Quality Ranking Algorithms for Knowledge Objects in Knowledge Management Systems

Amal Al-Rasheed
Information Systems Department
Princess Nourah Bint Abdulrahman Univesity (PNU)
Riyadh, Saudi Arabia

Jawad Berri
Information Systems Department
King Saud University (KSU)
Riyadh, Saudi Arabia

Abstract—The emergence of web-based Knowledge Management Systems (KMS) has raised several concerns about the quality of Knowledge Objects (KO), which are the building blocks of knowledge expertise. Web-based KMSs offer large knowledge repositories with millions of resources added by experts or uploaded by users, and their content must be assessed for accuracy and relevance. To improve the efficiency of ranking KOs, two models are proposed for KO evaluation. Both models are based on user interactions and exploit user reputation as an important factor in quality estimation. For the purpose of evaluating the performance of the two proposed models, the algorithms were implemented and incorporated in a KMS. The results of the experiment indicate that the two models are comparable in accuracy, and that the algorithms can be integrated in the search engine of a KMS to estimate the quality of KOs and accordingly rank the results of user searches.

Keywords—Knowledge Management System (KMS); Knowledge Object (KO); knowledge evaluation; quality indicator; recommender system

I. INTRODUCTION

The ever-increasing volume and diversity of knowledge in Knowledge Management Systems (KMSs) has required users to spend more time searching for the information they need. Searches of such knowledge repositories often yield a large number of results, making it difficult for users to choose items that will actually meets their requirements [1]-[3]. Ranking of knowledge objects (KOs) in search results is based on measurement of the degree of similarity between the query submitted by the user and topics in the knowledge repository, regardless of any consideration of quality [4]. Without an evaluation process that can determine the relevance significance and quality of KOs, most searches will be weak and of limited benefit [5]. For that reason, some knowledge bases have resorted to the use of expert evaluations. Although these are efficient, they necessarily encompass only a limited number of KOs because of the limited number of experts and the tediousness of manual evaluation [6]. Moreover, these evaluations are implemented individually, which limits their validity, owing to bias and differences of opinion. As a consequence, when searching KOs, resources that have not been evaluated will appear at the bottom of the list of search results, regardless of their actual quality.

To overcome the problem of the large number of KOs that remain unevaluated, there is a need to measure their quality automatically. To alleviate the problem of unbiased evaluations, the evaluation process needs to be based on collaboration, through which participants converge on more accurate evaluations [7]. To date, few studies have focused on the issue of quality of information in virtual communities. As members of these communities have more freedom to add new KOs, the quality of available knowledge tends to be lower than in knowledge repositories in organizations [8]. This paper proposes a general framework for quality evaluation that includes two models of quality measurement. The first of these recommends KOs for online communities on the basis of quality indicators that are grouped into four dimensions. The second model is based on estimates of user reputation and exploits a content-based recommender technique for estimating the quality of KOs.

The remaining of this paper is structured as follows. Next section provides a review of research work in the field. Section 3 proposes two models for KO evaluation. Section 4 explains the data set, evaluation metrics, evaluation procedure and the result of the experiment. Then, Section 5 illustrates the application of the models through a case study. Finally, last section concludes this paper and presents future work that can be done.

II. BACKGROUND AND RELATED WORK

A. User Reputation

According to the American Heritage dictionary, reputation is "the general opinion or judgment of the public about a person or thing" [9]. Reputation scores are utilized to motivate users to actively participate [10]. To enable users to comprehend them easily, these scores are generally simple and count-based. In practice, reputable users are considered to be among the most important assets of websites [11]. The present study focuses on user reputation in an online collaborative KMS, where users contribute, share, and rate knowledge. Online collaboration has become an important means of creating and organizing knowledge, but the approach presents challenges for both content creation and content use [12]. The process of content creation is open to abuse, and content consumers may have difficulty in distinguishing between high- and low-quality content. Reputation systems can help to prevent abuse and bring order to indications of content quality. One of the main objectives of a reputation system for collaborative content is to provide indications of content quality to users [12]. Reputation scores are computed

according to the quality and quantity of contributions made by individual users [10].

