Information Search and Retrieval in Microblogs

Miles Efron
Graduate School of Library and Information Science
University of Illinois, Urbana-Champaign
501 E. Daniel St.
Champaign, IL 61820
mefron@illinois.edu

Abstract

Modern information retrieval (IR) has come to terms with numerous new media in efforts to help people find information in increasingly diverse settings. Among these new media are so-called microblogs. A microblog is a stream of text that is written by an author over time. It consists of many very brief updates that are presented to the microblog's readers in reverse-chronological order. Today the service called Twitter is the most popular microblogging platform. While microblogging is increasingly popular, methods for organizing and providing access to microblog data are still new. This article offers an introduction to the problems that face researchers and developers of IR systems in microblog settings. After an overview of microblogs and the behavior surrounding them, the article describes established problems in microblog retrieval such as entity search, sentiment analysis, and modeling abstractions such as authority and quality. The article also treats user-created metadata that often appear in microblogs. Because the problem of microblog search is so new, the article concludes with a discussion of particularly pressing research issues yet to be studied in the field.

Introduction

The term *microblog* describes an increasingly common information medium. Usually consisting of brief textual entries written on an ongoing basis, a typical microblog is written by a single

person or entity and is read by anywhere from zero to hundreds of thousands of "followers." Users of microblog services such as Twitter[1] are likely to maintain a microblog of their own. Likewise, each user typically follows some number of other peoples' microblogs. As millions of users broadcast updates to their followers, microblogs appear attractive insofar as they promise access to timely information written by people we have chosen to pay attention to.

However, finding and managing information in growing masses of microblog data is non-trivial. This article discusses the most poignant problems faced by those of us who would like to help people use microblog data as they pursue their information needs. To contextualize this article's motivation, consider four questions that fall under the broad rubric of microblog information retrieval (IR).

1. There are many millions of microbloggers. Which of those millions of people write often enough on topics that interest me and with sufficient authority that I should follow them?

2. I follow over 1,000 peoples' feeds because they all post interesting material at least occasionally. But following 1,000 feeds, each of which is updated several times daily leads to information overload. How can I winnow my incoming microblog updates so that I see only the most interesting entries?

3. I plan to buy a new computer and am deciding between two models. I know that people voice opinions on Twitter about the merits of each model. How can I find trustworthy opinions on the matter, and how can I synthesize these opinions to help me choose a computer to buy?

[1] http://twitter.com

4. A wildfire is burning in my neighborhood. Many of my neighbors are posting information about the situation such as evacuation zones, aid station statuses, weather conditions in specific locations to Twitter. What information relevant to my situation is available in this setting? Also, what information could I contribute that would be of use to others?

Several points are worth noting with respect to these questions. In question 1 the unit of retrieval is a *person.* Instead of textual documents, the person asking question 1 expects an IR system to suggest people worth following. Question 2 imagines a scenario where the unit of filtering is the individual *post.* Thus in microblog IR, the unit of retrieval is often either non-obvious or in need of formalizing to permit tractable retrieval. Question 3 points to the role that subjectivity and opinion play in microblog IR. Because a great deal of microblog data does express opinions, IR systems built on these data invite hopes of answering questions related to abstract matters such as consensus or dispute. Finally, in Question 4 we see the role that *time* and *place* play in microblog IR. People post to microblogs at particular times in particular places. Rapidly updated data and knowledge of authors' locations suggest that microblog IR operates in an arena where relevance is often bounded by time and place.

The sections that follow elaborate on these ideas. But while this article focuses on information retrieval, I must stress that microblog IR is in its infancy and thus my discussion ranges to cognate problems, as well. Many of the problems that we face in microblog IR are part of established research areas. For instance, Question 1 above poses problems very similar to expert finding, a well-studied IR problem. Question 3 involves modeling opinion, a problem that is widely treated in the area of sentiment analysis. More broadly, the popularity of microblogs must be understood in the larger context of increasingly pervasive social computing in general. In the remainder of this article, then, I will articulate the main challenges and opportunities facing those

who study microblog IR. But doing so will require us to visit a broad palette of related literatures.

Overview of Microblogs

A microblog is usually maintained by an individual person or entity. It consists of periodic and brief posts (often called *status updates*) that collectively offer readers (usually called *followers*) timely information deemed interesting by the microblog's author. What constitutes interesting information is of course subjective. Many microblogs are highly personal and of interest only to their authors' close acquaintances[2]. Other microblogs have a wider audience. For instance thousands of people follow updates to celebrities' microblogs, affording these celebrities new venues for cultivating their fan bases. Other microblogs have wide audiences because their content has broad appeal. Many people follow a small number of elite Twitter users because these people post information that is consistently interesting or amusing.

At the time of this writing, several microblog services exist. For example, Facebook[3], Tumblr[4], Orkut[5], and FourSquare[6] all offer microblog-type services. However, at this time the most visible microblogging platform is Twitter. Because of its widespread use among microbloggers, in the remainder of this article I will use examples, data, and vocabulary taken from the context of Twitter interactions.

---

[2] The "value" of Twitter updates is of course subjective. A 2009 study (Kelly, 2009) classified 41% of tweets as "pointless babble," an opinion rejected by danah boyd in (2009).
[3] `http://facebook.com`
[4] `http://tumblr.com`
[5] `http://orkut.com`
[6] `http://foursquare.com`

Each user of Twitter maintains his or her own microblog. A user's microblog consists of individual posts, called status updates, each consisting of a text string of no more than 140 characters. Twitter posts are also known as *tweets.* While statuses are brief, they often point to external material by including a hyperlink to a document on the Web, or to an image or other uploaded content. In Twitter jargon, a given user's statuses are called his or her *timeline.* For our purposes, this is identical to his or her microblog.

Twitter was launched in 2006 and saw rapid growth in membership when a critical mass of participants in that year's South by Southwest conference in Austin, Texas adopted the service (Douglas, 2007; Terdiman, 2007). Since that time, Twitter's popularity has grown at an increasing rate (Beaumont, 2010; Twitter, 2010). Writing in *The Daily Telegraph,* Beaumont reports that as of February 2010, approximately users post fifty million tweets per day.

Peoples' motivations for microblogging are diverse (Jansen, M. Zhang, Sobel, & Chowdury, 2009; Java, Song, Finin, & Tseng, 2007; Krishnamurthy, Gill, & Arlitt, 2008; Zhao & Rosson, 2009), ranging from offering autobiographical updates to reporting on current news events (Diakopoulos & Shamma, 2010; Shamma, Kennedy, & Churchill, 2009) or crisis situations (Longueville, R. S. Smith, & Luraschi, 2009; Sakaki, Okazaki, & Matsuo, 2010; Vieweg, Hughes, Starbird, & Palen, 2010). Peoples' use of Twitter to communicate during political unrest (e.g. the disputed 2009 election in Iran) brought increased attention to microblogs as a communication medium (Evgeny, 2009; Grossman, 2009). Twitter has played a role in important events, but the service also allows people to communicate among a relatively small social circle, and a sizeable part of Twitter's success owes a debt to this function.

While the reasons for Twitter's popularity are complex, artifacts of peoples' interactions with Twitter are readily observable, thanks to the application programming interface (API) that allows

developers to download data from Twitter programmatically. The availability of Twitter data has led a host of researchers to pursue high-level questions regarding people's use of microblogs. At the broadest level, initial research addresses these questions:

- Who are the users of microblog services?
- What qualities characterize these users' social networks?
- How does information flow through microblog communities?

Two of the earliest studies of the demographics and network statistics of Twitter are (Java et al., 2007) and (Krishnamurthy et al., 2008). Both of these studies analyze samples (gathered in different ways) of tweets in efforts to characterize peoples' interactions with Twitter. Some findings here are not surprising. For example, Twitter data such as peoples' number of status updates and their number of friends and followers follow power law distributions. As in many informetric settings, a large proportion of data is due to a small number of highly engaged people. Java et al. note that the power law exponent for Twitter social graphs is approximately -2.4, a number that is very similar to the parameter for graphs on the Web (pp. 58-59).

However, less predictable results also emerged from these studies. As of 2008, Twitter use was far-reaching geographically speaking. While the US contained the lion's share of Twitter activity, these studies found heavy use internationally, as well. Of particular interest is the finding that many peoples' social networks contain geographically diverse entries, with connections often spanning several continents (Java et al., 2007). Additionally, Twitter networks tend to show strong reciprocity. According to Huberman, Robero, and Wu (2009), "90 perecent of a user's friends reciprocate attention by being friends of the user as well." That is, people who person *x* follows are likely also to be followers of person *x*.

