

Content Quality Assessment Related Frameworks for Social Media

Kevin Chai, Vidyasagar Potdar and Tharam Dillon

Digital Ecosystems and Business Intelligence Institute,
Curtin University of Technology, Perth, Australia
kevin.chai@postgrad.curtin.edu.au, {v.potdar, t.dillon}@curtin.edu.au

Abstract. The assessment of content quality (CQ) in social media adds a layer of complexity over traditional information quality assessment frameworks. Challenges arise in accurately evaluating the quality of content that has been created by users from different backgrounds, for different domains and consumed by users with different requirements. This paper presents a comprehensive review of 19 existing CQ assessment related frameworks for social media in addition to proposing directions for framework improvements.

Keywords: content quality assessment, social media, discussion forums, wikis, weblogs.

1 Introduction

Information quality (IQ) has been widely defined in literature by its fitness for use. Existing research in the field of assessing IQ within traditional information systems is relatively mature and has led to the discovery and validation of numerous quality dimensions and metrics [2, 15, 17, 29, 35].

Knight & Burn (2005) have presented a comprehensive review of literature in the realm of assessing IQ frameworks. Common IQ dimensions that were identified from their review are presented in Table 1. The definitions of these dimensions have been included from Wang & Strong (1996) [20, 37].

However, many IQ frameworks for traditional information systems fail to address quality problems caused through many-to-many mappings. This mapping refers to information created in different contexts to support the requirements of different activities and perspectives at the same time. Many-to-many mappings are common in social media [34].

Social media is an umbrella term that integrates technology, social interaction and *user generated content (UGC)* and differs from traditional broadcasting and publishing media such as television, radio and newspapers. Social media is commonly comprised of technologies such as instant messaging programs, discussion forums, weblogs and wikis [33].

Facebook, *YouTube* and *Wikipedia* are examples of popular social media websites that service a large population of users (ranging in the millions) who contribute and/or

consume content. Usage growths are predicted to continue in the near future as social media becomes more readily adopted [9, 22, 24, 32].

In most instances, social media user populations are diverse and consist of people from differing geographical locations, backgrounds, beliefs, motivations and expertise. As a result, the quality of UGC varies drastically within social media and can range from excellent to abuse and spam [1].

Excessive user contributions of low quality can also lead to what is termed as *information overload* which describes the situation when users feel they are swamped with unwanted information. Information overload can eventuate in user withdrawal from using a social media service [7, 31].

The assessment of IQ, or as we have termed as *content quality (CQ)*, in social media can allow for the identification and favourable presentation of high quality content over poor quality content. This assessment can therefore assist in handling the problem of information overload. However, certain difficulties arise in defining CQ for social media because content that is perceived as excellent quality by one user might be considered as poor quality to another user.

Due to the large population of social media users and vast volume of UGC in many social media services, it would be useful to automate the process of CQ assessment (CQA) over manual evaluation techniques. Furthermore, users that contribute quality content can be identified and possibly rewarded which can provide motivation for users to continue producing high quality content. Additionally, users that contribute low quality content could be identified and be educated on how they can improve the quality of their contributions [4, 6, 7, 10].

The ability to develop an accurate and objective measure for quality becomes challenging when one takes into consideration the diversity of user requirements apparent in most social media services. Naturally, one might consider for CQA to be evaluated at a community level in addition to integrating user feedback.

User feedback such as dwell time (implicit user feedback) or quality ratings for UGC contributed by other users (explicit user feedback) provides a decentralised approach to CQA and is the most popular mechanism used by the CQA frameworks evaluated in this paper (see Table 1) and many social media websites.

User feedback however should not be used as the only method in CQA. Such an approach puts forward the assumption that user feedback is honest and impartially represents the real opinions of users. However, not all users will genuinely provide feedback nor may they have the sufficient background and expertise to provide accurate CQ judgements [11, 21, 23].

Table 1 shows a summary of different CQ dimensions used for CQA within social media. This table has been populated by analysing the quality factors employed by 19 CQ related frameworks in recent literature and maps these factors to common IQ dimensions identified by [20].

This paper aims to provide a review of existing CQA related frameworks for social media. The term related is used to include frameworks that may not be designed exclusively for CQA such as [21] which proposes a model to validate user ratings used for quality assessment. It also aims to provide the necessary groundwork for other researchers to develop improved CQA frameworks.

