogs (also known as “splogs”). The last session was concluded by Doug Oard, who asked which aspects of research on search in social media connect to other areas of information retrieval and natural language processing, and challenged the workshop attendees to consider the definitions and boundaries of social media. This panel naturally transformed into an open discussion about research in social media and key research directions to pursue in the future. Summary of the tweets during these sessions is in Figure 4.

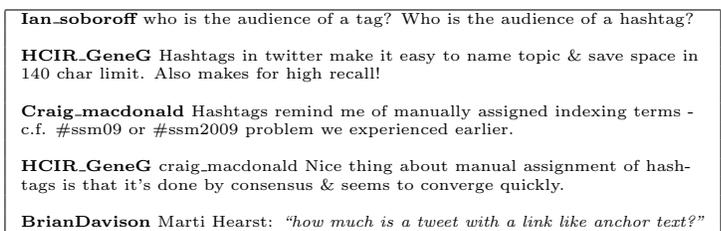


Figure 4: Selected tweets during the concluding workshop sessions

7 The Future

Although an important field, research is still sparse in this area. The third workshop on Search and Social Media will be held in conjunction with Third ACM International Conference on Web Search and Data Mining (WSDM 2010) in New York City in February 2010⁴. The SIGIR membership is urged to submit demos, posters, and position papers to this workshop.

Acknowledgements

The SSM workshop organizers thank the SIGIR 2009 organizers for hosting this workshop, and the members of the SSM'09 program committee for their work helping to make this workshop a success: Eytan Adar Ed Chi, Abdur Chowdhury, Natalie Glance, Bernardo Huberman, Matthew Hurst, Pranam Kolari, Craig Macdonald, Gilad Mishne, Nitya Narasimhan, Nicolas Nicolov, Doug Oard, Iadh Ounis, Maarten de Rijke, Markus Strohmaier, and Andrew Tomkins,

⁴<http://ir.mathcs.emory.edu/SSM2010/>