

Information Quality on Yahoo! Answers

Pnina Fichman

Indiana University, Bloomington, United States

ABSTRACT

Along with the proliferation of the social web, question and answer (QA) sites attract millions of user around the globe. On these sites users ask question while others provide answers. These QA site vary by their scope, size, and quality of answers; the most popular QA site is Yahoo! Answers. This chapter aims to examine the quality of information produced by the crowd on Yahoo! Answers, assuming that given enough eyeballs all questions can get good answers. Findings illustrate a process of answer quality improvement through crowdsourcing questions. Improvement is achieved by having multiple answers to any given question instead of a single answer, and through a mechanism of answer evaluation, by which users rank the best answer to any given question. Both processes contribute significantly to the quality of answers one can expect to find on Yahoo! Answers.

INTRODUCTION

Online question and answer (QA) sites are “websites, where members can post questions, answer other members' questions, and rate other members' answers to their questions--all for free.” (Rosenbaum & Shachaf, 2010, p. 1933). Visitors to these sites are looking for answers; these questions and answers are archived for future use. Under the assumption that everyone knows something (Roush, 2006), people answer the questions. On some QA sites, participants can rank the quality of answers and users build up their reputations on the site based on their past contributions to the site by providing good answers and exhibiting good behavior (Rosenbaum & Shachaf, 2010)

Research on QA sites focused attention on information retrieval (e.g., Bian, Liu, Agichtein, & Zha, 2008) and information seeking behavior (e.g., Gazan, 2007), or described as a socio-technical environment (e.g., Gazan, 2010; Rosenbaum & Shachaf, 2010; Shachaf & Rosenbaum, 2009). Q&A sites challenge information retrieval researchers, to incorporate a social dimension into the retrieval mechanism, reference scholars, to understand the nature of crowd sourced online question answering. However both camps have high stakes in identifying good answers. Thus, distinguishing between high and poor quality answers attracted much attention from the very beginning (e.g., Adamic et al., 2008; Agichtein et al., 2008; Bian et al., 2008).

Yahoo! Answers is among the most frequently consulted reference sites, second only to Wikipedia (Fichman, 2012). Its popularity and the fact that users return to the site frequently may indicate that answer quality is good. However, scholars raise questions about the reliability of user-created content on content repository sites, such as Yahoo! Answers, You Tube, and Wikipedia, while at the same time, many studies show that the crowd produces content that does not fall in quality from traditional publications (Giles, 2005). Prior research also indicates that Q&A sites, such as Yahoo! Answers and the Wikipedia Reference Desk, provide answers at a quality that is as good and even better than answers that librarian provide (Harper et al, 2008; Shachaf, 2009). Yet, even if the crowd provides good answers, it is unclear how this process of

quality improvement yields good answers and what conditions hinder or support high quality. It is also unclear how many amateurs does it take to provide one good answer, or in other words, how many answers per question are ‘enough’ to produce a good answer on Q&A sites.

This chapter tries to uncover the process by which crowdsourcing question answering can improve information quality and examine changes in answer quality as the number of answers increases. The chapter unpacks the black box of crowd-produced content, by conducting a content analysis of hundreds of questions/answers pairs from Yahoo! Answers, to assess answer reliability, as measured through accuracy, completeness, and verifiability.

BACKGROUND

There is a growing body of research on QA sites (e.g., Chua & Balkunje, 2012; Fichman, 2012; Rehavi & Refaeli, 2012; Wu & Korfiatis, 2013), showing that information quality can vary significantly between questions, topics, and sites. Yet, these scholars agree that the participatory nature of Web 2.0 provides an infrastructure for achieving high quality knowledge production.

In a recent papers, Bloom, Goh, and Chua (2012) identified predictors of answer quality under two categories: social (user interaction and feedback) and content features (intrinsic and extrinsic content quality). They argue that the features identified as strongly associated with high-quality answers include positive votes, completeness, presentation, reliability and accuracy, and that features weakly associated with high-quality answers were high frequency words, answer length, and best answers answered. Other scholars agree with Bloom, Goh, and Chau (2012) that content features are critical in assessing answer quality (Bloom, Chua, & Goh, 2008; Fichman, 2011; Harper, Raban, Rafaeli, & Konstan, 2008; Shachaf, 2009); yet, studies also found that better answers are longer (Adamic, Zhang, Bahshy, & Ackerman, 2008; Bloom, Chua, & Goh, 2008; Harper, et al., 2008), and include references to external sources (Gazan, 2006).