Table 1. Commonly used IQ Dimensions in evaluated CQA related frameworks adapted from Knight & Burn's (2005) [20]

Dimension	Freq	Definitions * Provided by Wang & Strong 1996 [37]	Sources	
1	User Feedback	14	users provides either an implicit (e.g. dwell time reading content) or explicit (e.g. provides a rating) quality evaluation of the content. Their quality evaluation may include the assessment of any or combination of any CQ dimensions.	[1, 3, 7, 11, 12, 14, 16, 18, 19, 21, 23, 29, 36, 38]
2	Amount of data	8	extent to which the quantity or volume of available data is appropriate*	[1, 3, 14, 16, 18, 19, 25, 36]
3	Reputation	6	extent to which information is highly regarded in terms of source or content*	[1, 11, 12, 14, 23, 36]
4	Objectivity	5	extent to which information is unbiased, unprejudiced and impartial*	[13, 21, 26, 27, 28]
5	Relevancy	5	extent to which information is applicable and helpful for the task at hand*	[5, 18, 26, 27, 29]
6	Reliability	4	extent to which information is correct and reliable*	[5, 26, 29, 36]
7	Completeness	3	extent to which information is not missing and is of sufficient breadth and depth for the task at hand*	[13, 19, 29]
8	Accuracy	2	extent to which data are correct, reliable and certified free of error*	[11, 29]
9	Timeliness	2	extent to which the information is sufficiently up-to-date for the task at hand*	[7, 36]
10	Understandability	2	extent to which data are clear without ambiguity and easily comprehended*	[5, 29]
11	Value-Added	2	extent to which information is beneficial, provides advantages from its use*	[26, 29]
12	Consistency	1	extent to which information is presented in the same format and compatible with previous data*	[29]
13	Security	1	extent to which access to information is restricted appropriately to maintain its security*	[29]
14	Accessibility	1	extent to which information is available, or easily and quickly retrievable*	[29]
15	Believability	1	extent to which information is regarded as true and credible*	[28]
16	Useful	1	extent to which information is applicable and helpful for the task at hand*	[29]

This paper is structured as follows. Section 2 categorises CQA related frameworks into social media platforms and explains how these frameworks will be evaluated and compared against one another. Section 3-9 surveys existing CQA related frameworks for social media, critically analyses these frameworks over eight comparative criteria

and proposes directions for improvement. Finally, Section 10 concludes the paper with future directions.

2 Content Quality Related Frameworks

In this section, we survey 19 different CQA related frameworks for social media. In this review we outline the different approaches applied in CQA of UGC and which domain the framework is designed and validated against within the evaluated literature. This survey includes the evaluation of frameworks that have been applied to different types of social media including discussion forums, peer-to-peer applications, question and answering portals, review portals, weblogs and wikis.

Every effort has been taken to study the most recent and relevant CQA related frameworks. An extensive review of CQA related frameworks is conducted but due to the scope of this paper, only a brief description is provided for each framework. These frameworks are evaluated within the following categories:

- Forums (Section 3)
- Question & Answering Portals (Section 4)
- Peer-to-peer (P2P) Applications (Section 5)
- Review Portals (Section 6)
- Weblogs (Section 7)
- Wikis (Section 8)
- Not Validated Against an Application Domain (Section 9)

A critical evaluation and comparison of these frameworks is presented after frameworks for each platform are described. A comparison of review portals is not presented as only one review portal framework is evaluated in this survey. This comparison evaluates CQA related frameworks along 8 different criteria (as shown in Table 2) to answer the following questions:

- Q1. Has the framework been developed into a prototype to automate CQA?
- Q2. Is the proposed framework domain independent?
- Q3. Are objective results from the CQ framework evaluated against subjective human judgment for validation purposes?
- Q4. Is theory used in the development of the CQ framework as opposed to only heuristics?
- Q5. What content items are evaluated in the framework?
- Q6. Are semantic content analysis techniques employed?
- Q7. Is explicit user feedback incorporated in CQA?
- Q8. Is user fraud handled either implicitly or explicitly?

The noteworthy answers to some of these questions will now be discussed in detail for selected frameworks throughout Sections 3-9.

