Chapter 18

Information Quality on Yahoo! Answers

Pnina Fichman
Indiana University – Bloomington, USA

ABSTRACT

Along with the proliferation of the Social Web, Question and Answer (QA) sites attract millions of users around the globe. On these sites, users ask questions while others provide answers. These QA sites 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 answersto their questions; 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, and 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, YouTube, 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 librarians 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 paper, 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 Blooma, Goh, and Chau (2012) that content features are critical in assessing answer quality (Blooma, Chua, & Goh, 2008; Fichman, 2011; Harper, Raban, Rafaeli, & Konstan, 2008; Shachaf, 2009); yet, studies also found that better answers are longer (Adamic, Zhang, Bakshy, & Ackerman, 2008; Blooma, Chua, & Goh, 2008; Harper, et al., 2008), and include references to external sources (Gazan, 2006).

Despite Blooma, 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).