Despite Bloom, Goh and Chau (2012) argument that best answers (ranked by users on Yahoo! Answers) do not correlate with answer quality, user rankings of “best answers” was heavily used in prior research to represent good answers (e.g., Adamic, Zhang, Bakshy, & Ackerman, 2008). Fichman (2012) suggests “user rankings are problematic because they provide a subjective measure of answer quality.” In line with this approach Kim and Oh (2009) found that in 29.8% of cases where users chose “best answers” in Yahoo! Answers, their selections were based on socio-emotional criteria rather than on the content or utility of the answer.

Interestingly, based on Bloom, Goh and Chau’s (2012) analysis, features related to the asker’s user history were found not to be associated with high-quality answers. However, based on the assumption that some users are more likely to provide better answers than others (Bouguessa, Dumoulin, & Wang, 2009), tracking user reputation was a common method to assess answer quality. This includes the ranking of authoritative responders using the number of “best answers” they have previously posted (Bloom, Chua, & Goh, 2008) or assessing user’s activity levels (e.g., Bloom, Chua, & Goh, 2008; Bouguessa, Dumoulin, & Wang, 2009). Fichman (2012, p. 263) argues, “even users with good reputations do not always provide high quality answers.”

In an effort to understand why the crowd produce good information, some suggest that the same principle that holds in Free and Open Source Software (FOSS) and known as Linus’ Law (Meneely & Williams, 2009; Schweik, et al., 2008), “given enough eyeballs, all bugs are shallow” is applicable also in open content repositories, such as Wikipedia. Thus, specifically for mass-content production sites, Arnison’s Law postulates, “given enough eyeballs, problematic content is shallow.” In the context of Q&A sites, and Yahoo! Answers in particular, one can claim that given enough eyeballs, all questions can get good answers. But, how much is ‘enough’?

Theoretically the claim that enough volunteers can produce good content cannot be rejected, since any identification of missing or inaccurate information over time only proves the claim by showing that additional review of the content leads to better content. Still, in reality this process of content improvement can take an endless number of users and a very long time before

‘enough’ users reviewed and modified to content to be good enough. Scholars, on the other hand, found empirical support for Linus’ Law (Meneely & Williams, 2009; Schweik et al. 2008). This chapter examines empirically a variation on Linus’ and Arnison’s Laws which is the claim that given enough eyeballs, all questions can get good answers by asking:

- 1) Does answer quality on Yahoo! Answers improve as the number of answers increases?
And if so;
- 2) How many answers are enough?

To answer these questions, a sample of 3,568 posts, which included 585 questions and their answers, was collected from Yahoo! Answers. Figure 1 shows the distribution of answers per question. Only a few questions received very high number of answers and many of the questions received either very few or no answers at all.

Insert figure 1 approximately here

Figure 1. Number of answers per question

A purposeful sample of 74 resolved informational transactions was drawn from this data set for further content analysis. Informational transactions are questions that “are asked with the intent of getting information that the asker hopes to learn or use via fact- or advice oriented answers” (Harper, Moy, & Konstan, 2009, p.759). A transaction includes a question and any number of answers (7.81 answer per question on average). Resolved transactions are question that include an answer that users on the site marked as “best answer”. Best answers are chosen by the user who asked the question or by a vote of other users.

Content analysis of 382 posts (74 questions and 308 answers) from Yahoo! Answers, was conducted to determine answer quality. Quality rates were marked for three types of answers in each transaction: 1) the whole answer (all the answers that were posted for a give question); 2) the first answer (which was posted in response to a given question); and 3) the “best answer” (which was chosen by the user who asked the question). Following Fichman (2011), three reliability measures were utilized information quality: 1) accuracy of an answer refers to a correct response; 2) completeness of an answer refers to an answer that is thorough, provides enough information, and answers all parts of a multi-part question; and 3) verifiability of an answer refers to an answer that provides a link or a reference to another source where the information can be found. These reliability measures are of particular importance in judging the credibility of answers to informational questions (Blooma, Chua, & Goh, 2008, 2012; Kim, 2010).

Two coders, who were graduate students, coded the entire data set, for levels of accuracy, completeness, and verifiability. Coders were instructed to determine the quality of the answers “on the surface” (Lombard, Snyder-Duch, & Bracken, 2002) on each of the three measures. Inter-coder reliability was determined using simple agreement and was 92%.

DOES ANSWER MULTIPLICATION IMPROVE ANSWER QUALITY? AND IF SO, HOW MANY ANSWERS ARE ENOUGH?