Table 2. Critical Evaluation of CQA Related Frameworks

Source	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8
Forums								
[3]	Yes	Possibly ¹	Yes	No	Yes	Posts, threads	Yes	Yes ²
[18]	Yes	No	Yes	Yes	Yes	Posts, threads	No	Yes ²
[19]	Yes	No	No	No	No	Posts	Yes	Yes ²
[25]	Yes	Possibly ¹	Yes	No	Yes	Posts	No	No
[26]	Yes	Yes	Yes	No	Yes	Posts	No	No
[38]	Yes	No	No	No	No	Posts	Yes	Yes ²
Question & Answering Portals								
[1]	Yes	Yes	Yes	No	Yes	Q&A pairs	Yes	Yes ²
[14]	Yes	Yes	Yes	No	No	Q&A pairs	Yes	Yes ²
Peer-to-peer Applications								
[7]	Yes	Yes	No	Yes	No	Files, user ratings	Yes	Yes ²
[11]	Yes	Yes	NA	No	No	Resources / files	Yes	Yes
Review Portals								
[21]	Yes	Yes	No	No	No	Reviews, user ratings	Yes	Yes
Weblogs								
[27]	Yes	Yes	Yes	No	Yes	Articles	No	No
[36]	Yes	Possibly ¹	Yes	No	Yes	Articles, profiles	Yes	No
Wikis								
[16]	Yes	Possibly ¹	Yes	No	No	Articles, ratings	Yes	Yes ²
[12, 23]	Yes	Yes	No	No	No	Article, edit history	Yes	No
Not Validated Against an Application Domain								
[5]	No	Yes	NA	Yes	NA	search results	NA	NA ³
[13]	No	Possibly ¹	NA	Yes	Yes	Messages	No	NA ³
[28]	No	Possibly ¹	NA	No	Yes	Messages	No	NA ³
[29]	No	Yes	NA	Yes	Yes	NA	NA	NA ³

¹ The CQA related framework may be applicable to other domains if minor refinements and fine tunings were made for the ported domain.

² Fraud is managed to a certain extent through the use of explicit user feedback or content filters but the system is still open to manipulation.

³ It is assumed that the manual approaches would be free of fraud as human judges would manually assess the content.

3 Forums

Chai (2007) [3] – online gaming forum

A user contribution measurement (UCM) model was developed to rank users based on both the quantity and quality of contributions within online discussion forums. The initial major significance of developing this model was for its introduction in revenue sharing forums. The model evaluates 16 features which are common and easily retrievable amongst various forum platforms. The main quality feature evaluated in this framework was the length of a post.

Kim et al. (2006) [18] – psychology and computing student forums

The contents of a message, relationships amongst the messages in a thread, and relations between each message and the thread to which it belongs are systematically

analysed to assess discussion quality. The authors propose the use of the Theory of Speech Acts which can represent a positive, negative or neutral response to a previous post depending on its attitude and recommendation. Additionally a Rocchio-style classifier is used to build topic profiles through automatic ontology induction. This provides a mechanism to assess topic shifts within a thread. Threads that remain topic focused may be inferred as higher quality than threads containing a large variety of loosely coupled topics.

Klamma et al. (2007) [19] –information systems learning design forum

This model was introduced as an incentive mechanism to encourage user contribution and participation within an online discussion forum. Quality ratings provided by other users was used as features in evaluating user contribution.

Lui, Li & Choy (2007) [25] - networking and multimedia student forum

An algebraic vector-space model coupled with latent semantic analysis (LSA) and a Naïve Bayes (NB) classifier is applied to classify discussion forum messages under different categories. This classification of messages can assist in content quality assessment as messages belonging to a particular category might be included or given a higher quality weighting than messages from other categories. An example would include classifying forum posts into knowledge seeking (questioning) and knowledge contributing (answering) classes for quality assessment over commentary posts.

McKlin et al. (2002) [26] - student discussion forum

The use of neural networks is proposed to analyse the cognitive effort behind textual messages within educational online discussion forums. First, messages were converted into a database. Secondly, a word count tool was used to count the number of messages within self-defined cognitive presence categories and the General Inquirer category of terms [8]. Thirdly, a group of messages were categorised by human judgement and then used as a training set to train a neural network. Lastly, for reliability purposes, the machine-classified messages were subjectively validated by human judgement. This method provides a mechanism at measuring the cognitive effort in posting messages and may be used to infer a quality measure for student contributions.

Wiertz & Ruyter (2007) [38] - firm-hosted commercial online discussion forum

A member-controlled point reward system was introduced into an online discussion forum. Points (ranging from 1 to 10) can be assigned by the user that posted a question based on the quality of answers they receive. When members accumulate a certain number of points they can be granted 6 different types of hats which represent 6 different membership levels. These membership levels are specifically designed to indicate the quality of knowledge contributions made by an individual user.

3.1 Evaluation of Forums

A large portion of the reviewed CQA related frameworks have been developed for online discussion forums. These frameworks have been developed and tested predominately for a teaching and learning domain. The main purpose for creating these frameworks was to provide an automated process of assessing the quality of student discussions. Two non-teaching learning frameworks [3] and [26] were validated against an online gaming forum and a firm-hosted product support forum respectively.

