

Research Trends of the Credibility of Information in Social Q&A

지식검색커뮤니티 정보의 신뢰성에 관한 연구 동향 분석

Soojung Kim(김수정)*

ABSTRACT

Social Q&A sites such as Yahoo! Answers and Naver Knowledge-iN have become a viable method for information seeking and sharing on the Web. Considering their immense popularity and growing concerns about their validity as information sources, questions about the credibility of the information provided on social Q&As are timely. Therefore, this paper summarizes recent research on credibility related to the social Q&A context, identifies research gaps, and presents a research agenda for future research to advance this newly developing area.

초 록

야후 앤서(Yahoo! Answers)와 네이버 지식인과 같은 지식검색 커뮤니티는 인터넷에서 정보를 찾고 공유하는 중요한 수단으로 부상하였다. 그러나 지식검색 커뮤니티의 인기가 날로 높아지는 것과 비례하여 정보자원으로서의 유효성에 대한 우려 또한 커지고 있는 것이 주지의 사실이다. 이러한 맥락에서 본 논문은 지식검색 커뮤니티와 관련된 신뢰성 문제에 대한 선행 연구들을 정리하고 향후 연구 과제를 제시함으로써 지식검색 커뮤니티 신뢰성에 관한 연구를 활성화시키는데 도움이 되고자 한다.

Keywords: credibility, question answering community, knowledge community, literature review, research trends, research agenda
신뢰성, 지식검색 커뮤니티, 문헌 조사, 연구 동향, 연구 과제

* Assistant professor, Department of Library and Information Science,
Chonbuk National University(kimsoojung@jbnu.ac.kr)

■ 논문접수일자 : 2012년 5월 19일 ■ 최초심사일자 : 2012년 5월 27일 ■ 게재확정일자 : 2012년 6월 16일
■ 정보관리학회지, 29(2): 135-154, 2012. [<http://dx.doi.org/10.3743/KOSIM.2012.29.2.135>]

1. Introduction

The recent past has witnessed the enormous growth of social media sites such as Wikipedia, YouTube, Yahoo!Answers, and more. These sites offer users opportunities to be creators as well as consumers of information by allowing them to generate content in various formats. User-generated content, which is individually experienced and contributed information, has natural advantages. First, individuals are in many cases in the best position to provide information that requires personal experience, opinions, and views (Flanagin & Metzger, 2008). Second, the movement towards user-generated content on the Web has led to collective intelligence, which enables people to make informed decision making.

Social question and answering (social Q&A) sites such as Yahoo!Answers and Naver Knowledge-iN are a phenomenal example of the user-generated content and have emerged as an alternative tool to automatic Web searches. The lack of professional editors to monitor these sites, however, has raised concerns about their credibility and possible hazards that could be encountered by information seekers. Accordingly, in the mid-2000s, researchers began their investigation by searching for answers to as to how credible social Q&As are and how people perceive the credibility of the information provided by fellow users in that environment. Given that social Q&As have become a viable method for information sharing, and difficulties involved in the credibility evaluation of answers provided by the public are high, it is timely to discuss the issue of credibility in the context

of social Q&As. Additionally, in spite of extensive literature on credibility in general, few comprehensive reviews are to be found on credibility within the social Q&A context. Therefore, this paper aims to offer a review that explores the main trends relating to credibility on social Q&As in the literature from 2004 to 2012.

This paper will, first, provide an overview of the concept of credibility with an introduction to the framework that guides the review of literature. It will then review literature on the credibility of information in the social Q&A environment to identify research gaps. Finally, it will suggest future research directions to advance the area.

2. Background

2.1 Credibility

While credibility is a multidimensional construct encompassing believability, trust, reliability, accuracy, fairness, objectivity, and others (Self, 1996), many researchers define credibility as believability (Tseng & Fogg, 1999) having two primary dimensions, namely trustworthiness and expertise. Trustworthiness refers to a source's willingness to provide accurate information and expertise refers to a source's ability to provide accurate information (Hovland, Janis, & Kelley, 1953). Within information science, credibility is generally related to information quality. Taylor (1986) defined quality as "a user criterion which has to do with excellence or in some cases truthfulness

in labeling” and identified five components of quality: accuracy, comprehensiveness, currency, reliability, and validity. Although Taylor did not explicitly use the term ‘credibility’, the notion is embedded in his derivation of quality from reliability and validity (Rieh & Danielson, 2007).

Another related concept to credibility is cognitive authority. Credibility and cognitive authority are interrelated in that both have trustworthiness and expertise as two main components. Wilson (1983) coined the term ‘cognitive authority’ to explain the kind of authority that not only possess expertise and trustworthiness, but also influences thoughts that people would recognize as being proper. This notion helps people consider the criteria they should use when evaluating information sources. Finally, credibility has been viewed as one of relevance criteria in information science (Barry, 1994; Bateman, 1999). Previous research shows that there is a core set of criteria that consistently appears across contexts: topic, accuracy/quality, authority, completeness/depth, document type, and belief. These relevance criteria, the dimensions of credibility, and the components of quality significantly overlap without having the clear-cut lines that separate them.

When evaluating information credibility, there are two contrasting perspectives (Flanagin & Metzger, 2008). The first perspective views credibility as an inherent, objective property of information and seeks to measure the accuracy of information by accepted standards or experts in a particular domain. By nature, this objective perspective is content-oriented and suitable for evaluating scientific knowledge. The second

perspective, on the other hand, views credibility as a perceived quality on the users’ side. Considerable empirical studies have taken this user-oriented perspective to understand users’ perceptions and factors influencing their credibility evaluation.