Digression: A Microblog Corpus and a Vocabulary for Discussing IR

During this essay, I will make reference to statistics calculated from a corpus of microblog data that I have collected and used in my own research. Twitter has made acquiring microblog data trivial by exposing their content via a streaming API[7]. The streaming API allows a developer to harvest a putatively random sample of tweets very quickly. However, this stream ("the garden hose" in Twitter jargon) delivers tweets out of context. If tweet *a* by author *A* is sampled, it is unlikely that other tweets by *A* will find their way into the sample. Nor is it likely that tweets by other people *B, C, D,* etc. who are members of author *A*'s social network will be represented in the sample. It is true that this context could be recovered by using other methods in the Twitter API, but on a large scale, such contextualization becomes an engineering burden, exacerbated by API rate limits and latency in API responses.

To build a more naturalistic test bed of microblog data, I decided to focus on a single (but broadly construed) community. The core of this collection is Twitter users who share interests in matters relating to information retrieval, human-computer interaction and information science.

My goal in collecting these data was twofold:

1. I wanted to capture data that allows analysis of the social aspects of microblogging.
2. I wanted the corpus to contain a large, heterogeneous topic base.

I aimed to build a collection where interactions between people and the diffusion of ideas within social circles could be observed. However, to allow for realistic analysis, the data set should not hew too narrow a line with respect to topical coverage.

---

[7] `http://dev.twitter.com/pages/streaming_api`

I identified the core set of authors in this corpus by tracking tweets that were written in a two-week period; one week before and after the date of the announcement of paper acceptances and rejections for the 2010 ACM SIGIR conference (March 24, 2010). My hope was to use this event as an opportunity to identify people in the IR community who use Twitter. To accomplish this, I used the Twitter streaming API, tracking all tweets using any of the words in Table 1:

sigir	sigir2010	geneva	workshop
retrieval	iiix	cfp	hcir

Table 1. Words used to track tweets using Twitter's streaming API during the initial data collection phase.

The words *geneva, workshop,* and *retrieval* intentionally admitted tweets by authors not involved in information retrieval research. The last three search terms (*iiix, cfp,* and *hcir*) were included because a good deal of discussion about SIGIR treated related conferences and calls for papers. Tracking this activity identified 49 individual users.

To augment this set, I harvested each of the 49 users' on July 17, 2010, for a total of 10,111 users. After removing accounts marked as "private" the corpus contained 9,274 authors in the "community."[8]

In this paper I report data obtained by tracking all tweets written by these 9,274 people beginning at midnight (CDT) on July 18, 2010. Though I continue to store this community's output, for the purposes of this paper, I limit analysis to tweets written before 23:59 (CDT) on August 28, 2010. Summary statistics for this corpus appear in Table 1. To be clear, it bears stressing that the sample is not random.

[8] N.B. I omitted the followers of one core community member because his social network was unmanageably large.

The final 2 rows of Table two report counts for two types of user-generated metadata. Each of these data types is described in detail below. But for the sake of contextualization, a brief definition here is in order. Hashtags are simply character strings that begin with the # character. They are often topical in nature and serve to collocate tweets related to a particular subject. Mentions are strings preceded with the @ sign followed by a Twitter user's screen name. They mark a tweet as being directed to a particular user.

Number of tweets	1,769041
Number of users	9,274
Median tweet length (characters)	88
Median number of tweets per user	90
Median number of friends per user	447
Median number of followers per user	888
Tweets containing at least one hashtag	279,549
Tweets containing at least one mention	1,051,875

Table 2. Summary statistics for corpus used to motivate examples in this article.

Given a corpus of microblog data, the IR literature gives us many approaches to conducting search and retrieval based on textual queries. To situate our discussion, I will rely primarily on the language modeling framework when discussing IR operations (Ponte & Croft, 1998; Zhai & Lafferty, 2004). I will focus on language modeling due to its simplicity, its pervasiveness in the IR literature, and its demonstrated effectiveness. While a full treatment of language modeling IR is beyond the scope of this essay, I will present a brief overview in the following paragraphs.

Language modeling IR relies on generative models to rank documents against queries. Under this approach, we assume that the terms in a document d_i were generated by a particular probability distribution M_i. Given a keyword query q we examine the quantity:

$$\Pr(d_i \mid q) \propto \Pr(q \mid d_i)\Pr(d_i) \qquad \textbf{(1}$$

where the proportionality is due to Bayes' rule and $\Pr(q \mid d_i)$ is taken to be $\Pr(q \mid M_i)$. The prior probability for a document $\Pr(d_i)$ is often assumed to be uniform and is thus omitted. In this case we simply rank documents by the likelihood that their underlying language models generated the words in q.

The likelihood of q given M_i is typically calculated under the assumption that terms are statistically independent. This assumption gives:

$$\Pr(q \mid d_i) = \prod_{w \in q} \Pr(w \mid M_i) \ . \qquad \textbf{(2}$$

To avoid numeric underflow, we typically use log-probabilities, replacing the product in Eq. 2 with a summation of logarithms. In either case, in the language modeling approach, documents whose corresponding language models have a high probability of generating q are ranked higher than those with low probability.

The family and fitting methods for inducing $\Pr(w \mid M_i)$ is treated in the language modeling literature. But typically, M_i is a multinomial distribution over the p terms in the indexing vocabulary, where the distribution is characterized by the multinomial probability distribution function and $\mathbf{\Theta}_i$, a p-dimensional probability vector where θ_{ij} gives the probability of observing the *jth* term on an observation from M_i. We assume that $\mathbf{\Theta}_i$ is estimated using maximum likelihood with Bayesian smoothing using Dirichlet priors and hyperparameter μ=1500. Interested readers may find more detail on estimation, smoothing, and document ranking using language models in (Zhai & Lafferty, 2004).

Two Types of Search in Microblog Systems

The term *microblog IR* easily leads to confusion. This because people search for microblog information in at least two ways:

1. Broadcasting questions to their followers in hopes that people in their social network will answer them.
2. Conducting searches over pre-existing microblog data in an effort to discover relevant information that has already been written.

This distinction has been made before (Sullivan, 2009a, 2009b). Sullivan discusses "how people use Twitter itself…to ask for help directly, especially when in the past, they might have first turned to a search engine." This is in contrast to "Twitter search…,where people can explicitly do a search against past tweets to find information." I will refer to the first of these modes of microblog search as *asking,* while I will call the second approach *retrieving.*

Of course asking and retrieving aren't mutually exclusive. It is conceivable, that a system could aid a user in asking by helping him or her to retrieve a group of people who are likely to lend helpful answers. Alternatively, a person might ask his followers to recommend hashtags that collocate tweets related to his information need. Nonetheless, the difference between asking and retrieving in microblogs is worth our attention.

Asking for Information

The activity that I have called *asking* has been shown to be common in peoples' use of microblogs as well as other social media. Research in this vein has shown many tweets involve asking or answering questions. Morris, Teevan and Panovich (2010) report results from a survey in which half of respondents had used Facebook or Twitter to ask questions. Efron and Winget

(2010) analyzed two corpora of tweets and found that in one corpus 13% of tweets contained a question, while the other corpus consisted of 16% questions. It bears noting, though, that DiMicco *et al.*, using a stricter definition of questions, found only three questions in a sample of 200 tweets (DiMicco et al., 2008).

The information interaction entailed by asking in microblogs is similar to the operations of online Q&A sites such as Yahoo Answers[9] as well as older systems such as email forums. In these domains, the searcher posts a question, typically articulated in some detail to an online space where other users can peruse questions and offer answers. Typically online Q&A sites connect askers and answerers who do not know each other and who may not be identifiable.

Research into online Q&A sites has shown that this approach to addressing information needs, though popular, suffers several drawbacks. Zhang *et al.* report lengthy delays between when questions were asked and when they were answered (J. Zhang, Ackerman, Adamic, & Nam, 2007). In the sample analyzed by Hsieh and Counts 20% of questions were never answered. Of course the quality of answers in online Q&A sites varies, and Shah and Pomerantz found it difficult to predict which answers users would actually find helpful (Shah & Pomerantz, 2010). While Q&A sites are similar to microblog asking, the shortcomings outlined here suggest room for improvement.

Why should microblog asking offer different results than online Q&A services? Asking questions via microblogs seems *prima facie* like nothing more than an informal online Q&A activity. However several points distinguish microblog asking from most established Q&A sites:

1. Questions are posed only to those people who have chosen to follow the asker (unless the question is re-tweeted).

[9] `http://answers.yahoo.com`

2. The asker is identified not only by his or her screen name, but also by the history of tweets that he or she has posted, lending context to a given question.
3. Due to the compressed format of microblogs, questions and answers must be expressed succinctly.