While the majority of these frameworks are domain dependent (i.e. it is unlikely a teaching and learning framework could be applied to an online gaming forum), it is believed that these two frameworks, [3, 26], could be refined and validated against other domains. However, these frameworks are not believed to be applicable to other social media platforms such as wikis as they predominantly evaluate forum based content items such as posts and threads.

The CQA frameworks developed by [18] proposed the use of the Theory of Speech Acts to infer what a forum message might be stating about another forum message (e.g. acknowledge, criticise etc...). The use of this theory can assist in discovering rich collaborative discussion threads which could be inferred as high quality threads based on CQ dimensions such as objectivity and completeness and believability.

Additionally, [26] developed a neural network to estimate the amount of cognitive effort expended by students in creating discussion forum messages. It is believed that their framework would have been used to reward students that put more thought into their postings. However, a counter argument to such an approach is that more cognitive effort does not necessarily result in the creation of higher quality content.

The content item evaluated by the majority of frameworks were forum posts with the exception of [3] and [18] which also evaluated thread related factors such as coherence, the number of responses within a thread and whether a thread was marked as a sticky thread. Through semantic analysis, [18] developed a metric to determine thread coherence which describes whether messages within a thread are focused on a particular topic or on a large range of unrelated topics. Such a metric may need to be further improved to handle concept drift, where threads may naturally expand into other topic areas as they evolve over time.

Frameworks developed by [3], [25] and [26] employed classification techniques to categorise forum messages (UGC) into categories including but not limited to knowledge seeking and knowledge creation messages.

Explicit user feedback was only utilised by [19] in which user ratings could be provided for forum messages and [38] in which a user could pose a question and could allocate ratings to answers they would receive from other users. However, it was discovered with the latter framework that users would often give good ratings for low quality answers as users were sometimes happy just to receive a reply to their question. Therefore the accuracy of quality ratings within this forum is questionable.

Whilst not previously mentioned, [18] employed Rhetorical Structure Theory analysis to verify student reports that tutors assisted in facilitating discussions. This mechanism could be used to verify potential fraudulent student feedback for tutor assessment. Other frameworks employed explicit user feedback and/or content filters to manage fraud to some extent.

4 Question & Answering Portals

Aigitech et al. (2008) [1] - Yahoo! Answers

The authors of this paper employ semantic content analysis with tag n-grams, part-of-speech (POS) tags and language models to assess the quality of question and answer text. This quality analysis includes the evaluation of punctuation, typos, syntactic and semantic complexity as well as grammar. Link analysis algorithms such as HITS and PageRank are used to infer quality by evaluating relationship between users and content. Usage statistics are also incorporated as quality evidence in CQA.

Jeon et al. (2006) [14] - Naver – Korean question and answering portal

Kernel density estimation (KDE) and the maximum entropy approach are used to handle different types of non-textual features and to build a stochastic process to predict the quality of an answer. 13 non-textual features are implemented to estimate answer quality.

4.1 Evaluation of Question & Answering Portals

CQA frameworks developed by [1] and [14] have been validated against question and answering portals. These frameworks are likely to be domain independent as content (questions and answers) are generated from a wide range of knowledge domains (e.g. teaching and learning, online gaming and so forth). However, it is also believed by [1] that their results and insights are applicable to other social media platforms. However, future research is required to validate their claim. Forums may be an applicable platform for validation as forums often contain questions and answers.

Both frameworks apply heuristics for CQA. Agichtein et al. (2008) evaluates multiple sources of quality evidence gathered through content analysis, usage statistics and user feedback to train a classifier that can identify high quality UGC [1]. Jeon et al. (2006) evaluates non-textual features to determine the quality of answers. Non-textual features can be termed as *metacontent* which is information that describes the content and could include features such as answer length and usage statistics rather than evaluating the actual content. A major benefit of evaluating non-textual features is that the CQA framework can avoid the complexity of linguistic content analysis in addition to being language independent. However such an approach may be open to fraud and spam without the use of semantic content analysis [14].

Additionally, these two frameworks employ both implicit (e.g. number of views for a Q&A) and explicit (e.g. number of votes received for a Q&A) user feedback in their CQA process. In essence, the evaluation of complimentary quality evidence sources, i.e. both types of user feedback can help capture high quality content from different perspectives and therefore assist in managing fraud [1, 14]

5 Peer-to-peer Applications

Cheng & Vassileva (2006) [7] - Comtella – student resource sharing system

A user contribution measurement model is developed to calculate the contribution made by students in an online resource sharing system. The average quality of user's contribution is defined as the average summative rating of all the resources s/he has shared so far. User ratings are used to assess the quality of contributions and users that rate other users' content are rewarded with c-points (a virtual currency). These c-points can be used to provide visibility for a user's contribution within search results.