The question of which perspective to choose is critical to understanding a credibility study because the selected perspective determines the definition of credibility and subsequently, research design, and the interpretation of findings.

In addition, a number of studies have suggested new ways to automatically identify credible information on social Q&A sites over recent years. These studies have formed a body of knowledge mainly in the computer science field with the purpose of automating a credibility evaluation process. Considering that users do not make aggressive efforts in evaluating the credibility of information online, developing automatic systems that could assist users in identifying credible information in social Q&A will continue to be an important topic in the realm of credibility research.

Therefore, this paper divides previous studies on credibility in social Q&As into three broad domains: 1) content-oriented approach, 2) user-oriented approach, and 3) system-oriented approach.

2.2 Social Q&A and credibility

In social Q&As (also called as collaborative question and answer communities or knowledge communities), answers are provided by a large community of users who actively engage in answering a question

on any topic, irrespective of their level of expertise (Rodrigues & Milic-Frayling, 2009). Yahoo! Answers, the most popular social Q&A site in the U.S., counts more than 179 million users and around 15 million users visit the site everyday (Yahoo! Answers Team, 2009). Naver Knowledge-iN, the world's first and Korea's largest social Q&A site, contains more than 70 million questions as of May 7, 2012.¹⁾

Despite their massive popularity, social Q&A sites have been criticized for a lack of credibility. In fact, concerns about Web information credibility are not a new issue. However, social Q&A sites provide a new avenue for credibility research with unique characteristics. First, as Lazar, Meiselwitz, and Feng (2007) noted, there are differences between the credibility of a website and the credibility of a person in an online community: credibility of a website is usually established through the physical nature of the website, whereas credibility of a person is related to trust toward a specific individual. A social Q&A site has characteristics of both a website and an online community, which makes credibility evaluation more complicated. Second, traditionally, authoritative organizations such as universities and companies endeavor to provide relatively objective information in a reliable manner on the Web. In such cases, credibility is granted based on the perceived authority of these organizations. In social Q&As, however, the identity of an information provider is often anonymous or hidden. Third, previous research on Web credibility shows that the design or appearance of a website is a quick

and useful cue for credibility evaluation. In a social Q&A site, however, the visual design of the site does not help at all because people evaluate individual answers within the same site, not individual web pages.

3. Scope

To set boundaries of literature review, this study focuses on previous research in the above-mentioned three domains: 1) research that empirically measured the credibility of information on social Q&As in an objective manner by established standards or experts (content-oriented approach), 2) research that examined users' perceived credibility or identified factors influencing users' credibility evaluation (user-oriented approach), and 3) research that suggested a new way to automate the process of finding high quality information on social Q&A sites (system-oriented approach). Excluded were studies that analyzed the phenomenon of social Q&As theoretically or reported users' general question asking and answer behaviors (e.g., frequency of use, answerers' motivations to share information) with no relation to credibility.

To find articles related to the credibility issue in the context of social Q&As, the author searched Library and Information Science and Technology Abstracts (LISTA), the Association for Computing Machinery (ACM) Digital Library, and Communication and Mass Media (CMM) for English articles,

1) The number of questions in each topic category was added from <http://kin.naver.com/qna/list.nhn>.

<Table 1> Terms that represent ‘credibility’ and ‘social Q&A’ concepts

Credibility	Social Q&A
<ul style="list-style-type: none"> • credibility • quality • authority • relevance • accuracy • expertise • trust 	<ul style="list-style-type: none"> • social question and answer • knowledge community • collaborative knowledge community • collaborative question and answer • community-based question answering • CQA • Q&A

and DBPIA and RISS for Korean articles. Searches were initially conducted from January 15 to 17 in 2012 and confirmed on February 27 in the same year.

In LISTA, ACM, and CMM, search queries were developed to capture two key concepts - ‘credibility’ and ‘social Q&A’. The terms that represent each concept were ANDed to retrieve articles containing both concepts (see Table 1). After browsing the abstract of reviewed articles, relevant articles were selected for review.

In DBPIA and RISS, search terms, ‘지식검색’ and ‘지식포털’ pulled up a large number of articles across various disciplines. Again, by browsing the abstract of each article, only relevant articles were selected for analysis.

4. Literature Review

4.1 Overview

Table 2 shows the number of retrieved articles by publication year.

Eighteen articles have been published in Korean

journals related to credibility on social Q&As. It seems that the popularity of this topic faded in the last year, but it is uncertain if this phenomenon will continue. Outside Korea, the popularity of the topic has remained undiminished as evidenced by the growing number of articles on this topic since 2006.

When comparing the number of articles between the two countries by research areas, the biggest gap exists in the number of articles taking the system-oriented approach (Table 3).

While articles in the system-oriented approach domain accounts for 76% of the retrieved articles outside Korea, it is only 33% in Korea. The discrepancy may reflect different research trends on this topic: computer science researchers in the U.S. are more interested in social Q&As than in Korea. It may be also due to the dissimilar coverage of the databases used. DBPIA and RISS cover a wide array of disciplines, whereas LISTA, ACM, and CMM focus mainly on library and information science, computer science, and communication respectively. Therefore, more comprehensive searches encompassing all disciplines would be needed in the future to produce an exhaustive list of English articles on this topic.

〈Table 2〉 Number of retrieved articles by publication year

	LISTA, ACM & CMM	DBPIA & RISS	Total
2012*	3	0	3
2011	25	1	26
2010	23	3	26
2009	17	3	20
2008	12	2	14
2007	4	4	8
2006	2	4	6
2005	0	0	0
2004	0	1	1
Total	86	18	104

Note: *Articles published between Jan. to Feb. in 2012 were included.