Point 1 above locates microblog asking in the domain of so-called social search (Evans & Chi, 2008; Horowitz & Kamvar, 2010). In the sense suggested by Evans and Chi, social search entails pursuing an information need by consulting (via whatever medium) with people whom the searcher knows. Evans and Chi found that survey respondents often engaged socially before, during, and after the search process itself. They argue that social interaction is a key factor in the search experience.

Discussing a problem with a colleague down the hall and emailing a question to friends are both social search actions. However, social search may also be computationally mediated (Horowitz & Kamvar, 2010), as in the case of the service called Aardvark[10]. Here, a searcher poses a question to the vark system. The system predicts which other users of vark are likely to have expertise on the question's topic, routing the question to people who are likely to provide useful answers.

Asking questions via microblogs shares characteristics with the notion of social search. In the case of microblogs, a searcher's question reaches an audience limited to those people who have chosen to read his or her tweets. If I broadcast a question via Twitter, I am asking a group of friends (loosely construed) for responses. Thus the mode of microblog search that I call *asking* entails an avenue to pursue social search that complements venues such as face-to-face discussion, emailed questions, etc.

[10] `http://vark.com`

Retrieving Information

What I have called retrieving is similar to traditional, ad hoc IR. Interactions of this type are likely to involve a "query" that is posed against an index of microblog data. Based on some matching criteria, the system finds putatively relevant "documents" in the index. These results, organized by algorithmic and design principles, are then presented to the searcher.

I have put the terms *query* and *documents* in quotes above because, as our discussion will show later, what constitutes a query or a document in microblog retrieval systems is quite fluid. Additionally, the matching criteria and result presentation techniques may vary from system to system (perhaps one system returns a ranked list of tweets, while another presents a word cloud). However, the process outlined here captures a typical search process in the context of retrieving microblog data.

Key Problems in Microblog Search

In 2008, Hearst, Hurst and Dumais asked "what should blog search look like?" (Hearst, Hurst, & Dumais, 2008). In 2011 we may ask the same question of microblog search. The fundamental challenges in microblog search have yet to be detailed, as noted by Golovchinsky and Efron who write, "Little attention…has been paid to how people search Twitter, and to how they explore returned search result sets" (2010, p. 2). What information needs do people bring to microblogs and what forms might tractable queries take? What are useful units of retrieval in the microblog setting? What constitutes relevance in a microblog search session? Factors such as influence, authority, and timeliness probably weigh on what makes some microblog posts more useful than others, but how to account for these qualities is non-obvious.

But in setting an agenda for microblog search, the IR community need not start from scratch. The IR literature contains findings, techniques, and vocabularies that serve as good starting points for research on microblogs. The following subsections outline how established problems in information retrieval have made or could make their way into retrieval settings over microblogs.

Sentiment Analysis and Opinion Mining

A key challenge in contemporary text mining is so-called sentiment analysis (Pang & L. Lee, 2008). Sentiment analysis has informed IR research especially in the context of retrieval of blog data, as formalized in the TREC blog track (Macdonald, Ounis, & Soboroff, 2007; Ounis, Macdonald, de Rijke, Mishne, & Soboroff, 2006; Ounis, Macdonald, & Soboroff, 2008). Not surprisingly, findings from the blog retrieval community have ready interpretation in the context of microblogs.

Given a document *D*, the goal of sentiment analysis is twofold. First, sentiment analysis algorithms may classify *D* as either containing opinionated content or not (i.e. taking an objective tone). A related goal of sentiment analysis is to treat a different question: assuming that document *D* does contain opinionated language with respect to topic *T,* does *D* evince a positive or negative opinion of *T* overall? This type of analysis has found its way into information retrieval in effort to support queries such as *find documents that contain favorable or unfavorable discussion of the state of the union speech.*

The majority of sentiment analysis operates by identifying opinion-expressing terms and estimating those terms' importance in the texts to be analyzed. Common sources of opinion-expressing terms are the OpinionFinder lexicon as described in (Wilson, Wiebe, & Hoffmann,

2005) and the General Inquirer lexicon described in (Hatzivassiloglou & McKeown, 1997). These resources are ontologies that enumerate words whose presence in a text constitutes evidence for the expression of emotion or opinion. Precisely how the knowledge encoded in these lexicons is used by a sentiment analysis system varies. But a common approach to identifying the 'semantic orientation' (i.e. positive or negative sentiment) of a text lies in a supervised machine learning setting, where we train a model on labeled data, making predictions on the basis of the target text's evidentiary features.

Like blog data, microblog posts often express opinions or emotion (Cheong & V. Lee, 2009; Diakopoulos & Shamma, 2010; Jansen et al., 2009). Analyzing a corpus of Twitter data, Diakopoulos and Shamma found that "the tenor of the tweets was distinctly negative" (p. 1197). In the context of e-commerce, Jansen et al. write, "microblogs offer immediate sentiment and provide insight in affective reactions toward products…." (p. 2170). Because tweets are often informal and expressive of opinions, the problems of sentiment detection and polarity identification have a clear role to play in microblog retrieval.

For instance, microblog data has shown promise as a gauge of political sentiment. Tumasjan et al. (2010) analyzed a sample of Twitter posts related to a German federal election. Constructing "sentiment profiles" for each candidate, Tumasjan et al. argue that "sentiment profiles of politicians and parties… plausibly reflect many nuances of the election campaign" (p. 183). Similar findings appear in the context of more abstract socio-political issues. In work such as (Bollen, Pepe, & Mao, 2009; O'Connor, Ramnath Balasubramanyan, Routledge, & N. A. Smith, 2010) researchers have found that phenomena such as consumer confidence at a future point in time can be predicted by analyzing Twitter text.

While microblogs' utility in gauging or predicting public opinion is compelling, other research has shown that Twitter data provides leverage in more general sentiment analysis tasks (Pak & Paroubek, 2010). Like other social media, microblogs afford data that allow researchers to answer questions such as *what is the public's current attitude towards universal healthcare?* In an IR setting, the sort of opinion detection that microblogs enable supports queries such as *which are the best Thai restaurants in San Francisco's marina district? Is the movie* Inception *worth seeing?*

Entity Search

One issue that makes microblog retrieval a compelling research area is its youth. Even the most basic problems in microblog IR remain to be identified. Among these basic problems is defining useful units of retrieval. Given a corpus of tweets and a user with a particular information need, precisely what should an IR system present to the searcher?

Early IR systems helped people find literature by searching textual surrogates of documents or books. Ad hoc IR as exemplified by early iterations of the Text REtrieval Conference, as well as a good deal of Web search is predicated on the idea that a retrieval system runs a user's keyword query against a collection of documents whose full text has been indexed, ultimately delivering a ranked list of documents.

Given a microblog corpus and a keyword query, returning individual tweets might be helpful. But it seems likely that showing a user a list of tweets taken out of context is not the best way to solve realistic information needs. If we suspect that a ranked list of tweets is not the best (or at least not the only) way to retrieve and present microblog information, what approach would be better?

We can find one answer to the question *what is a useful unit of retrieval in microblog IR* in the established field of entity search. Entity search is a type of IR in which we begin with a corpus of documents but return information about objects or actors that exert influence in those documents. Perhaps the best-known type of entity search is expert finding where the unit of retrieval is a person who has expertise on a topic indentified by a user's query. For example, Balog and de Rijke discuss the problem of finding topical experts in corpora of email (2006), a challenge that saw a great deal of attention in the context of the TREC enterprise track (Balog et al., 2008) and which has continued to generate research (Balog, Azzopardi, & de Rijke, 2006, 2009).

Entity search plays a natural role in microblog IR, where the appropriate unit of retrieval is likely to depend on the type of information need that the searcher brings to the search session. Current microblogging services such as Twitter are based on a publish/follow model where a person reads posts that are written by other people whom the reader "follows." This dynamic leads to the obvious problem of finding suitable people to follow. A keen question for users of microblogs is, *whose Twitter stream should I read and why?* Or, *I am interested in topic X. Who posts regularly and with authority on that topic?*

A key problem in expert finding is the matter of author representation. Given a corpus C containing N documents, N_a of which were written by author a, an expert finding system must induce a model for a's topical output. The task of finding people to follow in a microblogging environment presents a nearly identical problem.

Several methods of expert finding have been proposed in the IR literature. In what has been termed *the virtual document* approach, an author is represented simply by the totality of documents that he or she has written (Balog et al., 2006; Craswell, Hawking, Vercoustre, &

Wilkins, 2004). That is, for each author, we construct a 'virtual document' that consists of the concatenated text of his or her N_a documents. Based on this virtual document, we may conduct retrieval as usual. Given a query q and a corpus of virtual documents built from a collection of posts, we may rank the virtual documents in decreasing order of query likelihood using Eq. 2.