Han & Liu (2006) [11] - simulation of P2P network

This paper identifies and handles a problem many other social media services face in which they make the assumption that user feedback is honest and represents the real of opinions of peers within the online community. This model therefore validates the quality of user feedback (e.g. user ratings) through what the authors have termed as the, Feedback Valid Degree (FVD), and the result of this metric is used to refine a user's overall reputation score.

5.1 Evaluation of Peer-to-peer Applications

Cheng & Vassileva (2006) developed and validated an automated CQA framework for an educational social media application named *Comtella* while Han & Liu (2006) tested their framework against a simulated peer-to-peer network. The quality of UGC within both frameworks is predominately assessed through user ratings of shared resources [7, 11].

To the best of our knowledge, human judgement has not been incorporated to validate the reliability of the CQA results for [7] and is not required for [11] as quality influencing variables have been preset within their simulated network.

Fraud has been thoughtfully considered and managed in the design of both frameworks and has been referred to [7] as designs to prevent gaming the system. However, in terms of evaluating the quality of user ratings there are possible avenues to manipulate the system. The principles behind the frameworks developed in [11] and [21] can be used to validate user feedback and determine the leniency of user feedback respectively. These principles can be applied to [7] and many other CQA frameworks in determining whether user feedback is genuine through validation.

Another potential fraud issue with [7] is that resources contributed at the start of the week are given higher weights than resources contributed towards the end of the week. Although, timeliness may be a vastly important factor for this specific domain, their framework becomes open to fraud if semantic content analysis techniques are not employed to assess the quality of the resources contributed early in the week. This however may be difficult in analysing non-textual content such as images and videos.

6 Review Portals

Lauw, Lim & Wang (2007) [21] - Epinions (video reviews)

A differential model of leniency and quality is developed to assess content quality based on ratings provided by reviewers. The model considers modelling biases by the notion of leniency, which captures the tendency of a review to give a higher or lower score. This approach may be applicable in compensating for the under and over-estimation of quality from content ratings given by social media users.

7 Weblogs

Ni et al. 2007 [27] - MSN Space (Chinese weblog articles)

A model is proposed and various machine learning techniques are evaluated by the authors in classifying blog articles into two categories namely, informative and affective blog articles. Informative articles have content which are similar to news websites, provide technical descriptions (e.g. programming techniques), common sense knowledge and/or present objective comments about events in the world. Affective articles however include personal diaries and contain articles relating to a blogger's feelings and emotions. The researchers of this article propose that informative articles represent higher quality and more desirable content when compared with affective articles.

Ulicny, Baclawski & Magnus (2007) [36] - PoliticalTheory.info weblogs

A prototype has been developed that employs a number of metrics to mine news blog articles. In terms of quality assessment, the most important metric proposed by the authors is credibility in which they put forward the assumption that authors of high quality blogs will have high credibility. 50 source, message and reception features for measuring a blog author's credibility is stated but not all of these features disclosed in their paper.

7.1 Evaluation of Weblogs

[27] proposed a model that was able to classify weblog articles into two categories namely, informative and affective weblog articles. Informative articles have content similar to news websites, provide technical descriptions (e.g. programming techniques), common sense knowledge and/or present objective comments about events in the world. Adversely, affective articles comprise of content similar to personal diaries and can contain information relating to the weblog author's feelings and emotions on certain subjects. This framework puts forward the assumption that informative articles represent higher quality articles than affective articles.

However, [36] developed a prototype that evaluates a number of metrics to mine news weblog articles. In terms of quality assessment, the most important metric proposed by the authors is credibility in which they put forward the assumption that authors of high quality weblogs will have high credibility. 50 source, message and

reception features for measuring a weblog author's credibility include but are not limited to number of posts an author writes per month, the ratio of original text to advertisement (a feature of weblog spam), number of original (non-quoted) sentences in a post, number of subscribers to a weblog such as *del.icio.us* and *Digg*.

Both of these frameworks are not founded on strong theories but rather heuristics. The framework proposed by [27] is considered domain independent as it only judges quality through the informative and affective criteria regardless of the article topic domain. However, the framework developed by [36] has been specifically designed towards news weblog articles but also evaluates user feedback through evaluating the number of trackbacks (links from other weblogs) a weblog possesses as an additional quality evidence source. Results from both of these frameworks have been evaluated through human judgement which strengthens the reliability of both prototypes.