〈Table 3〉 Number of articles by research areas

	LISTA, ACM& CMM	DBPIA & RISS	Total
Content-oriented approach	8(9%)	5(27%)	13(12%)
User-oriented approach	13(15%)	7(39%)	20(19%)
System-oriented approach	66(76%)	6(33%)	72(69%)
Total	87*	18	105

Note: *One article mixed both the content-oriented and the user-oriented approaches, so it was counted in both.

4.2 Content-oriented approach

In this domain, there has been an ongoing discussion about the reliability of information produced by the public versus information produced by professionals or experts.

Harper et al. (2008) conducted a comparative study of three types of Q&A sites - digital reference services, Ask-A services, and social Q&As. They posted a set of test questions to the selected sites and a group of trained college students evaluated answer quality in two dimensions: 1) judged answer quality (e.g., answer correctness), and 2) judged answerer effort (e.g., degree of personalization in answer). Using

a similar research method, Shachaf (2009) analyzed the quality of answers on Wikipedia Reference Desk using three SERVQUAL measures - accuracy, completeness, verifiability - and compared it with library reference services. The findings of these studies are consistent in that the quality of answers on social Q&A sites in aggregate surpasses or matches the level of services found in other services because the social Q&A process results in fast and accurate answers by synthesizing the knowledge of all community members. Fichman (2011) also shows that social Q&As can provide better information, especially more complete and verifiable answers compared to the dyadic reference service. Librarians may

be disappointed by the result, but another comparative study by Tutos and Molla (2010) produced exactly the opposite result. They compared the performance of six search engines, databases, and social Q&A sites - Google, Google on PudMed, PubMed, OneLook, MedQA, and Answers.com Brain Boost - on answer quality across a sample of medical questions. Answers.com, the only social Q&A site, performed the worst of the six on questions requiring medical intervention. In sum, social Q&As do not have the capability as of yet of being a reliable source for medical topics and presumably, other specialized disciplines, but with less specialized topics, it is as good or better than similar services.

Some researchers compared multiple social Q&A sites under the assumption that social Q&A sites may provide answers of different quality because they have different communities and technological platforms. Fichman (2011), who compared four social Q&A sites, found that similar collaborative processes on these sites result in significant differences in answer accuracy, completeness, and verifiability. Park and Jeong (2004) compared three Korean social Q&A sites - Empas Knowledge, Naver Knowledge-iN, and Yahoo! Korea Answers - in terms of response rates, response time, and accuracy, and revealed the differences in their performance for each criterion. They attributed the differences to the number of visitors and answerers, and the number of existing questions and answers in each social Q&A site.

Other studies address a single social Q&A site's credibility. Park, Lee, and Jeon (2006) measured the relevance and credibility of answers in Naver

Knowledge-iN. They showed that while over 90% of the answers examined were relevant for the given questions, only 65.7% were highly credible. Since answerers provide information to a specific question on social Q&As, at least theoretically, every answer is topically relevant to the question. Therefore, topical relevance is not a useful criterion for evaluating answers on social Q&As. Other criteria such as credibility should come into play to identify high quality information. Several researchers measured the accuracy of answers in a particular category on a social Q&A site to determine if the site is a reliable source for the subject matter. For example, Kim (2007) examined the accuracy of answers and the types of cited references in the history category on Yahoo! Answers and suggested an alternative site for educating history.

The utilization of the content-oriented approach to the study of answer quality on Q&A sites can be useful, but poses certain challenges as the concept of information quality is dynamic and multifaceted. Without a universal agreement as to what constitutes a high quality answer, and what measures should be employed, each researcher has developed their own measures to evaluate answer quality. For example, Chen, Ho, and Kim's (2010) quality measures include the credibility and authority of a source and the presence of links to relevant websites. Park, Lee, and Jeon's (2006) credibility criteria include authority of a source, logical opinions, presence of a table/chart/picture, and so on. As a result, the interpretation of the findings requires a detailed account of the concept of information quality in each study.

Table 4 shows the objective measures used in literature. Despite the use of different measures, accuracy is the most common measure in the studies as it is the most important facet of information quality in this domain.

There are two methodological issues to consider when conducting credibility research taking the content-oriented approach: raters' background and types of test questions. Jeon, Kim, and Chen (2010) found

that the rater background makes a difference in evaluating answer quality: graduate students in Master of Science in Information (MSI) programs, who are considered semi-professionals, gave lower answer quality ratings than did undergraduate English majors in their study. This shows that experts tend to be more actively involved in information evaluation. Since the purpose of research in this domain is an objective evaluation, raters' expertise is critical in

<Table 4> Objective measures used in literature

Measures	Literature
Accuracy, correctness	Harper et al. (2008), Shachaf (2009), Fichman (2011), Tutos & Molla (2010), Kim (2007), Park (2008), Park & Jeong (2004), Chang & Lee (2006), Oh et al. (2011), Liu & Agichtein (2008)
Source credibility	Harper et al. (2008), Kim (2007), Park (2008), Park et al. (2006), Park et al. (2006), Chen et al. (2010)
Completeness	Shachaf (2009), Fichman (2011), Park et al. (2006), Oh et al. (2011), Chen et al. (2010)
Verifiability	Shachaf (2009), Fichman (2011), Park et al. (2006)
Response time, timeliness	Park & Jeong (2004), Chang & Lee (2006)
Response rate	Park & Jeong (2004), Chang & Lee (2006)
Relevance, pertinence	Oh et al. (2011), Park et al. (2006), Chen et al. (2010)
Answerer effort	Harper et al. (2008), Park et al. (2006)
Answerer confidence	Harper et al. (2008), Oh et al. (2011)
Being to the point	Park et al. (2006), Chen et al. (2010)
Logicality	Park et al. (2006)
Helpfulness	Harper et al. (2008)
Ease of use of answer	Harper et al. (2008)
Answer friendliness	Harper et al. (2008)
Answerer knowledge	Oh et al. (2011)
Answerer politeness	Oh et al. (2011)
Objectivity	Oh et al. (2011)
Monetary worth of answer	Harper et al. (2008)
Degree of personalization	Harper et al. (2008)
Progress towards receiving an answer	Harper et al. (2008)
Information cited summarized	Chen et al. (2010)
Well-organized and written clearly	Chen et al. (2010)
Conciseness	Chen et al. (2010)

ensuring the validity of research.