B. Knowledge Evaluation

The process of evaluating knowledge quality is difficult and complicated because multiple aspects must be considered. To the best of our knowledge, few existing studies have focused on the automatic evaluation of the quality of KOs; in this section, we review those of immediate relevance. One important evaluation and ranking algorithm is EigenRumor, which uses link analysis to calculate scores for community contributions, based on links from contributors to information objects. These scores can be used to classify information and contributors and are used as incentives to stimulate ongoing contributions to the community [13]. In [6], another evaluation algorithm assesses the value of knowledge and contributors according to common contribution actions, in which the dissimilar evaluation capabilities of contributors are reflected in the weightings of evaluated items. In [14], a general model is advanced for the automatic calculation of reputation scores based on the ratings given to knowledge resources, integrating these reputation scores to create value-added information about the rated resources. Other studies in related fields regarding evaluation of online resources are also of relevance here; for instance, in the context of e-learning, Ochoa and Duval proposed a number of quality metrics for the automatic evaluation of metadata characteristics [15]. In a subsequent study [16], the same authors proposed another set of metrics for deriving measures of personal, topical, and situational relevance [17].

III. Proposed Evaluation Models

A. Quality Indicators-Based Evaluation Model (QIEM)

The first proposed model captures four dimensions of quality to recommend knowledge objects to users. To our knowledge, this model is the first to consider contributor reputation as a measure of quality to enhance evaluation of KOs. We propose that reputation score can enhance quality and alleviate the effects of shortages of ratings or usage that undermine other models. The aggregation of all the scores from previous indicators support computing a general rating for a specific KO.

1) Quality indicators: Quality indicators are "statistical measures that give an indication of output quality;" the quality of outputs is best defined in terms of how well outputs address user needs, or whether they fit the user's purpose [18]. Given the increasing importance of properly defined quality indicators in the knowledge management field, academics from Loughborough University developed and facilitated a workshop entitled "The Use of Indicators for Monitoring and Evaluation of Knowledge Management and Knowledge Brokering in International Development," in association with the Institute for Development Studies. The workshop brought together thirty knowledge researchers and academics from twenty organisations to discuss quality indicators, and a resource pool of 100 such indicators are presented in the workshop report [19]. Of these, the indicators for quality of knowledge in virtual community include the following:

- Number of created knowledge objects
- Percentage of users who rate knowledge objects
- Number of citations of knowledge objects
- Number of downloads of knowledge objects
- Number of views of knowledge objects
- Percentage of readers of knowledge objects
- Number of items of relevance to one's work
- Number of channels that provide a knowledge object
- Availability of discussions of knowledge objects
- Number of recommendations of knowledge objects
- Usefulness of knowledge objects as perceived by target audience (5-point Likert scale)
- Number of examples where work has been cited.

It is noteworthy that while the majority of indicators are quantitative measures of how much or how many, relatively few are qualitative indices of how or why. In practice, indicators gain in strength when used as part of a basket of indicators [19]. For that reason, we will use some indicators from the above list that meet our requirements here and benefit from other indicators in adjacent fields, integrating them in a quantitative measure that serves to clarify the level of KO quality in online communities.

2) Quality indicators for KO evaluation: In evaluating the quality of KOs, four major dimensions will be taken into account: social, usage, characteristic, and contributor (Fig. 1).

a) Indicators of social quality

Social indicators are metrics that track users' explicit feedback and interaction [20]. Common forms of explicit feedback include comments, vote up, and vote down, as well as star-based ratings [21]. Previous examples of this kind give a good indication of the quality of content [22]. Explicit rating is the most effective way of capturing the user's judgments of KOs because it reflects the user's own evaluation of importance and quality. However, explicit feedback requires users to perform extra rating actions, which users may find inconvenient. Arising from this, the main limitations of explicit rating are that (i) new KOs that have not been evaluated appear at the bottom of the list of search results, as if their quality is low; and (ii) users have little interest in rating content because of a lack of incentives. To mitigate these limitations, we assign a neutral rating (3 out of 5) to new KOs until they attract user evaluations. Moreover, according to their interaction with the application, we award users extra points to encourage them to evaluate KOs.
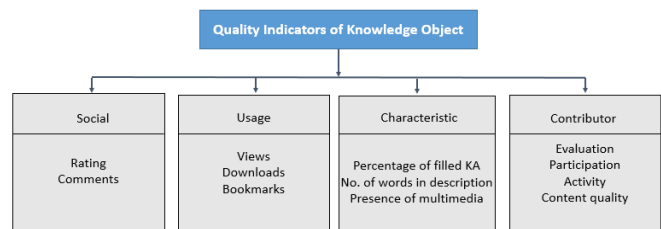

