

# Diversity and Relevance in Social Search

Omar Alonso<sup>†</sup>, Michael Gamon<sup>‡</sup>, Kevin Haas<sup>†</sup>, Patrick Pantel<sup>‡</sup>

<sup>†</sup>Microsoft Corp.

1075 La Avenida

Mountain View, CA 94043, USA

{omalonso,kevinhaa}@microsoft.com

<sup>‡</sup>Microsoft Research

One Microsoft Way

Redmond, WA 98052, USA

{mgamon,ppantel}@microsoft.com

## ABSTRACT

We propose the use of social features in studying diversity. The surge of social network engagements has led to growing insights on how an individual's connections view the world via opinions, likes, status updates, and their own connections. In information seeking scenarios that require sensemaking, these social signals can be leveraged to increase the diversity of search results and summarize one's network view on the query topic by pivoting on various social cues such as friend circle, expertise, and affinity. This position paper makes the case for incorporating social cues as a diversification tool and through user scenarios discusses the interplay between social cues and the sensemaking process.

## Categories and Subject Descriptors

H.3.3 [Information Search and Retrieval]: Search process

## General Terms

Design, Human Factors.

## Keywords

Social search, social relevance, annotations, sensemaking.

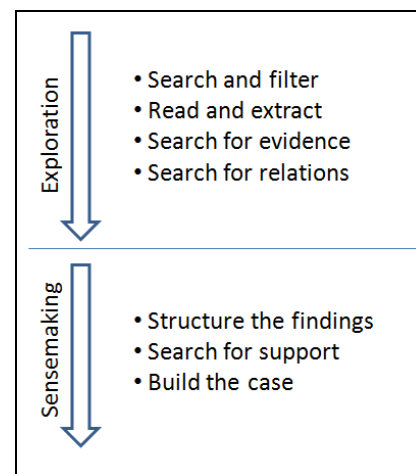
## 1. INTRODUCTION

The pervasive use of social networks has caused commercial search engines to rapidly integrate social content within their search results. Two notable examples are Bing, which includes Facebook and Twitter feeds, and Google with their recently launched Google+ network.

Social networks provide key signals for diversifying search results. On one hand, network preferences can be highlighted within the search results page (SERP) by re-ranking algorithmic results and by consolidating the network view in a social answer (direct display). On the other hand, within the social answer itself, diverse views can be represented through pivots on social cues such as a connection's expertise, friend circle, and affinity.

In the last few years there has been a lot of research trying to get a better understanding of how people search for information in different scenarios, see Hearst's chapter in [1] for a detailed

summary. Information seeking can be seen as being part of a larger process called *sensemaking* [3]: a task-centric holistic view of information larger than that normally associated with information retrieval. For example, when shopping for an SLR camera, the sensemaking process includes several information gathering tasks such as determining the most appropriate model, price range, and vendor. Each user involved in an activity that includes some sort of analysis is performing sensemaking. The goal is not just to retrieve information but to make sense of it. Evans and Chi [2] present the results of a large survey in the context of sensemaking and information seeking models. Over half of the search experiences reported were informational needs in nature (as opposed to transactional or navigational) and involve sensemaking. Search and presentation of important relevance cues are a core part of this process. We argue that social signals can be an important facilitator for the sensemaking process. Figure 1 shows a simplified version of sensemaking.



**Figure 1. From information retrieval and exploration to sensemaking.**

*Social search* is a general term used to describe searches that utilize social networks or involve a collective intelligence process to help the user satisfy an information need.

Many relevance cues are encoded on the social graph, examples include social connections such as followers and followees as well as past conversations. We also find many instances where the graph does not capture the wide range of *social cues*, for example when it comes to more fine-grained information about areas of expertise or particular tastes of members of a friend network. By social cues we denote those social interactions (in the physical and virtual world) that the user has captured over time like social

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

DDR '12, Feb 12, 2011 Seattle WA, USA.

Copyright 2012 ACM 1-58113-000-0/00/0010...\$10.00.

circle, affinity, expertise, location, etc. Depending on the user's need, leveraging those connections is important for determining potential relevant content.

## 2. VEHICLES FOR SOCIAL DIVERSITY: MOTIVATING SCENARIOS

Most of the motivating examples that are commonly used to showcase diversity such as {jaguar} or {apple} are somewhat limited and probably applicable to a web search context. While these are valid cases for providing diverse results due to the multiple interpretations, little has been done to include the most prominent source of user engagement on the Web nowadays: social behavior.

We envision two basic ways to leverage social cues for diversity. First, social cues can be used directly within the SERP as a mechanism for re-ranking. Second, social cues can be aggregated in a social answer. If enough information is available within the social answer, diversity comes into play in that answer.

To delineate the class of queries that can benefit from social cues, we introduce the notion of a *social intent*. A social intent is an intent that is *post-informational*. That is, the user has already performed searches, filtered certain results and read some of them. However, the user has not reached a decision stage yet. We now describe three scenarios involving a social intent to illustrate different ways in which a social signal can increase diversity and facilitate the sensemaking process.

In these scenarios we assume that the user has a presence in a social network and his/her social graph includes friends, colleagues, and family members with different degrees of affinity that have liked (or tagged) web pages or other properties.