Types of test questions influence the outcome of answer quality evaluation as well. Measures such as accuracy are better suited in evaluation of answers to informational questions (e.g., Who is the president of Mexico?) rather than conversational or opinion questions (e.g., Do you believe in evolution?). Thus, it is possible that social Q&A sites that attract fewer informational questions have a lower level of answer quality when using accuracy as a measure compared with other sites that attract more informational questions (Fichman, 2011).

To overcome the methodological limitations, researchers should hire domain experts or train competent raters to ensure the accuracy and consistency of evaluation, and explain the characteristics of test questions so that readers can judge the validity of the study.

4.3 User-oriented approach

Research in this domain addresses users' judgment of information credibility in two areas: 1) users' perceived credibility of an overall social Q&A site, and 2) users' judgments of individual answers on a particular site.

In the first line of research, Park and Jeong (2004) surveyed 253 users of social Q&As in Korea and found that a majority of the participants were satisfied with the quality of answers on social Q&As. Those who were not satisfied reported insufficient information and lack of accuracy and relevance as major reasons of dissatisfaction. When users are sat-

isfied with answer quality on a social Q&A site, their perceived credibility of the overall site increases, which in turn increases users' intentions to continue to use the site (Kim & Han, 2009). This confirms prior research indicating that credibility is a significant predictor of an online community members' desire to get and provide information in the community (Ridings, Gefen, & Arinze, 2002). However, even those who do not give great credence to social Q&As still use them for many reasons: they may want to collect opinions from others who have a similar problem or to find information that is not easily retrieved from a traditional Web search engine (Kim, 2010). Briefly, it is evident that the perceived credibility of a social Q&A site influences one's decision to use it, but it is not always the first and foremost factor. People have other good reasons to use it.

In the second line of research, which focuses on the subjective evaluation of individual answers, a significant barrier is the difficulty of gleaning credibility evaluation from actual users because in the majority of cases, social Q&A site users cannot be observed directly. Accordingly, many researchers based their analyses on user ratings of answers or the comments provided by askers when selecting a best answer. "Best answers" is a feature available in Yahoo! Answers: an asker can select one of the answers as the best from the set of answers they received or let the community vote and select one. Many researchers have considered the best answers chosen by the askers as an appropriate subjective measure of high-quality answers.

In an early study, Gazan (2006) analyzed the con-

tent of high-rating answers in Answerbag to identify factors that make a good answer. He found that questioners generally ranked higher those answerers who did not claim expertise, but provided links to external sources than those who provided information based on their expertise without a reference. This finding is consistent with Harper, Moy, and Konstan (2009), who demonstrated that answers with citations or links to other sites tend to receive higher ratings than those without. In addition to citations and references, askers' preferences for answers with emotional connotation were also observed in the study of Kim and Oh (2009), which analyzed 2,140 comments left on the best answers from Yahoo! Answers. Their findings show that in almost one third of the cases where users select best answers, their selection was based on socio-emotional criteria (e.g., answerer's emotional support, agreement) rather than the content or utility of the answer.

Although these studies provide useful insights into the characteristics of best answers or high-rating answers, the research method they used, which is the content analysis of the answers or comments, is purely descriptive and cannot reveal the underlying motives of the observed pattern accurately. To overcome the limitations, several user studies have been conducted where users were interviewed, observed, or experimented on directly to evaluate the credibility of information.

After interviewing 36 askers of Yahoo! Answers, Kim (2010) identified 22 criteria people use when evaluating the credibility of given answers and grouped them into three classes. This study is meaningful

in that it identified a comprehensive list of credibility criteria askers actually use on a social Q&A site through their own words.

A notable finding in the study is that the askers gauged an answerer's expertise using various cues. For example, a user profile was the most frequently consulted information about an answerer's credentials because it provides the history of answers including the best answer ratings. The content of the answers and the answerers' self-claimed expertise also serve as cues of expertise. In the same vein, Golbeck and Fleischmann (2011) found that text cues containing an answerer's connection to the topic at hand help build trust between the questioner and the answerer. Since it is difficult to evaluate an anonymous answerer's trustworthiness in a social Q&A environment, it is natural that askers rely on the perceived expertise more heavily than trustworthiness by using available cues.

Another important credibility criterion in Kim's (2010) study is 'ratings' on answers given by the members of a social Q&A site. Even though it is not frequently used, it is noteworthy because askers take advantage of the nature of the social Q&A site by relying on fellow users' ratings. All in all, the use of source criteria and user ratings points out the social and collaborative aspect of social Q&A sites where people interact with other people through the question and answering process, as opposed to a typical credibility valuation situation where people interact with individual websites.

Table 5 summarizes the credibility criteria identified in the user-oriented research.