The content analysis of 382 posts reveals that answer multiplication improves answer quality and that seven answers are enough. Table 1 summarizes the results of the content analysis for all three reliability measures (accuracy, completeness, and verifiability) on three types of answers (best answer, first answer, and whole answer) per transaction. Table 2 shows the results of cross tabulation, comparing rates on the three measures between the three types of answers as well as follow-up analysis on two variables, accuracy and completeness.

Insert Table 1 approximately here

Table 1. Rates on reliability measures

The results show that “best answer” and whole answer are significantly better than first answer on two of the three measures, accuracy and completeness. The quality of “best answers” is equal to that of whole answers; no significant differences were found. It is possible that fewer answers than the average are enough to produce a significantly better answer than the first answer without the need for evaluation or unlimited number of answers. To examine this possibility further investigation into the location of the ‘best answer’ in the transaction was conducted.

Insert Table 2 approximately here

Table 2. Cross-tabulation results for differences in rates of reliability measures between “Best Answer”, First Answer, and Whole Answer

The cumulative percentage and frequency of “best answer” location are reported in Table 3; Figure 2 illustrates the distribution of “best answer” location in the sample. In most of the transactions the “best answer” was one of the first three answers (59.45%), and in many cases the “best answer” was the first answer (28%).

Insert Figure 2 approximately here

Figure 2. Location of “Best Answer”

Eighty percent of the resolved transactions have a “best answer” chosen by the seventh answer (Table 3). That indicates that to achieve accuracy rate of 95% it is enough to have seven answers in 80% of the transactions; eighty percent was marked based on Pareto principal for optimization.

Insert Table 3 approximately here

Table 3. “Best answer” location in transaction

Further analysis shows that many high quality answers are posted before the “best answer” (Table 4). The number of accurate and correct answers that are posted before the “best answer” correlates with the “best answer” location ($r=.86$), with the total number of answers ($r=.90$), and with total number of answers ($r=.87$); there is a strong Pearson product-moment correlation coefficient between the number of answers and the location of the “best answer” in the transactions ($r=.93$).

Insert Table 4 approximately here

Table 4. Accurate and complete answers per transaction (resolved sample)

Next, the number and percentage of accurate answers and complete answers per transaction were marked, based on analysis of all question-answer pairs (Table 5). Findings show that only two answers were needed to achieve high accuracy in more than 80% of the transactions (Table 5).

Insert Table 5 approximately here

Table 5. Location of first accurate answer

Figure 3 illustrates the percentage distribution of accurate and complete answers. It is interesting to note that in 30% of the transactions all the answers are accurate and complete (i.e., 100% of the answers in the transaction are accurate and complete). Another 30% of the transactions have 50% or lower rates of accurate and complete answers per transaction; yet, half of the transactions have 70% or more accurate and complete answers per transaction.

Insert figure 3 approximately here

Figure 3. Percent of accurate and complete answers per transaction.

Discussion

The chapter aimed to answer two questions:

- 1) Does answer quality on Yahoo! Answers improve as the number of answers increases?
And if so;
- 2) How many answers are enough?

The findings indicate that:

- 1) Answer quality improved with additional answers. In other words, answer multiplication, with or without the “best answer” feature, results in higher answer quality than the first answer.
- 2) Seven answers are enough to get good answers. In other words, to achieve 95% accuracy rate 80% of the time, seven answers are enough.

Two processes contribute to quality improvement:

- 1) Answer multiplication, which was measured by whole answer.
- 2) Answer evaluation, which was measured by best answer.

The whole answer includes on average seven answers -- it can be argued then, that with seven answers on average the crowd produces an answer that is significantly better than the first answer. Thus, if answer multiplication includes seven users, answer accuracy and completeness will be significantly better than an answer provided by a single user. The findings of the present study show that additional answers only improved answer quality. Thus it can be concluded that the findings are in alignment with prior FOSS research that found evidence in support of Linus’s Law (Schweik, English, Kisting, & Haire, 2008). It is also supportive of the claim that given enough eyeballs all questions can get a good answer.

It is important to note that significant improvement in answer quality, in terms of accuracy and completeness, was also associated with “best answers.” While “best answers” are individual answers, their quality was equal to that of whole answers. This may be due to the selection process of “best answers”, which involves an additional step of information processing that includes feedback about the quality of the answer in light of the question. In support of the argument that given enough eyeballs all questions are can get a good answer, the findings reveal that it takes seven answers to achieve a 95% accuracy level.