An alternative approach is offered in (Macdonald, 2009), where several voting models are applied to the expert search problem. In these approaches, we perform a search against a corpus of documents (not virtual documents). We then process the result set by analyzing the authors of each document. As we proceed, each author is given a 'vote' when we encounter a document written by him or her. Precisely how we tally the vote varies from method to method—e.g. simple majority, Borda counting, Condorcet voting—but the idea is the same: authors whose output is strongly represented in topically relevant documents are likely to be experts in the area specified by the query.

While finding people is certainly an important problem in microblog IR, entities lend themselves to consideration for search in other microblog settings as well. Lists of authors, groups of tweets that comprise a multi-tweet conversation, question and answer pairs are all candidates for treatment of this kind.

User-Generated Metadata

A sea change in peoples' interaction with information came in the early 2000's. This decade saw the development and maturation of social media, online environments that allow people to interact in novel ways. With respect to information science, a key development brought by these media was the proliferation of user-generated content. In contrast to early Web resources, newer services (characterized by the ubiquitous label Web 2.0) allow and often rely upon their users to

create, edit, and mediate access to information. The popularity of user-generated content, particularly in the context of social media has spawned a host of novel IR problems (King, Li, Xue, & Tang, 2009).

An important aspect of microblog ecology is authors' use of informal metadata. Services such as Twitter make few demands on content. But in the course of their interactions, user communities have invented and adopted a variety of metadata conventions aimed at extending their texts' expressiveness.

For instance, the convention of using so-called hashtags has become pervasive in the Twitter community[11]. Hashtags are simply character strings preceded by a # sign. Peoples' reasons for applying hashtags vary from enhancing topical access to lending tweets rhetorical flourish. For instance, many people tweeting from the 2010 SIGIR conference added the hashtag `#sigir2010` to their posts in order to help others collocate this information. On the other hand, a SIGIR attendee added the tag `#genevaishotandhasnoairconditioning` to one tweet. Clearly the motivation for using these hashtags differs. But hashtags, whatever their reason for appearing, do constitute a privileged class of microblog content.

Less widely used, but similar to hashtags, are metadata that Chris Messina proposed (Messina, 2009) and which blogger Chris Blow terms "slash tags" (Blow, 2009). A slash tag simply begins with the `/` character. Messina originally proposed a modest three slash tags: `/via`, `/by`, and `/cc` (a re-posting of someone's earlier tweet, a citation to a possibly external resource, and carbon copy, respectively). Only 7,172 tweets in the corpus used in this paper contain at least one of these slash tags.

---

[11] The motivation and convention for using hashtags was described by the blogger Chris Messina (Messina, 2007).

Precisely how to marshal these user-generated metadata is an open question. But the matter has seen some treatment in the literature.

In (Efron, 2010) Efron describes the problem of 'hashtag retrieval,' a type of entity search. The goal of hashtag retrieval is to find, for a topical query *q,* a ranked list of hashtags, such that highly ranked tags have a strong tendency to mark tweets on *q*'s topic. This task has at least three potential uses:

1. *Ad hoc tag retrieval:* Help a searcher find tags that are often applied to tweets on a topic that he or she wishes to keep abreast of. This allows the user to 'follow' particular tags.
2. *Query expansion:* In the paper cited above, Efron found that during pseudo-relevance feedback, hashtags provide especially strong data for query expansion. By limiting expansion terms to hashtags, a significant improvement in three effectiveness metrics was observed with respect to both a baseline (non-expanded) model, as well as a model using standard terms for query expansion.
3. *Result display:* Hashtags could be used to arrange results of searches for tweets (or other entities), providing a de facto clustering mechanism for organizing returned documents.

Hashtag retrieval is a concrete example of the entity search problem described in the previous subsection. Given a query, we wish to find a ranked list of hashtags, where our knowledge about each hashtag is induced from an analysis of its use in observed tweets. A natural way to approach this problem is via language modeling, using Eq. 2. Here we create a virtual document for a hashtag *h*, where the virtual document consists of the concatenated text of all tweets containing *h*. We use this virtual document to estimate the parameters of a language model. Retrieval then proceeds by calculating the (log) likelihood of the query given the estimated model for each hashtag in our corpus.

hadoop nosql	semantic web	gulf oil spill
-10.20 #hadoop	-10.43 #semantic	-10.86 #oilspill
-10.53 #nosql	-10.94 #linkeddata	-12.30 #blacktide
-11.87 #hbase	-11.38 #semweb	-12.81 #eco
-12.36 #bigdata	-12.17 #rdfa	-13.04 #environment
-12.38 #mapreduce	-12.88 #a11y	-13.47 #ocean

Table 3. Results for three ad hoc hashtag searches. Each column lists the top five hashtags retrieved for the query shown in its heading, along with the query log-likelihood calculated from each tag's estimated language model.

Table 3 shows the top five results of three ad hoc searches over hashtags in the data described in Table 1. Without investigation into the typical usage of each tag, it is difficult to assess the quality of these retrievals. They do, however, appear to be plausibly useful to a person who is interested in the topics that generated each query. For instance, in the summer of 2010, a person eager to keep abreast of news on the Gulf oil spill might indeed have wished to read tweets that has been tagged with `#blacktide` by their authors.

Other user-embedded metadata involves explicit representations of social linkages. An author may direct a tweet to a particular user by including `@<user>` in the tweet where `<user>` is the screen name of the person to whom the tweet is directed. Huberman, Robero and Wu find that about 25 percent of tweets contain an `@` direction (Huberman, Robero, & Wu, 2009). Most third-party microblog clients display any `@` mentions directed to their owner prominently, even if the owner does not follow the tweet's author.

Direction via the @ sign is used in a variety of ways on Twitter. It is common to see tweets with a structure such as:

@milesEfron are you nearly done with your AIS article?

Likewise, we often seen tweets with this type of structure:

Congratulations to @userX on his new article in JASIS&T.

In the first example, the tweet author is asking a question of the particular user *milesEfron*. It is worth noting, that such a message will be visible to all of the author's followers (though not all of *milesEfron*'s followers). Presumably, the tweet author believes that this message is at least of passing interest to his or her other followers[12].

The second example sheds some light on the motivation for broadcasting `@`-directed messages. Tweets of this kind alert other users that a social linkage exists between the author and the user to whom the message is sent. Thus such a message demonstrates a social bond between two users, a function alluded to in (boyd, 2009). Additionally, `@` mentions allow followers of a tweet's author to learn about potentially interesting users (i.e. the user to whom the message is directed).

The `@` directive may also be applied in a more rhetorical fashion, along the lines of:

Why is the @twitter search failing to find tweets from last month?

In this case, the author directs the message to Twitter's account, although there is no implication that Twitter follows this author. In some sense this mention is acting like a hashtag. In all likelihood, managers of the `twitter` user account track all such mentions, 'following' them via a persistent search. More rhetorically, the author of this tweet is signaling that Twitter is especially important to the topical nature of the text. Here, in other words, it is difficult to articulate the difference between discussing `@twitter search` and `#twitter search`.

[12] Authors may contact other users privately using a so-called *direct message.*

These ambiguities notwithstanding, microblog posts are often replete with metadata despite their brevity. Marshaling these metadata to improve retrieval entails a promising avenue in microblog IR research.

Authority and Influence

It is well known that modern Web retrieval relies on hyperlink structure to induce models of authority that help deliver retrieval sets that are not only on-topic but that are also of high quality. Algorithms such as PageRank (Brin & Page, 1998) and HITS (Kleinberg, 1999) build models of authority by finding the steady state of differently defined Markov processes. The intuition behind these approaches is that a given resource is likely to be of high quality if it has inbound links from many other high-quality resources.

A similar intuition has found its way into microblog analysis. The matter has been most widely studied in the estimation of the influence of particular people in microblog environments. In a 2009 blog post (Tunkelang, 2009) Daniel Tunkelang proposed "a Twitter analog to PageRank." His metric, dubbed TunkRank, is indeed similar to PageRank, lending each person influence by summing the influence of the people who follow him or her. Thus the influence of a user *u* would be estimated by:

$$Influence(u) = \sum_{v \in Followers(u)} \frac{1 + p * Influence(v)}{\|Following(v)\|} \quad \textbf{(3}$$

where *Followers(.)* is the set of people following a given user and *Following(.)* is the set of people a given user follows and *p* is a real-valued number corresponding to the chances that a

given tweet is re-tweeted. A person's TunkRank score is similar to a Web resource's PageRank, where the score reflects the probability that a tweet by that person will be read. Though TunkRank isn't a proper probability, it could easily be normalized to behave as one, allowing us to incorporate its model of influence into document ranking by using Eq. 2 as the prior probability of a tweet in Eq. 1.