<Table 5> Credibility criteria used by askers in social Q&As

Class	Criteria
Accuracy	Kim (2010), Kim & Oh (2009)
Clarity	Kim (2010), Kim & Oh (2009)
Completeness	Kim (2010), Kim & Oh (2009)
Detail	Kim (2010), Kim & Oh (2009)
Fact	Kim (2010)
Layout	Kim (2010)
Answer length	Kim (2010), Kim & Oh (2009)
Logic	Kim (2010), Kim & Oh (2009)
Novelty	Kim (2010), Kim & Oh (2009)
Spelling/grammar	Kim (2010)
Tone of writing	Kim (2010), Kim & Oh (2009)
Topicality	Kim (2010)
Answerer's attitude	Kim (2010), Kim & Oh (2009)
Known answerer	Kim (2010)
Perceived expertise based on the answer	Kim (2010), Kim & Oh (2009)
Perceived expertise based on the answerer's profile	Kim (2010)
Reference to external sources	Kim (2010), Gazan (2006)
Self-claimed expertise/qualification	Kim (2010), Kim & Oh (2009), Gazan (2006)
Available alternatives	Kim & Oh (2009)
Ratings on the answer	Kim (2010)
Usefulness	Kim (2010), Kim & Oh (2009)
Verifiability	Kim (2010), Kim & Oh (2009)
Answerer's experience	Golbeck & Fleischmann (2011), Kim & Oh (2009)
Presence of an answerer's photo or avatar	Golbeck & Fleischmann (2011), Laramie (2009)
Quickness	Kim & Oh (2009)
Answerer's effort	Kim & Oh (2009)
Emotional support	Kim & Oh (2009)
Agreement	Kim & Oh (2009)
Taste	Kim & Oh (2009)
Humor	Kim & Oh (2009)

Before moving on to the next section, one study should be noted. Previous studies took either the content-oriented approach or the user-oriented approach. Exceptionally, Oh, Warroll, and Yi (2011) attempted to identify the gap between these two approaches. They recruited three groups of evaluators - users of Yahoo! Answers, health reference li-

brarians, and nurses - with a view to comparing the evaluation of general users against experts in the health domain. Preliminary results indicate that general users were more generous and rated answer quality higher than health reference librarians. The authors warn that the general public may encounter problems when making health decision using the

information obtained from a social Q&A site because they do not have sufficient evaluation skills.

4.4 System-oriented approach

In this domain, there are two primary approaches: 1) a feature-based approach and 2) a graph-based approach.

The feature-based approach analyzes answer quality or answerer's authority based on textual features (e.g., typos, syntactic and semantic complexity) or non-textual features (e.g., click count, N of best answers).

Relying on textual features, some researchers developed complicated language models to model the interests of an answerer or a questioner on a social Q&A site. Wang et al. (2009) calculated associations between questions and answers in Q&A systems under the assumption that several different types of direct associations exist between questions and answers. Guo et al. (2008) implemented the probabilistic latent semantic analysis (PLSA) model to analyze the questions addressed by answerers and built a model that depicts the interests of the answerers. The similarity between answerers' interests and questions for recommendation was calculated. One of the limitations of existing language models is that since question description on a social Q&A site is usually shorter than ordinary documents, it is very hard to build a question model for similarity calculation (Zheng et al., 2012). More importantly, recommending answerers solely on the basis of a language model is not a sufficiently accurate ap-

proach because it does not discriminate the relative expertise levels of answerers.

On the other hand, non-textual features such as users' recommendations, click count, and best answers can estimate answer quality better because those features typically represent the expertise of the answerer although it is difficult to get enough non-textual information in the early stage of a social Q&A site (Kim, Park, & Lee, 2010) and even on a mature social Q&A site, it is not strong enough during the early participation of a user. A major methodological benefit of evaluating non-textual features is that it can avoid the complexity of linguistic content analysis in addition to being language independent (Jeon et al., 2006). Bouguessa, Dumoulin, and Wang (2008) and Jeon et al. (2006) incorporated the non-textual features into the language model-based retrieval model and achieved a significant improvement in retrieval performance.

Among various non-textual features, the number of "best answers" given by community users is often considered as an indicator of authority of the answerer (e.g., Dom & Paranjpe, 2008). Unfortunately, developing an algorithm that can pick the best answer from a set of answers for a given question is extremely difficult because the act of selecting the best answer depends on many factors including socio-emotional criteria as shown in Kim and Oh (2009).

When answer quality is characterized by best answers or other non-textual features, answer length is the most significant feature for predicting answer quality (Agichtein et al., 2008; Adamic et al., 2008). Bouguessa, Dumoulin, and Wang (2008) corroborate

that just the very basic metric of reply length was most predictive of whether the answer would be selected as the best answer. In certain topic categories, however, the number of competing answers and the history of the answerer were more likely to predict answer quality. It is inferred that a lengthy answer is more complete, providing more accurate and detailed information. It is also possible that an asker appreciates the answerer's effort and time to create such a long answer, so selects it as the best answer.

Another line of work in this domain explores the rankings of social Q&A users through the construction of the user graphs and the use of well-known link analysis algorithms such as PageRank (Page et al., 1998) and HITS (Kleinberg, 1999). These algorithms make use of the question-reply structural information in an online community, but not the content of questions or answers. Jurczyk and Agichtein (2007) developed an application of the HITS algorithm to a social Q&A site, especially the user interactions graph, and found a positive correlation between authority calculated with the HITS algorithm and answer quality. Agichtein et al. (2008) further derived multiple user interaction graphs for different kinds of relationships, such as asking-answering, selecting best answers, abuse reporting and answer voting/rating. A primary limitation of these studies is their definition of expertise is question-independent by relying on the global ranking in the community. However, it is more logical to assume that each answerer has different levels of expertise for different topics. Therefore, Suryanto et al. (2009) suggested

question-dependent expertise-based methods and found that the methods outperform methods using answer features only. Kim, Park, and Lee (2010) created an algorithm that considers user centrality based on Social Network Analysis (SNS). Their algorithm which combines textual, non-textual information as well as link analysis outperforms the feature-based algorithm only.