FUTURE RESEARCH DIRECTIONS

Future research may be done to overcome the limitations inherent to the study described in this chapter.

One of the limitations of this study is that the analysis was limited to transactions from one Q&A site, Yahoo! Answers. This was done in effort to control for site variations in user demographics, policies, and technological infrastructure (Fichman, 2011; Rosenbaum & Shachaf, 2010). Still, Yahoo! Answers is the most popular QA site, where the vast majority of collaborative question answering is taking place. But, because the study uses data only from one QA site, generalizations should be made with caution (Fichman, 2011). Future research should look into these questions in Q&A sites that are different or less popular. It would be interesting to know if answer multiplication always improves answer quality, if enough answer multiplication depends on the size of the community, or if the focus of the community leads to lower number of answers required. Shachaf (2009), for example, report an average of 4.5 answers on the Wikipedia reference desk, which led to high answer quality.

Another limitation is the choice of informational questions, which count for 73% of all questions asked in this sample; success rates and evaluation criteria for informational questions may differ from conversational questions. This type of question was chosen to facilitate the use of evaluation criteria across all the transactions. Nevertheless, these transactions and their success rates characterize the vast majority of transactions on Yahoo! Answers. However, future research can examine the number of answers that lead to success in conversational questions -- does more

answers/ opinions means always better? Do all Q&A sites encourage both types of questions, and if not, how does the range of question type affect answer multiplication and answer quality?

Finally, one of the reliability measures, verifiability, did not improve with answer multiplication, and the levels reported are very low. This suggests that future research should investigate why verifiability levels are so low on Yahoo! Answers. It would also be beneficial to examine how do various reliability measure relates to each other. What is the weight that each should be given respectively – does the low level indicate low quality or lack of attention to and interest in this dimension of answer quality.

CONCLUSION

The chapter focused attention on information quality on Yahoo! Answers using three measures of answer reliability (accuracy, completeness, and verifiability), on three types of answers (first answer, best answer, and whole answer). By analyzing a sample of 382 posts, the study found that answer quality improved with additional answers, and that there was no evidence of a number of answers, after which additional answers reduce quality. The findings also indicate that seven answers are enough to yield good answers for over 80% the questions. This supports the argument that given enough eyeballs all questions can get a good answer.

REFERENCES

- Adamic, L.A., Zhang, J., Bakshy, E., & Ackerman, M.S. (2008). Knowledge sharing and Yahoo! Answers: Everyone knows something. In *Proceedings of the International World Wide Web Conference* (pp. 665-674). Beijing: ACM.
- Agichtein, E., Castillo, C., Donato, D., Gionides, A., & Mishne, G. (2008). Finding high-quality content in social media. In *Proceedings of the International Conference on Web Search and Web Data Mining*, (pp. 183-194). Palo Alto: ACM.
- Bian, J., Liu, Y., Agichtein, E., & Zha, H. (2008). Finding the right facts in the crowd: factoid question answering over social media. In *Proceedings of the International World Wide Web Conference* (pp. 467-476). Beijing: ACM.
- Blooma, M.J., Chua, A.Y., & Goh, D.H. (2008). A predictive framework for retrieving the best answer. In *Proceedings of the 2008 ACM Symposium on Applied Computing* (pp. 1107-1111). Fortaleza, Ceara, Brazil: ACM.
- Blooma, M.J. Goh, D.H., & Chua, A.Y. (2012). Predictors of high-quality answers. *Online Information Review*, 36(3), 383 – 400.
- Bouguessa, M., Dumoulin, B., & Wang, S. (2009). Identifying authoritative actors in question-answering forums: The case of Yahoo! Answers. In *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 866-874). Las Vegas: ACM.
- Chua, A.Y.K., Balkunje, R.S. (2012). Comparative evaluation of community question answering websites. *The outreach of digital libraries: A globalized resource network. Lecture Notes in Computer Science*, 7634 (pp. 209-218).
- Fichman, P. (2011). A comparative assessment of answer quality on four question answering sites. *Journal of Information Science*, 37(5), 476-486.
- Fichman, P. (2012). How many answers are enough? Optimal number of answers for Q&A sites. *Social Informatics*, 260-274.
- Gazan, R. (2006). Specialists and synthesists in a question answering community. In *Proceedings of the American Society for Information Science & Technology Annual Meeting* (pp. 1-10). Austin: ASIST.
- Giles, J. (2005). Internet encyclopedias go head to head. *Nature*, 438, 900-901 (available at <http://www.nature.com/news/2005/051212/full/438900a.html>).