8 Wikis

Katakula (2008) [16] - Simulated wiki

An employee performance measurement model was built to assess the quality of employee contributions within wikis. The result of this assessment was then used to categorise employees into 5 different grades for performance evaluation and rewards distribution. Their model evaluates 12 features attributing to the quantity and quality of UGC.

Hu et al. (2007) & Lim et al. (2006) [12, 23] - Wikipedia (articles of countries)

The Basic, PeerReview and ProbReview models are developed based upon what the authors have termed as the mutual reinforcement principle between quality and authority. This principle suggests that in order to measure quality, the contributor's authority must also be measured. The authority of contributors is determined by adopting ideas from the HITS link analysis algorithm.

8.1 Evaluation of Wikis

Hu et al. (2007) and Lim et al. (2006) proposed a CQA framework that mines *Wikipedia* article edit histories to determine the quality of its encyclopaedic articles based on an authority score calculated for article editors and the words they contribute to an article [12, 23]. [16] developed a CQA framework as a means to measure the quality of employee contributions within a wiki.

It is believed that these frameworks are only applicable to wikis that are built on *Wikipedia's* wiki platform, *Mediawiki*. As such, a number of the ideas presented in their frameworks can not be easily adopted for other social media types such as forums. These frameworks however are considered domain independent as they need to evaluate the quality of articles from many knowledge domains.

The CQA framework results generated from [12] was not validated through subjective analysis to ensure its reliability. For example, related research conducted in [23], ranked *Wikipedia* country articles in terms of their quality. Based on the experimental results of two proposed CQ frameworks, Benin was ranked as top

quality country article. However, the contents of the Benin article could have been evaluated through human judgment and compared against country articles that were ranked lower. These evaluations can ensure the reliability of the framework in addition to identifying areas for framework improvement.

No strong theories have been used in the development of either framework but heuristics have been largely adopted. For example, quality related factors such as the word length of an article, words that survive article revisions and the number of ratings received per page were evaluated for CQA.

Neither framework has adopted semantic content analysis in their CQA framework [12, 16]. [12] evaluated words within articles but have not extended their evaluation to capture concepts. Additionally, [12] has excluded the evaluation of article edits that fix syntactical errors such as correcting punctuation, typos and adding stop words. However, one might argue that these corrections do in fact impact quality as these fixes can improve the overall readability of an article.

Interestingly, [12] propose that in order to measure quality, the contributor's authority (reputation) must also be determined. Therefore, this framework assumes that a high authority author will produce high quality content. However, people who have high authority and knowledge may only possess that for a specific domain and this distinction has not but should be made in the CQA process. For example, a highly reputable computer scientist that contributes to a fashion related article should not necessarily carry over their authority score to an article on fashion. This framework enhancement could also allow for the discovery of knowledge or related knowledge domain experts.

No specific details are provided by [12] and [16] on how fraud is managed. However, wiki articles can be considered as less susceptible to fraud because such activity can often be policed and managed by other article editors. Other social media types such as forums and weblogs allow users to create their own content and users may therefore be more inclined to engage in fraudulent activities to improve the quality ranking of their content.

9 Not Validated Against an Application Domain

Chen & Xu (2005) [5] - Undergraduate and graduate student search results

Theory-driven approach used to develop a conceptual model of relevance based on Grice's theory of human communication. The relevance model consists of 5 factors which include topicality, reliability, understandability, novelty and scope. Relevancy, understandability and reliability are defined as a common IQ dimensions in existing literature and these metrics are applicable to quality assessment.

Inch, Warnick & Endres (2005) [13]

Argument analysis helps in identifying a student's point of view and supplies information not stated in the message. Message quality can be categorised into 4 categories using premise and claim as a classifier. These categories supplement the Levels of Disagreement and are used in evaluating quality. Two models are proposed which are the General and Toulmin models. To the best of our knowledge, these

models have not been realised into a working prototype. One difficulty of implementing this model is clearly distinguishing between what is a premise and a claim.

Nussbaum et al. (2002) [28]

This paper proposes a coding system to represent levels of disagreement. Participants that engage in online debates (preferably civil) are generally seen as productive. Therefore, one way of identifying quality discussions is to code messages according to their level of disagreement that is exhibited in relation to the referred posting. 4 levels of disagreement are proposed by the author. To the best of our knowledge this paper does not implement the model into a working prototype and thus has no experimental results.