What should be noted at this point is that there are conflicting results in research using the graph-based approach. For example, Jurczyk and Agichtein (2007) performed link analysis on a dataset from Yahoo! Answers by using a slight adaptation of the HITS algorithm. Their results indicate that the HITS algorithm outperformed simple graph measures such as InDegree. On the contrary, Bouguessa, Dumoulin, and Wang (2008) show that a simple technique such as InDegree is the most appropriate for rating the authority level of each user in Yahoo! Answers. Further research should reanalyze the characteristics of the data sets and methodological procedures of the studies to reconcile the conflicting results.

5. Agendas for Future Research

Reviewing previous research on the credibility of information on social Q&As lead to the identification the following avenues for future research for each of the three domains as well as across the domains.

5.1 Across the domains

From the review, several future research areas are suggested across the domains.

First, almost all previous studies attempted to measure answer quality at a specific time moment. An exception is Liu and Agichtein's study (2008), which investigated the temporal changes of Yahoo! Answers during the years of 2006 and 2007 with respect to its effectiveness in answering factoid, opinion, and complex questions. Their finding is that as the site keeps growing rapidly in size, its answering quality for factoid question degrades. This study is meaningful in that it is the first attempt to investigate answer quality from more than one time spot. However, it gathered the data only between two years, so a longitudinal study for a longer time period is needed to confirm the finding. Moreover, the reasons of the degrading quality should be further analyzed to suggest ways to improve answer quality on social Q&A sites.

Additionally, longitudinal studies should provide more comprehensive observations to understand the changes in people's perception of credibility over time. Considering that the amount of time people spend on a particular social medium influences credibility perception, it would be interesting to see, for example, if Yahoo! Answers' users regard the site as more credible as they get familiar with the site even though the objective quality of answers degrades over time, as shown in Liu and Agichtein (2008).

Second, most previous research examined Yahoo! Answers as a research setting. Considering its domi-

nant status in the U.S. and around the world, Yahoo! Answers' popularity in the academia is understandable, but more diverse social Q&A sites should be explored and compared in the future. In connection to the first research area suggested above, it would be interesting to see how the answer quality of diverse social Q&A sites change over time. The different collaborative processes provided by different social Q&A sites may influence the temporal changes of answer quality. For example, while Yahoo! Answers maintains a repository of answers that includes incorrect information, WikiAnswers allows for answer modification by the members of the site. Does the answer quality of WikiAnswers improve with the self-editing mechanism as time goes by? Which mechanism produces more accurate information in the long run? These are the example of research questions that should be examined in the future.

Third, further research regarding the gap of subjective evaluation and objective evaluation would be of great help in educating users on effectively evaluating information on social Q&A sites. It is alarming that users who do not have sufficient evaluation skills to filter out non-credible information act upon the information obtained from a social Q&A site as shown in Oh, Warroll, and Yi (2011). By comparing the evaluation of users against experts in more diverse topic categories other than medical topics, this line of research would help design instruction materials and services for educating social Q&A users.

Fourth, a primary limitation of system-oriented research is they do not elucidate the underlying rea-

sons why certain features are associated with answer quality. Research consistently shows that lengthier answers tend to be selected as the best answers. Although it is conjectured that answer length may be a proxy for other variables such as completeness, it is not known exactly what mediating factors account for the observed relationship. Future research taking the user-oriented approach could reveal the underlying reasons for the relationship and inform researchers working in this domain so that they can build more sophisticated credibility evaluation model.

5.2 Content-oriented approach

In this domain, a comparison across diverse topic areas is needed to determine if certain topic areas are more credible than others on a social Q&A site. For example, answers for scientific topics may be more credible than for political topics on a social Q&A site. Since social Q&A sites cover a wide range of topics and each site tends to attract certain topics of questions more than others, it is necessary to examine which topic categories include more credible answers on a specific social Q&A site or across the sites. The identification of relatively credible topics on a social Q&A site can help users select an appropriate social Q&A site for their needs and help researchers characterize social Q&A sites compared with other similar services.

5.3 User-oriented approach

First, the novelty of social Q&As has yielded most-

ly exploratory and descriptive studies. Although they are fruitful for understanding how people perceive answer quality on social Q&As, prior research fails to illuminate specific factors associated with the perception. In the future, a more systematic approach should uncover factors leading to credibility perception, and measure their effect, going beyond a mere description of users' credibility perceptions.

Second, previous research identified a rich array of individual credibility criteria, but did not reveal the complicated relationship among the criteria. It is unclear whether askers favor answers containing certain criteria at the expense of other criteria (e.g., selecting an answer lacking completeness because of the answerer's good manners) or whether they consider the answerer's manners after other criteria are satisfied. More obtrusive research methods such as observation would provide an integrated picture of how people actually evaluate answers in social Q&A.

Third, as researchers compared the answer quality of social Q&As with similar services in the objective manner, comparative studies should be conducted to know people's perception of the credibility of diverse Q&A services. Since people may apply different dimensions of credibility to different services, more research is needed to examine how one's construct of credibility is related to credibility evaluation in a specific type of Q&A service. For example, is accuracy more important when users evaluate digital reference services than social Q&A sites? If not accuracy, what is the most important dimension of credibility on social Q&A sites?