Harper, F.M., Moy, D., & Konstan, J.A. (2009). Facts or friends?: Distinguishing informational and conversational questions in social Q&A sites. In *Conference on Human Factors in Computing Systems* (pp. 759-768). Boston: ACM.

Harper, F.M., Raban, D., Rafaeli, S., & Konstan, J.A. (2008). Predictors of answer quality in online Q&A sites. In *Proceedings of the Conference on Human Factors in Computing Systems* (pp. 865-874). Florence, Italy: ACM.

Kim, S., & Oh, S. (2009). Users' relevance criteria for evaluating answers in social Q&A site. *Journal of the American Society for Information Science and Technology*, 60(4), 716-727.

Lombard, M., Snyder-Duch, J., & Bracken, C. C. (2002). Content analysis in mass communication: Assessment and reporting of intercoder reliability. *Human Communication Research*, 28(4), 587-604.

Meneely, A., & Williams, L. (2009). Secure open source collaboration: An empirical study of Linus' Law. In *Proceedings of the 16th ACM Conference on Computer and Communications Security* (pp. 453-462). New York: ACM.

Rehavi, A., & Refaeli, S. (2012). Knowledge and social networks in Yahoo! Answers. *45th Hawaii International Conference on System Science (HICSS)*, pp. 781-789.

Rosenbaum, H., & Shachaf, P. (2010). A structuration approach to online communities of practice: The case of Q&A communities. *Journal of the American Society for Information Science and Technology*, 61(9), 1933-1944.

Schweik, C.M., English, R.C., Kisting, M., & Haire, S. (2008). Brooks' versus Linus' law: an empirical test of open source projects. In *Proceedings of the 2008 International Conference on Digital Government Research* (pp. 423-424). Montreal, Canada: Digital Government Society of North America, ACM.

Shachaf, P. (2009). The paradox of expertise: Is the Wikipedia Reference Desk as good as your library? *Journal of Documentation*, 65(6), 977-963.

Shachaf, P., & Rosenbaum, H. (2009). Online social reference: A research agenda through a STIN framework. *Proceedings of the iConference '09*, Chapel Hill, NC, US.

Wu, P.F., & Korfiatis, Ni. (2013). You scratch someone's back and we'll scratch yours: Collective reciprocity in social Q&A communities. *Journal of the American Society for Information Science and Technology*.

ADDITIONAL READING SECTION

1. Adamic, L.A., Zhang, J., Bakshy, E., & Ackerman, M.S. (2008). Knowledge sharing and Yahoo! Answers: Everyone knows something. In *Proceedings of the International World Wide Web Conference* (pp. 665-674). Beijing: ACM.
2. Blooma, M.J., Chua, A.Y., & Goh, D.H. (2008). A predictive framework for retrieving the best answer. In *Proceedings of the 2008 ACM Symposium on Applied Computing* (pp. 1107-1111). Fortaleza, Ceara, Brazil: ACM.
3. Blooma, M.J. Goh, D.H., & Chua, A.Y. (2012). Predictors of high-quality answers. *Online Information Review*, 36(3), 383 – 400.
4. Bouguessa, M., Dumoulin, B., & Wang, S. (2009). Identifying authoritative actors in question-answering forums: The case of Yahoo! Answers. In *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 866-874). Las Vegas: ACM.
5. Chen, W., Zeng, Q., & Wenyn, L. (2006) A user reputation model for a user-interactive question answering system. In *Proceedings of the Second International Conference on Semantics, Knowledge, and Grid* (pp. 40-45). Washington D.C.: IEEE Computer Society.
6. Chua, A.Y.K., Balkunje, R.S. (2012). Comparative evaluation of community question answering websites. *The outreach of digital libraries: A globalized resource network. Lecture Notes in Computer Science*, 7634 (pp. 209-218).