TunkRank is only one among many so-called 'peopleRank' metrics (Kamvar, Schlosser, & Garcia-Molina, 2003; C. Lee, Kwak, Park, & Moon, 2010; Weng, Lim, Jiang, & He, 2010), but in recent research it showed robustness for the problem of identifying accounts that post spam on Twitter (Gayo-Avello, 2010). Gayo-Avello found that among five prestige algorithms based on users' social graphs, TunkRank was the least sensitive to the presence of 'nepotistic links,' follower/following relationships that give a false impression of a user's network centrality.

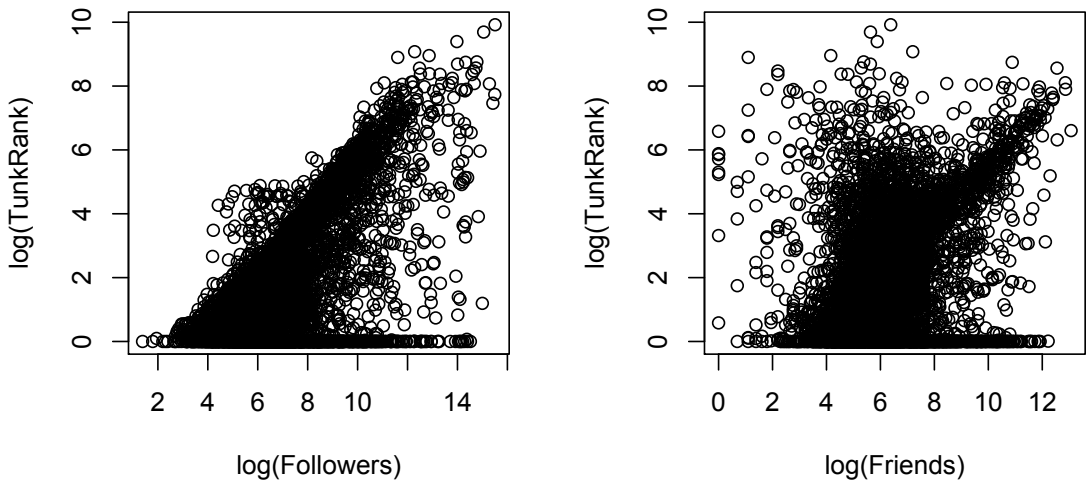

Figure 2. The relationship between users' number of followers (friends) and TunkRank. The *x*-axis of the left panel is the log of the number of people who follow each user. The *y*-axis gives the log TunkRank. The right panel gives data for each user's number of "friends" (i.e. people whom they follow).

Figure 2 visualizes the relationship between users' TunkRank scores[13] and their immediately observable social statistics—number of friends and number of followers on Twitter. The panels in Figure 2 show that a positive correlation between both friend- and follower-count and TunkRank exists[14], but that popularity does not necessarily translate to strong influence as measured by TunkRank. In particular, the bottom of each panel shows that people with a great breadth of friends or followers are non-influential according to TunkRank, a fact that bears on the results reported by Gayo-Avello.

Temporal Issues

No branch of IR can afford to neglect the matter of time in information interactions. But time has a prominence in microblog search that is especially keen. A great deal of recent research and development treats so-called real-time search. The term real-time search refers to the problem of keeping indexes and the search engines they support current to within a very low tolerance for latency. However, time plays a role in microblog IR that is not limited to real-time search. Many tweets treat topics that have a limited time horizon; conferences come and go, news stories fade in importance, newly released movies are eclipsed by still newer movies. Additionally, people's social networks change over time, providing information that could prove useful in retrieval.

Twitter has brought new research areas into the broad field of real-time search. The flood of data coming from Twitter and other social networking services has brought new challenges to commercial search engines (not least among them, Twitter's own search service). According to David Geer, search engine developers hope to bring microblog data into their indexes quickly both to enable direct access to those data, and to improve results for other time-critical arenas

---

[13] TunkRank scores were obtained from (Adams, 2009).
[14] The Pearson correlation between log-followers and log-TunkRank is 0.519. Correlation between log-friends and log-Tunkrank is 0.264.

such as news search (Geer, 2010). Similar approaches have been treated in (Dong et al., 2010). But it should be stressed that while microblogs pose novel challenges for search engines, real-time search follows on an established body of literature on so-called streaming algorithms (Alon, Matias, & Szegedy, 1996; Charikar, Chekuri, Feder, & Motwani, 1997; Charikar, O'Callaghan, & Panigrahy, 2003; Henzinger, Raghavan, & Rajagopalan, 1999).

Immediacy is one important facet of time's role in microblog search. However, temporal concerns enter into microblog IR in other senses, too. For instance, Lee et al. bring a temporal approach to bear on the problem of user influence modeling as described above (C. Lee et al., 2010). They note that while social networks play an important role in user influence, users who report important stories early in the stories' life cycle are likely to be especially influential. Alonzo, Gertz and Baeza-Yates also introduce a nuanced temporal dimension into IR (2009). They use documents' lexical content to construct "temporal profiles" that in turn allow them to cluster search hit results (in Twitter and other domains) by time.

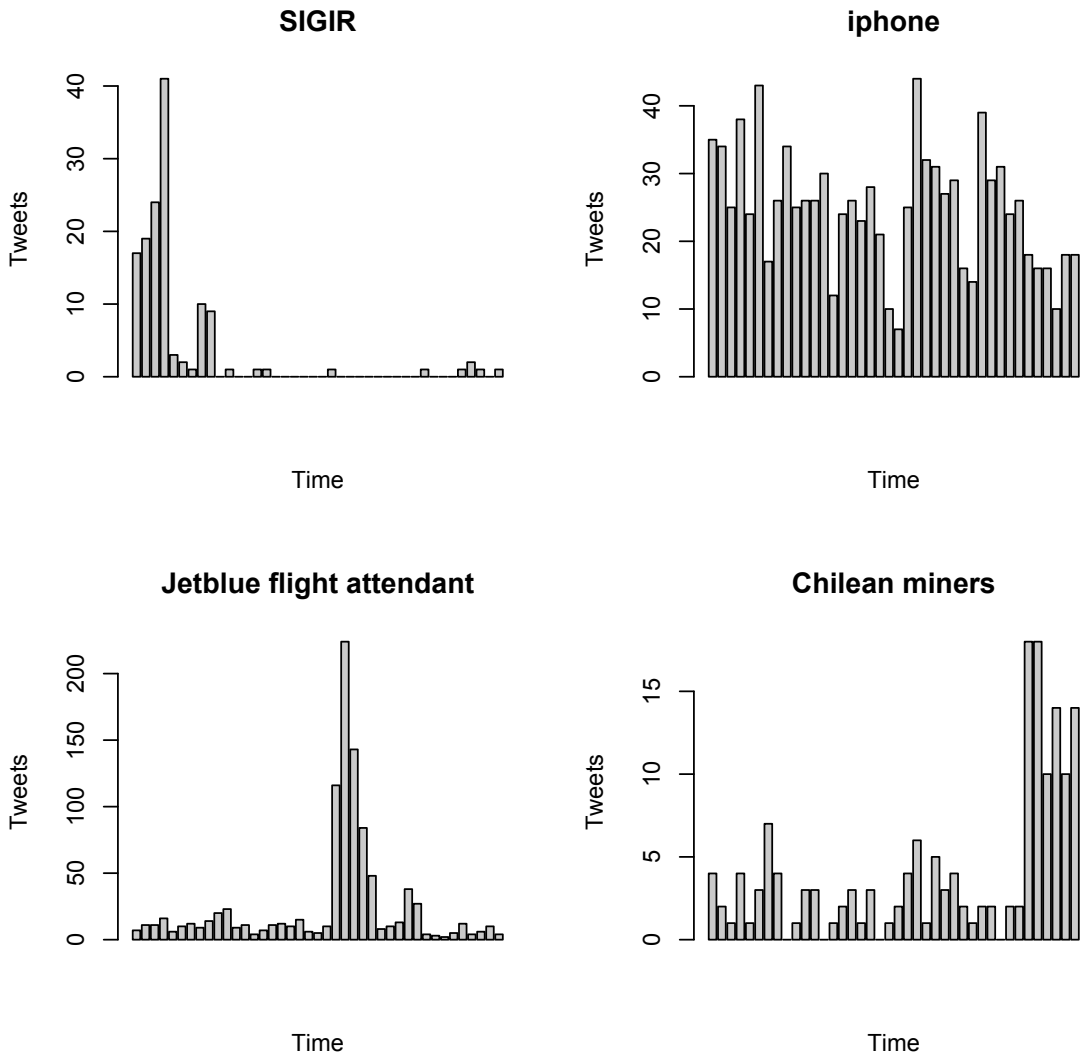

Figure 2. Frequency of hits on four searches. Taken from the data described in Table 1. The *x*-axis of each panel is the date (from July18-August 28, 2010). Each bar is the number of tweets returned for the query shown in the panel title on each day.