5.4 System-oriented approach

First, technologies do not yet exist to automatically answer open domain complex and subjective question (Bian et al., 2008). Recent research on automatic evaluation shows that even for factual questions, evaluation is extremely difficult and has many challenges. It was natural, therefore, that prior research focused on factual questions to identify high quality answers or authoritative answerers. While the notion of best answer may apply well to factual questions, research should be expanded to other types of questions which expect answers tailored to different personal preferences.

Second, while link analysis techniques have proven to be a useful method to identify authoritative users, more research is needed to reconcile conflicting results regarding the effectiveness of individual link analysis techniques. Researchers often compare the techniques they developed with a baseline model, for example, feature-based algorithms. It is already known that link analysis techniques outperform the feature-based algorithms only, but what is still not known is which link analysis techniques work the best under what conditions on a social Q&A site. By reanalyzing the methodological procedures and research settings of previous research or conducting a new study that delves into this issue, the conflicting results regarding which link analysis techniques work the best should be reconciled to suggest the best technique for the social Q&A environment.

Third, social Q&A sites exhibit a rich variety of information sources: in addition to the answer itself,

there is a wide array of non-textual information available, such as links between items, explicit quality ratings from members of the community, and user profiles. Although some non-textual features have been utilized in prior research, there are other possible features worth exploring. For example, Yahoo! Answers offers “Answer Network”, in which people create their own private networks with other users for question and answering. Perhaps, people regard the members of the private network more credible than anonymous answerers outside the network. Exploring such site-specific features could identify important features that can predict answer quality more accurately.

It is possible that researchers conduct studies outside the above-discussed three domains. For example, previous research has revolved around information seekers, examining how they evaluate credibility or automating the process of retrieving credible information for them. Another side of credibility that has received less attention is lay information providers who post answers in social Q&As. Research can be conducted, for example, to link their perception of credibility and other variables to the practices by which they establish credentials and to develop an information system for them.

6. Conclusion

The amount of user-generated content in social media is exploding, completely changing the way people seek, share, and evaluate information. In terms

of credibility evaluation, understanding the methodologies of objective credibility, developing a model for subjective credibility, and creating effective credibility assessment systems are all crucial questions to understand the credibility phenomenon in social Q&As.

This paper summarizes previous research on credibility conducted in the context of social Q&As in three broad domains. Through the literature review, some research themes have emerged and they suggest promising directions for future research in each domain.

References

- Adamic, L. A., Zhang, J., Bakshy, E., & Ackerman, M. S. (2008). Knowledge sharing and Yahoo Answers: Everyone knows something. *Proceedings of the 17th International Conference on World Wide Web*, 665-674.
- Agichtein, E., Castillo C., Donato, D., Gionis, A., & Mishne, G. (2008). Finding high-quality content in social media. *Proceedings of the International Conference on Web Search and Data Mining*, 183-194.
- Barry, C. L. (1994). User-defined relevance criteria: An exploratory study. *Journal of the American Society for Information Science*, 45(3), 149-159.
- Bateman, J. (1999). Modeling the importance of end-user relevance criteria. *Proceedings of the 62nd Annual Meeting of the American Society for Information Science*, 36, 396-406.
- Bian J., Liu, Y., Agichtein, E., & Zha, H. (2008). Finding the right facts in the crowd: factoid question answering over social media. *Proceedings of the 17th International Conference on World Wide Web*, 467-476.
- Bouguessa, M., Dumoulin, B., & Wang, S. (2008). Identifying authoritative actors in question-answering forums: the case of Yahoo! Answers. *Proceedings of the 14th International Conference on Knowledge Discovery and Data Mining*, 866-874.
- Chen, Y., Ho, T., & Kim, Y. (2010). Knowledge market design: A field experiment on Google Answers. *Journal of Public Economics Theory*, 12(4), 641-664.
- Dom, B., & Paranjpe, D. (2008). A bayesian technique for estimating the credibility of question answerers. *Proceedings of the Society for Industrial and Applied Mathematics on Data Mining*, 399-409.
- Fichman, P. (2011). A comparative assessment of answer quality on four question answering sites. *Journal of Information Science*, 37(5), 476-486.
- Flanagin, A. J., & Metzger, M. J. (2000). Perceptions of Internet information credibility. *Digital Media, Youth, and Credibility*, 77(3), 515-540.

- Flanagin, A. J., & Metzger, M. J. (2008). The credibility of volunteered geographic information. *GeoJournal*, 72(3-4), 137-148.
- Gazan, R. (2006). Specialists and synthesists in a question answering community. Proceedings of the 69th Annual Meeting of the American Society for Information Science and Technology. Retrieved from <http://eprints.rclis.org/handle/10760/8793>
- Golbeck, J., & Fleischmann, K. R. (2010). Trust in social Q&A: The impact of text and photo cues of expertise. Proceedings of the 73th Annual Meeting of the American Society for Information Science and Technology, 47(1), 1-10.
- Guo, J., Xu, S., Bao, S., & Yu, Y. (2008). Tapping on the potential of Q&A community by recommending answer providers. Proceedings of the 17th ACM Conference on Information and Knowledge Management, 921-930.
- Harper, F. M., Moy, D., & Konstan, J. A. (2009). Facts or friends?: Distinguishing informational and conversational questions in social Q&A sites. Proceedings of the ACM Conference on Human Factors in Computing Systems, 759-768.
- Harper, F. M., Raban, D., Rafaeli, S., & Konstan, J. A. (2008). Predictors of answer quality in online Q&A sites. Proceeding of the 26th Annual SIGCHI Conference on Human Factors in Computing Systems, 865-874.
- Hovland, C. I., Janis, I. L., & Kelley, H. H. (1953). *Communication and persuasion*. New Haven, CT: Yale University Press.
- Jeon, G. Y., Kim, Y. M., & Chen, Y. (2010). Re-examining price as a predictor of answer quality in an online Q&A site. Proceedings of the 28th Annual SIGCHI Conference on Human Factors in Computing Systems, 325-328.
- Jeon, J., Croft, B., Lee, J. H., & Park, S. (2006). A framework to predict the quality of answers with non-textual features. Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, 228-235.
- Jurczyk, P., & Agichtein, E. (2007). Discovering authorities in question answer communities by using link analysis. Proceedings of the 16th ACM Conference on Information and Knowledge Management, 919-922.
- Kim, B., & Han, I. (2009). The role of trust belief and its antecedents in a community-driven knowledge environment. *Journal of the American Society for Information Science and Technology*, 60(5), 1012-1026.
- Kim, D. J., Park, G. W., & Lee, S. H. (2010). QualityRank: Measuring authority of answer in Q&A community using social network analysis. *Journal of KIISE: Databases*, 37(6), 343-350.
- Kim, M. K. (2007). The current status and problems of a knowledge community in history: Focusing on