Fig. 1. Proposed indicators for knowledge object evaluation.

b) Indicators of usage quality

Usage indicators are metrics that track users' implicit feedback and behaviors. Claypool et al. reported that implied data acquired from user behaviour is very effective for sorting lists of search results [23]. As implicit feedback is based on search behaviour, there are many possible sources of such information. A number of studies have looked at the classification of possible sources of implicit feedback [23]-[26]. Implicit feedback systems commonly use such measures as document reading time, interaction, and scrolling, as these measures reflect the concerns of users and their satisfaction level and cost less than explicit evaluation [23], [27]. However, these systems are built on the assumption that relevant documents will be viewed and interacted with more than those that are less relevant.

In the particular case of knowledge bases. Data such as popularity, number of views, or number of bookmarks can be utilized to complement information on the quality of knowledge objects.

c) Indicators of characteristic quality

The characteristic dimension includes quality indicators based on the capability of information to describe a knowledge attribute. Some quality characteristics proposed in [28] include accuracy, provenance, completeness, consistency and coherence, timeliness and accessibility, and conformance to expectations. For instance, to evaluate the completeness of a knowledge attributes record (quality characteristic), we can check how many attributes have been filled with information (metric).

It is important to note that these metrics relate to the quality of knowledge attributes but not to the quality of the KO itself. The completeness metric has been selected for present purposes, as it is convenient to implement for the available information in real knowledge repositories. Completeness is the extent to which knowledge attributes contain all the information needed to provide a complete representation of the labelled resource.

When all the attributes have non-null values, the value of this metric will be 1 (maximum value); in cases of an empty resource, the value will be 0 (minimum value). While this seems straightforward, not all knowledge attributes are equally relevant for all resources. For this reason, it is better to use a weighting factor to illustrate the significance of the attribute.
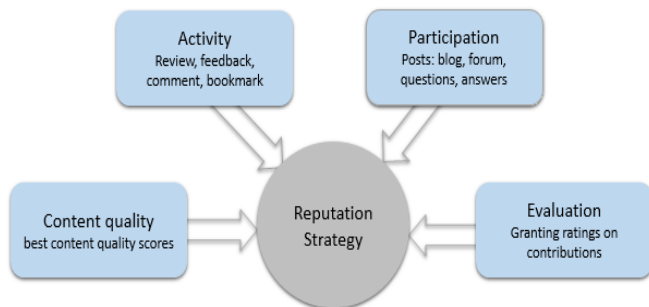
To measure the characteristic dimension, we define indicators based on the more general indicators discussed in the preceding section. In addition to the number of filled attributes, we added the number of words in the description attribute and the presence of multimedia.

d) Indicators of contributor quality

Reputation is defined by Alfarez et al. as "an expectation about an agent's behaviour based on information about or observation of its past behaviour" [29]. A number of previous studies have demonstrated that user reputation is a good indicator of the reliability and quality of content [30]. Within an online community, users can build their reputation through condensing their activities. User reputation score can be calculated by reference to features that denote the user's authority and influence within the community [31]. In the present case, we have identified four features to determine user reputation score (Fig. 2): contributor's evaluation, participation, activity, and content quality. During the evaluation process, there is a clear association between a contributor's reputation score and knowledge value score, as contributors with high reputation scores are likely to supply the knowledge repository with high quality KOs [32]. The four features used to estimate user's reputation can be described as follows:

- Evaluation: The average of all ratings of the contributor for their contributions