Figure 2 suggests another way in which time bears on microblog search. The figure shows the results of four queries posed against the corpus described in Table 2. The number of hits (i.e. tweets) are displayed in bar plots, where each bar is a day (the left of the *x*-axis is July 18 and the right is August 28). Each bar's height gives the number of tweets in our corpus matching the query that were written on the corresponding day.

The panels in Figure 2 show several distinct distributions with respect to time. The queries *sigir* and *jetblue flight attendant* are strongly unimodal. *sigir* was mentioned commonly during the 2010 ACM SIGIR conference. The distribution for *jetblue flight attendant* shows a similar rise and drop of activity. But instead of a formal event, the peak for this query owes its strength to the outburst of a flight attendant on a Jetblue Airline flight, an event that drew strong and amused attention on Twitter. A query such as *iphone* has a less visibly temporal character, with a number of daily hits that does not change systematically over time (although the distribution does appear to have several weakly evident modes). Finally the story of the miners stranded underground in Chile emerged immediately after the cave-in that trapped them. Though the graph ends mid-story, we can see that the hits for this query rise suddenly. After this point, the count decays slowly: the story continues, but discussion of it gradually wanes.

Figure 2 suggests that time plays a strong role not only in finding trending topics, but also in terms of more retrospective queries. We might, for instance, wish to arrange tweets returned for a given query differently for a news story than for a query such as *iphone*, or even for a query about an academic conference. In addition to fielding the common "recency queries" that we anticipate in a microblog environment, we might also support queries showing precedence such as, *who were the first people to report on the 2010 earthquake in Haiti?*

Concluding Remarks: Continuing Issues in Microblog Search

In this article I have chosen to focus on a few topics in some detail. Of course this approach comes at the expense of a treatment of many other important considerations. By way of conclusion, this section discusses several problems that I believe are important, but that have yet to be extensively studied.

Geographical Information in Microblog Search.

Locale bears on microblog search in several ways. Most obviously, relevance for many plausible queries over microblog data will be contingent on geographical concerns (this is no different from many other types of IR). A query such as *movie times* issued in San Francisco is ostensibly different from the same lexical query issued from Champaign, Illinois. Queries such as *downtown traffic, allergen levels,* and *where are all of these ambulances going* all entail implicit geographical references. Earlier in this article I discussed user-generated metadata in microblogs. Geographic metadata also finds its way into microblog posts in the form of the latitude and longitude coordinates from which a given tweet was sent—the so-called geocode data for the tweet. Though many users opt not to include geocode data in their tweets, the ability to insert this information is available in many microblog client applications.

Microblogs, Information Needs, and Queries

In our discussion of how people use Twitter to ask questions, I reviewed literature concerned with the nature of the information needs that people bring to microblogs. But these findings do not readily translate to the other sense in which we considered IR—retrieving existent microblog data. Teevan, Ramage and Morris (2011) show that queries issued to a Twitter search engine differ in marked ways from standard Web queries. Twitter queries, these authors find, tend to be shorter than Web queries. They also tend to be repeated often. Twitter queries often relate to people—Twitter users or celebrities, in particular. People use microblog search, the authors found, largely to stay abreast of ongoing or emerging topics. While these findings shed crucial and actionable light on microblog search, we should also bear in mind that observing users' searches lends information about interactions with systems as they are now. But how might interactions change if microblog search systems operated in qualitatively different ways?

One issue that complicates our anticipation of how people might search microblogs lies in the fluidity of the unit of retrieval in these settings. It is true that we could retrieve and rank tweets in decreasing likelihood (based on lexical evidence) that they will satisfy a person's information need. But the entity search scenarios that we have outlined invite a different experience and thus a different type of information need. For example, it is one thing to build a system that lets people find tweets that mention a particular piece of consumer electronics. It is another matter to help people find people or communities of people who write often and clearly on the matter of consumer electronics.

Because microblog search is in its infancy, it is important to be mindful of realistic, compelling use cases now. As we build pilot systems and undertake research agendas, we implicitly take as axiomatic what it is that people will search for and how they will engage in this searching. Queries by example over a variety of entities using many criteria (e.g. time, place, topic, authority) for organizing results seem like plausible interactions, at least as plausible as simple keyword searches for indexed tweets. There is a great deal of creativity to be applied to design and research in the space of supporting user needs in microblog search.

Microblog Search Evaluation

For many years variations on the Cranfield model of evaluation have allowed IR researches to assess the merits of proposed innovations (Voorhees, 2007). The Cranfield model, with its reliance on access to canonical corpora, queries and corresponding relevance judgments, has always engendered debate with respect to its relation to actual search effectiveness. Nonetheless, Cranfield-style experiments have dominated modern IR research in an era that has seen innovations in the field that are difficult to discount.

Assessing the effectiveness of searches in microblog environments does not preclude Cranfield-type analysis. Indeed we could craft a set of queries to run against a corpus of documents (or entities) and acquire query-document relevance judgments on which to base calculations of such familiar statistics as precision and recall. But a number of issues must guide such an approach:

1. *Relevance.* The matter of how relevant a document is to a query has always been crucial to successful Cranfield-style evaluation and it has always been controversial; microblogs don't change this. What is different in the context of microblogs is the array of new problems and criteria that enter into the matter of retrieval. We know that what task a user is trying to complete bears on what constitutes relevance. To this consideration we must add factors such as temporality, locality, authority, etc., which I have discussed already in this article. All of these factors are conflated in a naturalistic idea of successful retrieval. This does not qualitatively set microblogs apart from other data. But it is the case that microblogs' idiosyncrasies merit careful deliberation in designing experimental settings.
2. *Corpora.* Thanks to microblog services' APIs, acquiring microblog data is easy. However, it is not clear how we should gather these data when creating corpora for IR experimentation (using the Cranfield model or any other approach). The putatively random sample that comes from the Twitter API's "garden hose" feed gives us a cross-section of Twitter data. Setting aside the question of how random this sample actually is, we might also ask if a random sample is desirable. The answer to this question surely hinges on what we are studying. Analyses that take into account the experience of particular users, for instance, should presumably harvest data that comprises those users' "social horizons" (i.e. the people writing "close to" that person, with respect to social network proximity, topic, geographic location, etc.).
3. *Recency.* If we built a microblog test collection, would it be useful in a year? Would it be useful in ten years? Of course any test collection ages: the Reagan-era queries of the

> early TREC collections have little topical resonance today. Yet these early news wire collections still have obvious value. It is less clear that the work of building a microblog test collection would yield the enduring value that we have seen with the TREC collections. Not only would topics of interest come and go, but peoples' use of microblogs would be likely to change over time. Today we use hashtags and mentions in our tweets. How will we tweet a year from now? Different conventions and different rhetorical and topical preoccupations stand to make microblogs (and queries we would pose against them) very different over a short time.

None of this is to say that Cranfield-style evaluation does not have a place in microblog IR research. Surely it does. But we should be strategic in crafting assessment methodologies at this early stage of research and development in microblog retrieval. Serious consideration of naturalistic and behavioral methods of assessing system performance will no doubt have a large impact on future research, as we work to make our studies both realistic and generalizaeble. The organization of the first TREC microblog track is underway at the time of this writing. In this setting and in the work of the field of microblog researchers this agenda will begin to emerge.

Microblogs form part of the vital and rapidly changing landscape of technologies that mediate peoples' contemporary interactions with information. People post to Twitter; they write on Facebook; they search the Web, browsing articles on Wikipedia. They do all this while also finding information offline, via face-to-face relationships, in newspapers, in libraries, etc. Building search services using microblog data is challenging for many reasons outlined in this article. But chief among those challenges is the fact that microblogs already act as de facto, informal search services. They provide venues for people to ask questions, answer questions, and make suggestions. As microblog IR research moves forward it will be important to avoid simply mimicking this "guerilla" social search function of microblogs. The question that faces us is,

given so much data created by so many actively engaged people, how can we apply our expertise as IR researchers to identify and to solve information problems whose solutions are latent in the microblog arena?