- Naver Knowledge-iN. Sangmyung University. The Graduation School of Education. Master Thesis.
- Kim, S. (2010). Questioners' credibility judgments of answers in a question answering community site. *Information Research*, 15(2) paper 432. Retrieved from <http://InformationR.net/ir/15-2/paper432.html>
- Kim, S., & Oh, S. (2009). Users' relevance criteria for evaluating answers in a social Q&A site. *Journal of the American Society for Information Science and Technology*, 60(4), 716-727.
- Kleinberg, J. (1999). Authoritative sources in a hyperlinked environment. *Journal of the ACM*, 46(5), 604-632.
- Lazar, J., Meiselwitz, G., & Feng, J. (2007). Understanding Web credibility: A synthesis of the research literature. *Foundation and Trends in Human-Computer Interaction*, 1(2), 139-202.
- Liu, Y., & Agichtein, E. (2008). On the evolution of the Yahoo! Answers QA community. *Proceedings of the 31st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 737-738.
- Oh, S., Warroll, A., & Yi, Y. (2011). Quality evaluation of health answers in Yahoo! Answers: A comparison between experts and users. *Proceedings of the 74th Annual Meeting of the American Society for Information Science and Technology*, 48(1), 1-3.
- Page, L., Brin, S., Motwani, R., & Winograd, T. (1998). The Pagerank citation ranking: Bringing order to the Web. Technical report, Stanford Digital Library Technologies Project.
- Park, J. B., & Jeong, D. Y. (2004). An empirical study on Web-based question-answer services. *Journal of the Korean Society for Information Management*, 21(3), 83-98.
- Park, S., Lee, J. H., & Jeon, J. (2006). Evaluation of the documents from the Web-based question and answer service. *Journal of the Korean Society for Library and Information Science*, 40(2), 299-314.
- Rieh, S. Y., & Danielson, D. R. (2007). Credibility: a multidisciplinary framework. *Annual Review of Information Science and Technology*, 41, 307-364.
- Ridings, C. M., Gefen, D., & Arinze, B. (2002). Some antecedents and effects of trust in virtual communities. *Journal of Strategic Information Systems*, 11(3/4), 271-295.
- Rodrigues, E. M., & Milic-Frayling, N. (2009). Socializing or knowledge sharing? Characterizing social intent in community question answering. *Proceedings of the 18th ACM Conference on Information and Knowledge Management*, 112-1136.
- Self, C. S. (1996). Credibility. In M. Salwen & D. Stacks (Eds.), *An integrated approach to communication theory and research* (pp. 421-441). Mahwah, NJ: Lawrence Erlbaum Associates.
- Suryanto, M. A., Lim, E. B., Sun, A., & Chiang, R. H. L. (2009). Quality-aware collaborative question answering: methods and evaluation. *Proceedings of the 2nd ACM International Conference on Web Search and Data Mining*, 142-151.
- Taylor, L. D. (2009, May). Static avatars, social presence, and online answers. Paper presented at the annual

- meeting of the International Communication Association, Marriott, Chicago, IL.
- Taylor, R. S. (1986). Value-added processes in information systems. Norwood, NJ: Ablex.
- Tseng, S., & Fogg, B. J. (1999). Credibility and computing technology. *Communications of the ACM*, 42(5), 39-44.
- Tutos, A., & Molla, D. (2010). A study on the use of search engines for answering clinical questions. *Proceedings of the 4th Australasian Workshop on Health Informatics and Knowledge Management*, 108, 61-68.
- Wang, X., Tu, X., Feng, D., & Zhang, L. (2009). Ranking community answers by modeling question-answer relationships via analogical reasoning. *Proceedings of the 32nd International ACM SIGIR Conference on Research and Development in Information retrieval*, 179-186.
- Wilson, P. (1983). *Second-hand knowledge: An inquiry into cognitive authority*. Westport, CT: Greenwood Press.
- Yahoo! Answers Team (2009, October 5). Did you know? Message posted to <http://yanswersblog.com/index.php/archives/2009/10/05/did-you-know/>
- Zhang, J., Ackerman, M. A., & Adamic, L. (2007). Expertise networks in online communities: Structure and algorithms. *Proceedings of the 16th International Conference on World Wide Web*, 221-230.
- Zheng, X., Hu, Z., Xu, A., & Chen, D. (2012). Algorithm for recommending answer providers in community-based question and answering. *Journal of Information Science*, 38(1), 3-14.