- Participation: The quantity of contributions made by the user, weighted according to type (article, blog, question, answer, forum post)

- User activity: The weighted average of three metrics: user ratings, user bookmarks, and user comments

- Content quality: The average of the contributor's best content quality scores

3) Evaluation model: Fig. 3 illustrates inputs and outputs of the evaluation system. Using collected data for predefined quality indicators and statistical measures, the quality manager assigns a score to each search result. First, values are individually calculated for each of the four indicators. These values are then added together, using their respective (and configurable) weighting percentages. After that, they are normalized using a scale between 0 and 1 to adjust for the disparity in the values measured on different scales. Finally, quality manager automatically re-rank the search result according to the calculated scores.
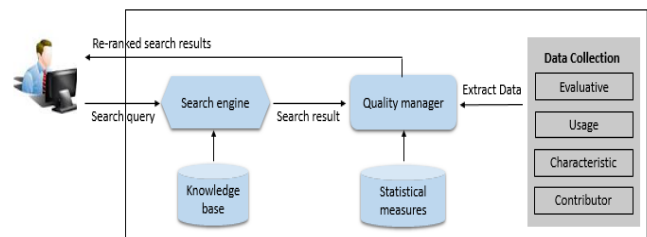

Fig. 2.   Reputation strategy.

Fig. 3.   Evaluation system.

The resulting score is an outline of numerical values assigned to all predefined quality indicators [14]. Measurement of the estimated quality score entails the following steps:

Step 1: Normalize indicators. The original data for all indicators should be normalized to eliminate the influence of any dimension over the others. The normalized value of e_i for indicator E is calculated as follows:

$$Normalized\ (e_i) = \frac{e_i - E_{min}}{E_{max} - E_{min}} \tag{1}$$

where: E_{min} = minimum value for indicator E and E_{max} = maximum value for indicator E If E_{max} is equal to E_{min} then Normalized (e_i) is set to 0.5.

Step 2: Weight indicators. To measure the effect of indicators in an evaluation system, assigning weights to all indicators is an essential procedure. An indicator whose weight is high will exert a greater effect on overall quality; otherwise, its effect is lower. In statistical terms, because standard deviation measures the distribution of numbers, it is one of the best weighting methods. The basic principle of standard deviation is that when the data of one indicator present large differences among multiple evaluated objects, the standard deviation of this indicator must be high [33]. The value of an indicator's standard deviation is directly proportional to its contribution to the integrated formula of all indicators.

Standard deviation determines the weights of indicators by means of the following equations:

$$w_j = {\sigma_j}\Big/{\sum_{j=1}^{n} \sigma_j} \qquad j = 1, \dots, n \tag{2}$$

$$\sigma_j = \sqrt{\frac{\sum_{i=1}^{m}(x_{ij}-x_j)^2}{m}} \qquad j1, \dots, n \tag{3}$$

where w_j is the weight of a criteria and σ_j is the standard deviation. Table I sets out the weight assigned to each quality indicator using this method.

Step 3: Integrate indicators. Now, the information about quality indicators can be integrated in one score, where each indicator makes a measured contribution to overall quality. The estimated quality score combines all quantitative information about quality indicators of a KO, which means that if a quality indicator does not exist, the estimated score can be calculated automatically from the existing indicators. Moreover, the new KOs will be assigned a neutral rating (3 out of 5) to resolve the critical problem of new KOs without ratings appearing at the end of the list of search results, increasing the reliability of recommendations. Estimation of the quality score of a knowledge resource is described in (4):

TABLE I.   STANDARD DEVIATION VALUES AND WEIGHTS OF QUALITY DIMENSIONS

	Social	Usage	Characteristic	Contributor
SD	0.144822	0.299779	0.096784	0.07456
Weighted SD	0.235122	0.486698	0.157132	0.121049

$$Score = \sum_{i=1}^{m} a_i * Social + \sum_{j=1}^{n} b_j * Usage +$$
$$\sum_{k=1}^{o} c_k * \ Characteristic + \sum_{l=1}^{p} d_l * Participant \tag{4}$$

where a, b, c, and d represent the respective weights of social, usage, characteristic, and participant indicators, and m, n, o, and p represent the indicator number in each quality dimension. Where any of these data are missing, the weights are adapted to compensate for this absence in calculating estimated quality. In addition, all indicators are normalized by scaling between 0 and 1 as described in (2), and their mean values are included in the final score.