References

Adams, J. (2009). TunkRank::Measuring Influence by How Much Attention Your Followers Can Give You. Retrieved September 29, 2010, from http://tunkrank.com/

Alon, N., Matias, Y., & Szegedy, M. (1996). The space complexity of approximating the frequency moments. In *Proceedings of the twenty-eighth annual ACM symposium on Theory of computing* (pp. 20-29). Philadelphia, Pennsylvania, United States: ACM. doi:10.1145/237814.237823

Alonso, O., Gertz, M., & Baeza-Yates, R. (2009). Clustering and exploring search results using timeline constructions. In *Proceeding of the 18th ACM conference on Information and knowledge management* (pp. 97-106). Hong Kong, China: ACM. doi:10.1145/1645953.1645968

Balog, K., Azzopardi, L., & de Rijke, M. (2006). Formal models for expert finding in enterprise corpora. *SIGIR '06: Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval*, 43-50.

Balog, K., Azzopardi, L., & de Rijke, M. (2009). A language modeling framework for expert finding. *Information Processing & Management*, *45*(1), 1-19.

Balog, K., & de Rijke, M. (2006). Finding experts and their eetails in e-mail corpora. *WWW '06: Proceedings of the 15th international conference on World Wide Web*, 1035-1036.

Balog, K., Soboroff, I., Thomas, P., Bailey, P., Craswell, N., & de Vries, A. P. (2008). Overview of the TREC 2008 Enterprise Track. *Proceedings of TREC 2008*.

Beaumont, C. (2010, February 23). Twitter users send 50 million tweets per day. *telegraph.co.uk*.

Retrieved August 24, 2010, from http://www.telegraph.co.uk/technology/twitter/7297541/Twitter-users-send-50-million-tweets-per-day.html

Blow, C. (2009, November 11). "Slashtags" for citizen journalists. *Unthinkingly.com*. Retrieved August 8, 2009, from http://unthinkingly.com/2009/11/09/slashtags-for-citizen-editors/

Bollen, J., Pepe, A., & Mao, H. (2009). Modeling public mood and emotion: Twitter sentiment and socio-economic phenomena. *0911.1583*. Retrieved from http://arxiv.org/abs/0911.1583

boyd, D. (2009, August 16). Twitter: "pointless babble" or peripheral awareness + social grooming. *apophenia*. Retrieved August 24, 2010, from http://www.zephoria.org/thoughts/archives/2009/08/16/twitter_pointle.html

Brin, S., & Page, L. (1998). The anatomy of a large-scale hypertextual Web search engine. *Comput. Netw. ISDN Syst.*, *30*(1-7), 107-117.

Charikar, M., Chekuri, C., Feder, T., & Motwani, R. (1997). Incremental clustering and dynamic information retrieval. In *Proceedings of the twenty-ninth annual ACM symposium on Theory of computing* (pp. 626-635). El Paso, Texas, United States: ACM. doi:10.1145/258533.258657

Charikar, M., O'Callaghan, L., & Panigrahy, R. (2003). Better streaming algorithms for clustering problems. In *Proceedings of the thirty-fifth annual ACM symposium on Theory of computing* (pp. 30-39). San Diego, CA, USA: ACM. doi:10.1145/780542.780548

Cheong, M., & Lee, V. (2009). Integrating web-based intelligence retrieval and decision-making from the twitter trends knowledge base. In *Proceeding of the 2nd ACM workshop on Social web search and mining* (pp. 1-8). Hong Kong, China: ACM. doi:10.1145/1651437.1651439

Craswell, N., Hawking, D., Vercoustre, A., & Wilkins, P. (2004). Panoptic expert: Searching for experts not just for documents. *Proceedings of the 7th Australasian World Wide Web*

Conference (AusWeb-04), (poster).

Diakopoulos, N. A., & Shamma, D. A. (2010). Characterizing debate performance via aggregated twitter sentiment. In *Proceedings of the 28th international conference on Human factors in computing systems* (pp. 1195-1198). Atlanta, Georgia, USA: ACM. doi:10.1145/1753326.1753504

DiMicco, J., Millen, D. R., Geyer, W., Dugan, C., Brownholtz, B., & Muller, M. (2008). Motivations for social networking at work. In *Proceedings of the 2008 ACM conference on Computer supported cooperative work* (pp. 711-720). San Diego, CA, USA: ACM. doi:10.1145/1460563.1460674

Dong, A., Zhang, R., Kolari, P., Bai, J., Diaz, F., Chang, Y., Zheng, Z., et al. (2010). Time is of the essence: improving recency ranking using Twitter data. In *Proceedings of the 19th international conference on World wide web* (pp. 331-340). Raleigh, North Carolina, USA: ACM. doi:10.1145/1772690.1772725

Douglas, N. (2007, March 12). Twitter blows up at SXSW Conference. *Gawker*. Retrieved August 24, 2010, from http://gawker.com/tech/next-big-thing/twitter-blows-up-at-sxsw-conference-243634.php

Efron, M. (2010). Hashtag retrieval in a microblogging environment. In *Proceeding of the 33rd international ACM SIGIR conference on Research and development in information retrieval* (pp. 787-788). Geneva, Switzerland: ACM. doi:10.1145/1835449.1835616

Efron, M., & Winget, M. (2010). Questions are content: A Taxonomy of Questions in a Microblogging Environment. *Proceedings of the 2010 Annual Meeting of the American Society for Information Science and Technology*.

Evans, B. M., & Chi, E. H. (2008). Towards a model of understanding social search. In *Proceedings of the 2008 ACM conference on Computer supported cooperative work* (pp. 485-494). San Diego, CA, USA: ACM. doi:10.1145/1460563.1460641

Evgeny, M. (2009). Iran: Downside to the "Twitter Revolution". *Dissent*, *56*(4), 10-14.

Gayo-Avello, D. (2010). Nepotistic Relationships in Twitter and their Impact on Rank Prestige Algorithms. *1004.0816*. Retrieved from http://arxiv.org/abs/1004.0816

Geer, D. (2010). Is It Really Time for Real-Time Search? *Computer*, *43*(3), 16-19. doi:10.1109/MC.2010.73

Grossman, L. (2009, June 17). Iran Protests: Twitter, the Medium of the Movement. *Time.com*. Retrieved August 24, 2010, from http://www.time.com/time/world/article/0,8599,1905125,00.html

Hatzivassiloglou, V., & McKeown, K. R. (1997). Predicting the semantic orientation of adjectives. *In Proceedings of the 8th conference on european chapter of the association for computational linguistics*, 174-181.

Hearst, M. A., Hurst, M., & Dumais, S. T. (2008). What should blog search look like? In *Proceeding of the 2008 ACM workshop on Search in social media* (pp. 95-98). Napa Valley, California, USA: ACM. doi:10.1145/1458583.1458599

Henzinger, M. R., Raghavan, P., & Rajagopalan, S. (1999). Computing on data streams. *DIMACS series in Discrete Mathematics and Theoretical computer Science*, *50*, 107--118.

Horowitz, D., & Kamvar, S. D. (2010). The anatomy of a large-scale social search engine. In *Proceedings of the 19th international conference on World wide web* (pp. 431-440). Raleigh, North Carolina, USA: ACM. doi:10.1145/1772690.1772735

Huberman, B., Robero, D. M., & Wu, F. (2009). Social networks that matter: Twitter under the microscope. *First Monday*, *14*(1-5). Retrieved from http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/2317/2063

Jansen, B. J., Zhang, M., Sobel, K., & Chowdury, A. (2009). Twitter power: Tweets as electronic word of mouth. *J. Am. Soc. Inf. Sci. Technol.*, *60*(11), 2169-2188.