Price & Shanks (2004) [29]

An IQ framework based on semiotic theory to describe the form, meaning and a use-related aspect of information is proposed. This framework provides a sound theoretical basis for defining quality categories and for integrating different approaches in deriving quality criteria for each category. To the best of our knowledge, this framework has not been implemented in a working prototype and tested on a social media platform.

9.1 Evaluation of Non-Validated Frameworks

To the best of our knowledge, these frameworks have not been validated against a social media domain [5, 13, 28, 29]. The future implementation and validation of these frameworks can provide potential insights for either improving existing frameworks or developing new CQA frameworks for social media.

10 Conclusion

This paper has surveyed different CQA related frameworks for social media. This has been conducted through evaluating common CQ dimensions, theories and heretics, domains and platforms in addition to presenting a critical assessment the CQA related frameworks over 8 evaluative criteria. Hopefully, this review has provided insights into existing CQA frameworks as well as to suggest areas of improvement for these frameworks. Our team is currently involved in developing a UCM for social media with CQA being a major component in our model.

References

1. Agichtein, E., Castillo, C., Donato, D., Gionis, A., Mishne, G.: Finding High-Quality Content in Social Media. In: Proceedings of the International Conference on Web Search and Web Data Mining, pp. 183-194 (2008)

2. Cappiello, C., Francalanci, C., Pernici, B.: Data Quality Assessment from the User's Perspective. In: Proceedings of the 2004 International Workshop on Information Quality in Information Systems, pp. 68-73 (2004)
3. Chai, K.: Development and Evaluation of User Contribution Scoring Models for Web-based Discussion Forums. In: Curtin University of Technology's Information Systems Honours Thesis (2007)
4. Chai, K., Potdar, V., Chang, E.: A Survey of Revenue Sharing Social Software's Systems. In: Proceedings of the International Workshop on Social Interaction and Mundane Technologies (2007)
5. Chen, Z., Xu, Y.: User-Oriented Relevance Judgement: A Conceptual Model. In: Proceedings of the 38th Hawaii International Conference on System Sciences, Track 4, pp. 101b (2005)
6. Cheng, R., Vassileva, J.: User Motivation and Persuasion Strategy for Peer-to-peer Communities. In: Proceedings of the 38th Annual Hawaii International Conference on System Sciences, pp. 193a (2005)
7. Cheng, R., Vassileva, J.: Design and Evaluation of an Adaptive Incentive Mechanism for Sustainable Online Educational Communities. In: Journal of User Modeling and User-Adapted Interaction, vol. 16, no. 3-4, pp. 321-348 (2006)
8. General Inquirer Categories (2008) (Retrieved March 19, 2008) from <http://www.wjh.harvard.edu/~inquirer/homecat.htm>
9. Goldberg, J.: Social Media's Population (2008) (Retrieved March 11, 2008) from http://www.socialmedian.com/2007/10/social_medias_population.html
10. Harper, M.F., Li, S.X., Chen, Y., Konstan, J.A.: Social Comparisons to Motivate Contributions to an Online Community. In: Proceedings of The Second International Conference on Persuasive Technology, pp. 148-159 (2007)
11. Han, J., Liu, Y.: Dubious Feedback: Fair or Not? In: Proceedings of the First International Conference on Scalable Information Systems, article no. 49 (2006)
12. Hu, M., Lim, E.P., Sun, A., Lauw, H.W., and Vuong, B.Q.: Measuring Article Quality in Wikipedia: Models and Evaluation. In: Proceedings of the Conference on Information and Knowledge Management, pp. 243-252 (2007)
13. Inch, E.S., Warnick, B., Endres, D.: Critical Thinking and Communication: The Use of Reason in Argument. Allyn & Bacon, Boston (2005)
14. Jeon, J., Croft, W.B., Lee, J.H., Park, S.: A Framework to Predict the Quality of Answers with Non-Textual Features. In: Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 228-235 (2006)
15. Kahn, B.K., Strong, D.M., Wang, R.Y.: Information Quality Benchmarks – Product and Service Performance. In: Communications of the ACM, vol. 45, no.4, pp. 184-192 (2002)
16. Katakula, J.: A Framework for Employee Performance Measurement in an Online Collaborative Environment like MediaWiki. In: Curtin University of Technology's Masters of Electronic Commerce Thesis (2008)
17. Katerattanakul, P., Siau, K.: Measuring Information Quality of Web Sites: Development of an Instrument. In: Proceedings of the 20th international conference on Information Systems, pp. 279-285 (1999)
18. Kim, J., Shaw, E., Feng, D., Beal, C., Hovy, E.: Modeling and Assessing Student Activities in On-Line Discussions. In: Proceedings of the AAAI Workshop on Educational Data Mining (2006)
19. Klamma, R., Chatti, M.A., Duval, E., Hummel, H., Hvannberg, E.T., Kravcik, M., Law, E., Naeve, A., Scott, P.: Social Software for Life-long Learning. In: Journal of Educational Technology & Society, vol. 10, no. 3, pp. 72-83 (2007)
20. Knight, S., Burn, J.: Developing a Framework for Assessing Information Quality on the World Wide Web. In: Informing Science Journal, vol. 8, pp. 159-172 (2005)