B. Reputation-Based Evaluation Model (REM)

This second model assesses the quality of knowledge objects automatically by exploiting the capabilities of recommender systems and user reputation scores. The proposed model is based on the concept of recommending KOs that are similar in content and specifying KO quality on the basis of ratings posted by reputable users to help other users to select the best KOs. This model is based on the premise that users with high reputation points are more reliable in evaluating KOs. Intuitively, as reputable users can be expected to submit high quality contributions and to attract high ratings from the user community, users can benefit from reputation scores identifying good contributions [10].

1) Evaluation model and algorithm: Recommending KOs involves two phases. In the first phase, relevant KOs are retrieved according to the keywords entered by the user. The second phase re-ranks search results according to the estimated quality score for each KO.

Algorithm 1 describes the reputation-based evaluation approach.

Algorithm 1. *Reputation-based evaluation*

Input: *S* = vector of *KOs*, as returned by search engine S
　　　　UP = users' profiles
　　　　R = ratings of *KOs*
Output: Vector of *KOs*, re-ranked according to *SCORE*

For each *KO* $\in S$
　　Set *n = 0*
　　For each rater *I*
　　　　Set *E* = the average of all ratings the rater *I* gains
　　　　Set *P* = the number of contributions made by the rater *I*
　　　　Set *A* = the weighted average of: user ratings, user bookmarks, and user comments
　　　　Set *Q* = the average of best content scores the rater *I* gains
　　　　REP = (E + P + A + Q)
　　　　If *REP > threshold* **Then**
　　　　　　n = n++
　　End For
　　If *n >= 3*
　　　　SCORE = R of reputable users */ n*
　　Else
　　　　Calculate cosine similarity between the *KO* and *KOs* in database
　　　　Set *X* = most similar *KO*
　　　　PRED = R of reputable users on *X*
　　　　SIM = cosine similarity between the user and the *KO* creator
　　　　SCORE = PRED+ SIM/2
　　End If
End For
Sort *S* based on *SCORE* in descending order

The evaluation process begins by retrieving ratings of the KO by reputable users. In the absence of sufficient ratings, the recommendation process will start by measuring the similarity of the KO in the search result and other KOs in the database, based on their attributes. The most similar KOs exceeding the threshold of similarity will be utilized to predict the ratings of reputable users. Where more than one KO has the same similarity value, will the result will be improved by calculating the similarity between the user and the creator of the KO. Below, we describe the REM in detail, using the content-based collaborative recommender and user reputation.

Step 1: Find ratings. To begin, historical ratings of reputable users are used to estimate the quality score that will assist the user in selecting high quality KOs. The quality estimation strategy begins by scanning the search results. If the KO returned by the search has attracted some ratings, the system will interrogate the repository of ratings for the reputation points of users who have rated that KO. In calculating the average rating for a particular KO, the system will consider only reputable users' ratings. Equation (5) specifies how the average rating is calculated.

$$Q_j = \frac{\sum_{i=1}^{n} r_{i,j}}{N_j} \qquad (5)$$

where $r_{i,j}$ is the rating of $user_i$ of KO_j, N_j is the total number of reputable users who rated KO_j, and Q_j is the estimated quality score.

Step 2: Retrieve the most similar KOs. If no ratings of users with high reputation points can be found for a particular KO, the system will calculate predicted ratings, using the content-based filtering algorithm to calculate similarity and to make quality predictions. The predicted rating is computed on the basis of (i) the similarity between the characteristics of the KO in the search result and other KOs in the database and (ii) neighbors of the user whose profiles reveal similar characteristics. In determining the user's neighbors and finding similar KOs, we use the cosine similarity measure to calculate both user-user similarity and item-item similarity [34]. Finally, the overall quality estimate is the linear combination of predicted rating and user similarity. Prediction of a KO's quality is then computed by performing a linear weighted average.