Java, A., Song, X., Finin, T., & Tseng, B. (2007). Why we twitter: understanding microblogging usage and communities. In *Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 workshop on Web mining and social network analysis* (pp. 56-65). San Jose, California:

ACM. doi:10.1145/1348549.1348556

Kamvar, S. D., Schlosser, M. T., & Garcia-Molina, H. (2003). The Eigentrust algorithm for reputation management in P2P networks. In *Proceedings of the 12th international conference on World Wide Web* (pp. 640-651). Budapest, Hungary: ACM. doi:10.1145/775152.775242

Kelly, R. (2009). *Twitter Study Reveals Interesting Results about Usage*. San Antonio, TX USA: Pear Analytics. Retrieved from http://www.pearanalytics.com/blog/wp-content/uploads/2010/05/Twitter-Study-August-2009.pdf

King, I., Li, J., Xue, G., & Tang, J. (2009). Proceeding of the 2nd ACM workshop on Social web search and mining (p. 70). Hong Kong, China: ACM. Retrieved from http://portal.acm.org/citation.cfm?id=1651437&coll=Portal&dl=GUIDE&CFID=99414250&CFTOKEN=48743584

Kleinberg, J. M. (1999). Authoritative sources in a hyperlinked environment. *J. ACM*, *46*(5), 604-632. doi:10.1145/324133.324140

Krishnamurthy, B., Gill, P., & Arlitt, M. (2008). A few chirps about twitter. In *Proceedings of the first workshop on Online social networks* (pp. 19-24). Seattle, WA, USA: ACM. doi:10.1145/1397735.1397741

Lee, C., Kwak, H., Park, H., & Moon, S. (2010). Finding influentials based on the temporal order of information adoption in twitter. In *Proceedings of the 19th international conference on World wide web* (pp. 1137-1138). Raleigh, North Carolina, USA: ACM. doi:10.1145/1772690.1772842

Longueville, B. D., Smith, R. S., & Luraschi, G. (2009). "OMG, from here, I can see the flames!": a use case of mining location based social networks to acquire spatio-temporal data on forest fires. In *Proceedings of the 2009 International Workshop on Location Based Social Networks* (pp. 73-80). Seattle, Washington: ACM. doi:10.1145/1629890.1629907

Macdonald, C. (2009). *The Voting Model for People Search* (Doctoral Dissertation). University of Glasgow. Retrieved from http://www.dcs.gla.ac.uk/~craigm/thesis.shtml

Macdonald, C., Ounis, I., & Soboroff, I. (2007). Overview of the TREC 2007 Blog Track. *Proceedings of TREC 2007*. Retrieved from http://trec.nist.gov/pubs/trec16/papers/BLOG.OVERVIEW16.pdf

Messina, C. (2007, August 25). Groups for Twitter; or A Proposal for Twitter Tag Channels. *FactoryCity*. Retrieved August 8, 2010, from http://factoryjoe.com/blog/2007/08/25/groups-for-twitter-or-a-proposal-for-twitter-tag-channels/

Messina, C. (2009, November 8). New microsyntax for Twitter: three pointers and the slasher. *FactoryCity*. Retrieved August 8, 2010, from http://factoryjoe.com/blog/2009/11/08/slashtags/

Morris, M. R., Teevan, J., & Panovich, K. (2010). What do people ask their social networks, and why?: a survey study of status message q&a behavior. In *Proceedings of the 28th international conference on Human factors in computing systems* (pp. 1739-1748). Atlanta, Georgia, USA: ACM. doi:10.1145/1753326.1753587

O'Connor, B., Ramnath Balasubramanyan, Routledge, B. R., & Smith, N. A. (2010). From Tweets to Polls: Linking Text Sentiment to Public Opinion Time Series. Text.Proceedings, . Retrieved August 12, 2010, from http://www.aaai.org/ocs/index.php/ICWSM/ICWSM10/paper/viewPaper/1536

Ounis, I., Macdonald, C., de Rijke, M., Mishne, G., & Soboroff, I. (2006). Overview of the TREC 2006 Blog Track. *Proceedings of TREC 2006*. Retrieved from http://trec.nist.gov/pubs/trec15/papers/BLOG06.OVERVIEW.pdf

Ounis, I., Macdonald, C., & Soboroff, I. (2008). Overview of the TREC 2008 Blog Track. *Proceedings of TREC 2008*. Retrieved from http://trec.nist.gov/pubs/trec17/papers/BLOG.OVERVIEW08.pdf

Pak, A., & Paroubek, P. (2010). Twitter as a corpus for sentiment mining. *Proceedings of the Seventh conference on International Language Resources and Evaluation (LREC'10)*, 1320-1326.

Pang, B., & Lee, L. (2008). *Opinion Mining and Sentiment Analysis*. Foundations and Trends in Information Retrieval (Vol. 1). Now Publishers, Delft.

Ponte, J. M., & Croft, W. B. (1998). A language modeling approach to information retrieval. *SIGIR 1998: Proceedings of the 21st annual international ACM SIGIR conference on Research and development in information retrieval*, 275-281.

Sakaki, T., Okazaki, M., & Matsuo, Y. (2010). Earthquake shakes Twitter users: real-time event detection by social sensors. In *Proceedings of the 19th international conference on World wide web* (pp. 851-860). Raleigh, North Carolina, USA: ACM. doi:10.1145/1772690.1772777

Shah, C., & Pomerantz, J. (2010). Evaluating and predicting answer quality in community QA. In *Proceeding of the 33rd international ACM SIGIR conference on Research and development in information retrieval* (pp. 411-418). Geneva, Switzerland: ACM. doi:10.1145/1835449.1835518

Shamma, D. A., Kennedy, L., & Churchill, E. F. (2009). Tweet the debates: understanding community annotation of uncollected sources. In *Proceedings of the first SIGMM workshop on Social media* (pp. 3-10). Beijing, China: ACM. doi:10.1145/1631144.1631148

Sullivan, D. (2009a, March 13). How We Search With The Twitter "Help Engine". *Search Engine Land*. Retrieved September 7, 2010, from http://searchengineland.com/how-we-search-with-twitter-16920

Sullivan, D. (2009b, March 13). The Rise of Help Engines: Twitter & Aardvark. *Search Engine Land*. Retrieved September 7, 2010, from http://searchengineland.com/the-rise-of-help-engines-16921

Teevan, J., Ramage, D., & Morris, M. R. (2011). #TwitterSearch: A Comparison of Microblog Search and Web Search. In *Proceedings of the Fourth ACM International Conference on Web Search and Data Mining (WSDM 2011)*. (to appear).

Terdiman, D. (2007, March 10). To Twitter or Dodgeball at SXSW. *CNET*. Retrieved from http://news.cnet.com/8301-17939_109-9696264-2.html

Tumasjan, A., Sprenger, T. O., Sandner, P. G., & Welpe, I. M. (2010). Predicting Elections with Twitter: What 140 Characters Reveal about Political Sentiment. Text.Proceedings, . Retrieved August 12, 2010, from http://www.aaai.org/ocs/index.php/ICWSM/ICWSM10/paper/viewPaper/1441

Tunkelang, D. (2009, January 13). A Twitter Analog to PageRank. *The Noisy Channel*. Retrieved September 22, 2010, from http://thenoisychannel.com/2009/01/13/a-twitter-analog-to-pagerank/

Twitter. (2010, June 18). Big Goals, Big Game, Big Records. *Twitter blog*. Retrieved August 24, 2010, from http://blog.twitter.com/2010/06/big-goals-big-game-big-records.html

Vieweg, S., Hughes, A. L., Starbird, K., & Palen, L. (2010). Microblogging during two natural hazards events: what twitter may contribute to situational awareness. In *Proceedings of the 28th international conference on Human factors in computing systems* (pp. 1079-1088). Atlanta, Georgia, USA: ACM. doi:10.1145/1753326.1753486

Voorhees, E. M. (2007). TREC: Continuing information retrieval's tradition of experimentation. *Commun. ACM*, *50*, 51–54. doi:http://doi.acm.org/10.1145/1297797.1297822

Weng, J., Lim, E., Jiang, J., & He, Q. (2010). TwitterRank: finding topic-sensitive influential twitterers. In *Proceedings of the third ACM international conference on Web search and data mining* (pp. 261-270). New York, New York, USA: ACM. doi:10.1145/1718487.1718520

Wilson, T., Wiebe, J., & Hoffmann, P. (2005). Recognizing contextual polarity in phrase-level sentiment analysis. In *Proceedings of the conference on Human Language Technology*

and Empirical Methods in Natural Language Processing (pp. 347-354). Vancouver, British Columbia, Canada: Association for Computational Linguistics. Retrieved from http://portal.acm.org/citation.cfm?id=1220619

Zhai, C., & Lafferty, J. (2004). A Study of Smoothing Methods for Language Models Applied to Information Retrieval. *ACM Transactions on Information Systems*, *2*(2), 179-214.

Zhang, J., Ackerman, M. S., Adamic, L., & Nam, K. K. (2007). QuME: a mechanism to support expertise finding in online help-seeking communities. In *Proceedings of the 20th annual ACM symposium on User interface software and technology* (pp. 111-114). Newport, Rhode Island, USA: ACM. doi:10.1145/1294211.1294230

Zhao, D., & Rosson, M. B. (2009). How and why people Twitter: the role that micro-blogging plays in informal communication at work. In *Proceedings of the ACM 2009 international conference on Supporting group work* (pp. 243-252). Sanibel Island, Florida, USA: ACM. doi:10.1145/1531674.1531710