21. Lauw, H.W., Lim, E.P., Wang, K.: Summarizing Review Scores of "Unequal" Reviewers. In: Proceedings of Society of Industrial and Applied Mathematics International Conference on Data Mining, pp. 26-28 (2007).
22. Liew, J.: Top social networks for engagement - some surprises (2007) (Retrieved January 7, 2008) from <http://lsvp.wordpress.com/2007/07/23/top-social-networks-for-engagement-some-suprises/>
23. Lim, E.P., Vuong, B.Q., Lauw, H.W., Sun, A.: Measuring Qualities of Articles Contributed by Online Communities. In: Proceedings of the 2006 IEEE/WIC/ACM International Conference on Web Intelligence, pp. 81-87 (2007)
24. Lipsman, A.: comScore: Social Networking Goes Global (2007) (Retrieved January 7, 2008) from <http://www.comscore.com/press/release.asp?press=1555>
25. Lui, A.K.F., Li, S.C., Choy, S.O.: An Evaluation of Automatic Text Categorization in Online Discussion Analysis. In: Proceedings of the Seventh IEEE International Conference on Advanced Learning Technologies (ICALT), pp. 205-209 (2007)
26. McKlin, T., Harmon, S.W., Evans, M., Jones, M.G.: Cognitive Presence in Web-Based Learning: A Content Analysis of Students' Online Discussions (2002) (Retrieved March 20, 2008) from <http://it.coe.uga.edu/itforum/paper60/paper60.htm>
27. Ni, X., Xue, G.R., Ling, X., Yu, Y., Yang, Q.: Exploring in the Weblog Space by Detecting Informative and Affective Articles. In: Proceedings of the 16th International World Wide Web Conference, pp. 281-290 (2007)
28. Nussbaum, M.E., Hartley, K., Sinatra, G.M., Reynolds, R.E., Bendix, L.D.: Enhancing the Quality of On-Line Discussions'. In: Paper presentation at the annual meeting of the American Educational Research Association, New Orleans, LA (2002)
29. Price, R., Shanks, G.: A Semiotic Information Quality Framework. In: Proceedings of the 2004 IFIP International Conference on Decision Support Systems, pp. 658-672 (2004)
30. Price, R., Shanks, G.: A Semiotic Information Quality Framework: Development and Comparative Analysis. In: Journal of Information Technology, vol. 20, no. 2, pp. 88-102 (2005)
31. Rafaeli, S., Ravid, G., Soroka, V.: De-lurking in Virtual Communities: A Social Communication Network Approach to Measuring the Effects of Social and Cultural Capital. In: Proceedings of the 37th Hawaii International Conference on System Sciences, Track 7, pp. 70203 (2004)
32. Schonfeld, E.: Social Site Rankings (2007) (Retrieved January 7, 2008) from <http://www.techcrunch.com/2007/11/28/social-site-rankings-october-2007/>
33. Social Media (2008) (Retrieved March 10, 2008) from http://en.wikipedia.org/wiki/Social_media
34. Stvilia, B., Twidale, M.B., Smith, L.C., Gasser, L.: Information Quality Work Organization in Wikipedia. In: Journal of the American Society for Information Science and Technology, vol. 59, no. 6, pp. 983-1001 (2008)
35. Strong, D., Lee, Y., Wang, R.: Data Quality in context. In: Communications of the ACM, vol. 40, no. 5, pp. 103-110 (1997)
36. Ulicny, B., Baclawski, K. & Magnus, A.: New Metrics for Blog Mining. In: Proceedings of SPIE Defense & Security Symposium, vol. 6570, no. 65700I (2007).
37. Wang, R.Y., Strong, D.M.: Beyond accuracy: What Data Quality Means to Data Consumers. In: Journal of Management Information Systems, vol. 12, no. 4, pp. 5-33 (1996)
38. Wiertz, C., Ruyter, K.D.: Beyond the Call of Duty: Why Customers Contribute to Firm-hosted Commercial Online Communities. In: Organization Studies, vol. 28, no. 3, pp. 347-376 (2007)