**Diversity within the SERP (re-ranking):** In this scenario let us assume the user is looking to purchase expensive electronic equipment such as a digital camera or plasma TV. After some exploratory search to test the waters, he/she settles for a brand and range of models like {nikon d700}. Because of the high price, the user wants to make sure that he/she is making a good purchase decision.

Most of the reviews are positive and customers are happy. However, this is also the case for other brands and models which makes it difficult to reach a purchase decision. Having a friend who liked a similar model may influence the situation. If this signal is used in re-ranking of results to promote the particular model(s) that a social cue is available for, the user's sensemaking process can be accelerated.

Furthermore, a friend's ultimate purchasing decision after researching the same brand and range of products may provide an unexpected new insight – possibly there is even another product out there that the user has overlooked but that a friend's research will bring to attention – in this case by potentially up-ranking a search result that otherwise would have been of much lower relevance.

**Diversity within a social answer:** Assume that the user needs to have cataract procedure. After consultation with a doctor, he/she issues a number of queries about {eye cataract surgery} to become more familiar with the details, options and recovery. The information available in the search results is very topical to the

medical domain but lacks a social connection. In the particular case of medicine, there are many web pages with footnotes noting that the information is not intended to replace the advice of a doctor and disclaim liabilities.

A social answer in the form of an annotation that shows cues from a social network can be very beneficial. For example, if a friend who has had the same procedure recommends a clinic or comments about his/her successful recovery, this information would be extremely relevant to the user. Multiple such comments and social cues – maybe from a larger and specialized social community such as a health-related patient community – could be used to create a diverse social answer with rich and faceted information.

Similarly to the diversity within the SERP scenario, this can again lead to the discovery of novel and unexpected information, for example through the mentioning of dietary changes that helped a friend during recovery, even if this was not the central point of the initial queries.

**Diversity within a social answer (pivoting on social cues):** Say that the user is planning a long vacation and he/she is looking for {pet sitters} for their cat. Assuming that there are a number of providers and sitters locally, friends' comments or information on availability and overall experience with this specific type of service can be useful for making a choice. Within a social answer the user may be able to pivot on various aspects of the social connection. For example, recommendations of close friends or family members may bear more weight than those of work colleagues. Similarly, cat owners in the friend circle are probably providing stronger signals than dog or reptile owners.

Once again, serendipity may come into play – maybe the user could discover that some pet-owning friends trade pet sitting services amongst themselves, without engaging a professional service.

## 3. FACETS OF SOCIAL CUES

While it is beyond the goal of this paper to produce a complete taxonomy of all the possible facets of social cues, we would like to point out at least a few prominent ones. Note also that with the emergence of new social networks the set of facets will remain fluent depending on the level of detail that is encoded in a network.

A first obvious facet is *affinity*. Close friends will tend to be more important for social cues than more distant friends, and family is likely to be important as well, but possibly the mechanism here could be more complicated (there may be tendencies to differentiate oneself from some members of one's family).

*Expertise* is likely to be an important facet as well, which in some cases may trump affinity. If a distant friend is a known expert on photography, their opinion may count more than that of a very close friend who only occasionally takes some blurry snapshots.

Another interesting facet (and probably hard to capture) is *taste*. Some friends may have very different tastes in movies, to the point where there is a clear negative correlation between their taste and a user's likelihood to heed their recommendations. In this case similar tastes will influence similar behavior while dissimilar tastes will influence the user to avoid movies that are recommended by certain friends.

Other examples of facets include *geography* (if a friend lives in Seattle, I am much more likely to trust their knowledge on attractions there), *age* (teenagers' recommendations versus recommendations from adults), *gender* (shopping recommendations from female friends will differ from those of male friends), etc.

All these facets can allow a very rich social answer experience, but their interplay and usefulness for a given query needs to be carefully studied in order to provide the right measure of diversity and to facilitate the sensemaking process without overloading the user with useless information.

We outline a number of problems and challenges. This is not an exhaustive list.

1. Relevance assessment: how to assess content and how to collect training data
2. Experiments: how to conduct experiments when the content is highly personalized
3. Metrics: what are the metrics that capture social relevance and diversity that we can use
4. Ranking: Re-ranking search results including a social display and ranking and presentation of social cues within the answer/social display.

## 4. CONCLUSIONS

In this position paper we have presented social search as a task that can benefit from a sensemaking perspective by providing social answers to help users reach a decision when trying to satisfy a need. We have presented a number of examples and scenarios where social answers can be very helpful. We also describe the notion of diversity in the context of situations where social answers are available.

Finally, we outline a number of problems and challenges in utilizing social cues for diversity in the SERP and for diversity within the answer.

## 5. REFERENCES

- [1] R. Baeza-Yates and B. Ribeiro-Neto. *Modern Information Retrieval*, 2<sup>nd</sup> Edition, 2011.
- [2] B. Evans and E. Chi. "An Elaborated Model of Social Search", *Information Processing and Management*, 46(6):656-678. 2010.
- [3] D. Russell and M. Stefik. "Beyond information retrieval: Sensemaking:" Xerox PARC ISTL Technical Report, 1992.