7. Dom, B., & Paranjpe, D. (2008) A Bayesian technique for estimating the credibility of question answerers. In *Proceedings of the Society for Industrial and Applied Mathematics (SIAM)* (pp. 399-409). Atlanta: SIAM, (available at http://www.siam.org/proceedings/datamining/2008/dm08_36_Dom.pdf).
8. Fichman, P. (2011). A comparative assessment of answer quality on four question answering sites. *Journal of Information Science*, 37(5), 476-486.
9. Fichman, P. (2012). How many answers are enough? Optimal number of answers for Q&A sites. *Social Informatics*, 260-274.
10. Gazan, R. (2006). Specialists and synthesists in a question answering community. In *Proceedings of the American Society for Information Science & Technology Annual Meeting* (pp. 1-10). Austin: ASIST.
11. Gazan, R. (2007). Seekers, sloths and social reference: Homework questions submitted to a question-answering community. *New Review of Hypermedia & Multimedia*, 13(2), 239-248.
12. Gazan, R. (2010). Microcollaborations in a social Q&A community. *Information Processing & Management*, 46(6), 693-702.
13. Harper, F.M., Moy, D., & Konstan, J.A. (2009). Facts or friends?: Distinguishing informational and conversational questions in social Q&A sites. In *Conference on Human Factors in Computing Systems* (pp. 759-768). Boston: ACM.
14. Harper, F.M., Raban, D., Rafaeli, S., & Konstan, J.A. (2008). Predictors of answer quality in online Q&A sites. In *Proceedings of the Conference on Human Factors in Computing Systems* (pp. 865-874). Florence, Italy: ACM.
15. Harper, F.M., Weinberg, J., Logie, J., & Konstan, J.A. (2010) Question types in social Q&A sites. *First Monday*, 15(7) (available at <http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/viewArticle/2913/2571>)
16. Jurczyk, P., & Agichtein, E. (2007a). Discovering authorities in question answer communities by using link analysis. In *Proceedings of the Sixteenth ACM Conference on Information and Knowledge Management* (pp. 919-922). New York: ACM.
17. Jurczyk, P. & Agichtein, E. (2007b). Hits on question answer portals: Exploration of link analysis for author ranking. In *Annual ACM Conference on Research and Development in Information Retrieval* (pp. 845-846). Amsterdam: ACM.
18. Kim, S., & Oh, S. (2009). Users' relevance criteria for evaluating answers in social Q&A site. *Journal of the American Society for Information Science and Technology*, 60(4), 716-727.
19. Kim, S. (2010). Questioners' credibility judgments of answers in a social question and answer site. *Information Research*, 15(2), paper 432 (available at <http://InformationR.net/ir/15-2/paper432.html>).
20. Li, B., Liu, Y., Ram, A., Garcia, E.V., & Agichtein, E. (2009). Exploring question subjectivity prediction in community QA. In *Proceedings of the 31st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 735-736). Singapore: ACM.
21. Liu, Y., Li, S., Cao, Y., Lin, C., Han, D., & Yu, Y. (2008). Understanding and summarizing answers in community-based question answering services. In *Proceedings of the 22nd International Conference on Computational Linguistics* (pp. 497-504). Manchester, UK.
22. Nam, K.K., Ackerman, M.S., & Adamic, L.A. (2009). Questions in, knowledge in?: A study of Naver's question answering community. In *Proceedings of the 27th International Conference on Human Factors in Computing Systems* (pp. 779-788). Boston: ACM.
23. Rehavi, A., & Refaeli, S. (2012). Knowledge and social networks in Yahoo! Answers. *45th Hawaii International Conference on System Science (HICSS)*, pp. 781-789.

24. Rosenbaum, H., & Shachaf, P. (2010). A structuration approach to online communities of practice: The case of Q&A communities. *Journal of the American Society for Information Science and Technology*, 61(9), 1933-1944.
25. Shachaf, P. (2009). The paradox of expertise: Is the Wikipedia Reference Desk as good as your library? *Journal of Documentation*, 65(6), 977-963.
26. Shachaf, P. (2009). The paradox of expertise: Is the Wikipedia Reference Desk as good as your library? *Journal of Documentation*, 65(6), 977-963.
27. Shachaf, P. (2010a). Answer reliability on Q&A sites. In *Proceedings of the Sixteenth Americas Conference on Information Systems* (paper 376). Lima, Peru: AMCIS.
28. Shachaf, P. (2010b). Social reference: A unifying theory. *Library & Information Science Research*, 32(1), 66-76.
29. Shachaf, P., & Rosenbaum, H. (2009). Online social reference: A research agenda through a STIN framework. *Proceedings of the iConference '09*, Chapel Hill, NC, US.
30. Suryanto, M.A., Sun, A., Lim E., & Chiang, R.H.L. (2009). Quality-aware collaborative question answering: methods and evaluation. In *Proceedings of the Second ACM International Conference on Web Search and Data Mining* (pp. 142-151). Barcelona, Spain: ACM.
31. Wu, P.F., & Korfiatis, Ni. (2013). You scratch someone's back and we'll scratch yours: Collective reciprocity in social Q&A communities. *Journal of the American Society for Information Science and Technology*.