

# The “Nays” Have It: Exploring Effects of Sentiment in Collaborative Knowledge Sharing

Ablimit Aji, Eugene Agichtein

Mathematics & Computer Science Department

Emory University

{aaji, eugene}@mathcs.emory.edu

## Abstract

In this paper we study what effects sentiment have on the temporal dynamics of user interaction and content generation in a knowledge sharing setting. We try to identify how sentiment influences interaction dynamics in terms of answer arrival, user ratings arrival, community agreement and content popularity. Our study suggests that “Negativity Bias” triggers more community attention and consequently more content contribution. Our findings provide insight into how users interact in online knowledge sharing communities, and helpful for improving existing systems.

## 1 Introduction

Recently, Collaborative Knowledge Sharing sites (or CQA sites), such as Naver and Yahoo! Answers have exploded in popularity. Already, for many information needs, these sites are becoming valuable alternatives to search engines. Previous studies identified visibility as an important factor for content popularity and developed models in static settings. However, when users post social media content, they might either explicitly or implicitly express their personal attitudes or sentiment. The following example illustrates a question with negative sentiment.

*Q*: Obama keeps saying we need to sacrifice. What sacrifices has he and the gov made collectively and individually?  
*A*<sub>1</sub>: Our hard earned tax dollars. 17 ↑, 2 ↓  
*A*<sub>2</sub>: None and they never will. 18 ↑, 2 ↓

Psychological studies (Smith et al., 2008) suggest that our brain has “Negativity Bias” - that is, people automatically devote more attention to negative information than to positive information. Thus, our attitudes may be more heavily influenced by negative opinions. Our hypothesis is that this kind of human cognitive bias would have measurable effects on how users respond to information need in CQA

communities. Our goal in this paper is to understand how question sentiment influence the *dynamics* of the user interactions in CQA - that is, to understand how users respond to questions of different sentiment, how question sentiment affects community agreement on best answer and question popularity.

## 2 Sentiment Influence

While (Aji et al., 2010) suggests that question category has a patent influence on interaction dynamics, we mainly focus on sentiment in this exploratory study, for the reason that sentiment is a high level but prominent facet in every piece of content. We focused on how may sentiment effect the following dimensions:

- **Answer Arrival:** Measured as number of answers arrived every minute.
- **Vote Arrival:** Measured as number of votes arrived per answer.
- **Community Agreement:** Mean Reciprocal Rank (MRR), computed by ranking the answers in order of decreasing “Thumbs up” ratings, and identifying the rank of the actual “best” answer, as selected by the asker.

$$MRR = \frac{1}{|Q|} \sum_{i=1}^N \frac{1}{rank_i} \quad (1)$$

where  $rank_i$  is the rank of the best answer among the answers submitted for question  $i$ .

- **Answer Length, Question Length:** We examine whether questions with different sentiment exhibit variations in question and answer length.
- **Interest “Stars”:** How many users marked question as interesting.

## 3 Dataset Description

For our study we tracked a total of approximately 10,000 questions, sampled from 20 categories from Yahoo! Answers. Specifically, each new question in our tracking list crawled every five minutes until it’s closed. As a result, we obtained approximately 22

million question-answer-feedback snapshots in total. Since labeling all the questions would be expensive, we randomly selected 2000 questions from this dataset for human labeling. We then utilized the Amazon Mechanical Turk Service<sup>1</sup>. Five workers labeled each question as either *positive*, *negative* or *neutral*; the ratings were filtered by using majority opinion (at least 3 out of 5 labels). Overall statistics of this dataset are reported in Table 1. The overall inter-rater agreement was 65%.

Positive	Negative	Neutral	Total
379	173	548	1,100

Table 1: Statistics of the *Temporal* dataset

## 4 Results and Discussion

Figure 1 reports answer arrival dynamics for question with varying sentiment. Answers to negative questions arrive substantially faster than answers to positive or neutral questions, whereas the difference between positive and neutral questions are minor. This strongly confirms the “Negative Bias” effect. Given the fact that questions stay in the category front page relatively same amount of time where their visibility contributes potential answers, on average, negative sentiment questions managed to get more answers than two other types of questions (4.3 vs. 3.3 and 3.5). It seems, sentiment expressed in a question contributes to the answer arrival more than visibility.

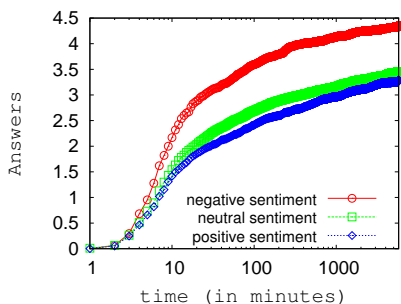


Figure 1: Cumulative answer arrival

Figure 2 reports rating arrival dynamics. Interestingly, positive *ratings* arrive much faster to negative questions, whereas positive and negative ratings arrive roughly at the same rate for positive and neutral questions. While this might be partially due to the fact that negative sentiment questions are more “attention grabbing” than other types of questions, we conjecture that this effect is caused by the selection bias of the raters participating in negative question threads, who tend to support answers that strongly

<sup>1</sup><http://www.mturk.com>

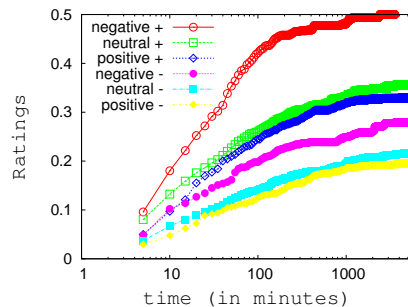


Figure 2: Cumulative user ratings arrival

agree (or strongly disagree) with the question asker. Surprisingly, community agreement(MRR) on the

Type	MRR	QLength	ALength	Stars
Negative	0.47	<b>78</b>	49	<b>0.25</b>
Positive	0.56	58	<b>52</b>	0.16
Neutral	<b>0.57</b>	52	47	0.15

Table 2: Agreement, Question length, Answer Length and Star count averaged over question type

best answer is lower for negative sentiment questions. On average, negative sentiment questions were marked as interesting more than positive or neutral questions were marked as interesting. Although this may sound counterintuitive, it is not surprising if we recall how the “Negative Bias” influences user behavior and may increase implicit “visibility”. All the above mentioned differences are statistically significant(t-test  $p = 0.05$ ).

In summary, our preliminary exploration indicates that sentiment may have a powerful effect on the content contribution dynamics in collaborative question answering, and is a promising direction for further study of knowledge sharing communities.

## Acknowledgments

We thank HP Social Computing Labs for support.

## References

- Ablimit Aji. Eugene Agichtein. 2010. *Deconstructing Interaction Dynamics in Knowledge Sharing Communities*. International Conference on Social Computing, Behavioral Modeling, & Prediction.
- Gabor Szabo. Bernardo Huberman. 2008. *Predicting the popularity of online content*. HP Labs Technical Report.
- Kristina Lerman. 2007. *Social Information Processing in Social News Aggregation*. IEEE Internet Computing: Special Issue on Social Search.
- N. Kyle Smith Jeff T. Larsen Tanya L. Chartrand John T. Cacioppo 2006. *Affective Context Moderates the Attention Bias Toward Negative Information*. Journal of Personality and Social Psychology.