A vector space model [35] is used to represent a KO as a vector of attributes. Weight is then calculated and included in the vector. In the vector, attribute value is (1) for presence or (0) for absence (0) of a term; binary weights are utilized to compute similarity between two KOs. Following the weighting of knowledge attributes, the similarity between two KOs can be computed using the following cosine similarity formula:

$$similarity = \cos(\vec{A}, \vec{B}) = \frac{\vec{A} \cdot \vec{B}}{\|\vec{A}\| \|\vec{B}\|} = \frac{\sum_{i=1}^{n}(A_i * B_i)}{\sqrt{\sum_{i=1}^{n} A_i^2} \sqrt{\sum_{i=1}^{n} B_i^2}} \qquad (5)$$

where $\vec{A}$ and $\vec{B}$ are treated as vectors of attributes of KOs and $\|\vec{A}\|$ and $\|\vec{B}\|$ are the magnitudes of vectors $\vec{A}$ and $\vec{B}$. The angle between the two vectors indicates their degree of similarity; a smaller angle signifies greater similarity.

For more accurate results in the absence of similar KOs with sufficient ratings, users' neighbors with similar characteristics can be identified on the basis of their profiles. In the same way, the system calculates the similarity of user and creator of a KO in the search result on the basis of their user profiles.

Step 2: Predict the quality score. The predicted quality score is computed using the similarity value for the user and creator of the KO and average ratings of users with high reputation points on the most similar KO. The equation is defined as follows:

$$Q_s \quad = \left(\frac{\sum_{i=1}^{n} r_{i,s}}{N_s} + Sim_{u,c}/2 \right) \qquad (7)$$

where $r_{i,s}$ is the rating of $ruser_i$ with high reputation points on the most similar KO, and $Sim_{u,c}$ is the similarity between the user who performs the search and the creator of KO.

Step 4: Re-rank the search results. Search results can now be re-ranked. Each KO in the search results is assigned a score, representing the estimated quality score in helping users to find valuable KOs, calculated either by (5) or (7), according to the availability of ratings. Search results are ranked in descending order from the highest score to the lowest for presentation to the user as the list of recommended KOs.

IV. EXPERIMENTS

In experiments to examine the performance of our proposed quality recommendation models, the following were the main objectives: 1) to assess the effectiveness of the proposed quality recommendation models and 2) to evaluate the accuracy of the proposed models in order to select an appropriate model for adoption in our KMS. The results are presented and discussed below, following a description of the data sets.

A. Data Sets

One of the challenges facing the implementation of KMS is to find appropriate data sets for experimentation. Although there are many available sources of data, most of these have not been defined or documented. Most KM websites allow users to see part of the data but do not generally offer open data sets or provide evaluation of KOs. Additionally, no data set contains all the quality indicators under consideration here. The two available options for constituting the data set, then, were to use a real data set that might (imperfectly) match the characteristics of the target domain and task, or to synthesize a data set specifically to match the required properties. The proposed models entail data sets that contain information about users and their action types, reputation scores, and knowledge resources, as well as explicit and implicit ratings. Fortunately, our search identified a website (www.teachability.com) containing some indicators from the four quality dimensions mentioned above, enabling us to run the experiments using a real data set, The website in question is a collaborative online sharing space for teachers, enabling them to connect, learn, and improve their capabilities.

The data set contained 217 resources and 58 users. The resources gathered were available on the website in the period between May 2011 and November 2015. The data set contained some quality indicators from each of the dimensions (social, usage, characteristic, and contributor). Each resource had a title, description, keywords, and information about the creator. User actions recorded on the Teachability website included accessing learning resources, bookmarking resources, adding a comment, adding a rating, and accessing user pages. These actions provided useful implicit and explicit knowledge about the quality of the resource. Teachability awards points to users for their actions, which is a pivotal factor for present purposes. To examine the proposed models, we could not use all the quality indicators as they were not found in a unique data set. Instead, we used the available indicators, which included KO description, author, reputation points, views, bookmarks, ratings, and comments).

B. Evaluation Metrics

In respect of the proposed models, we were interested in ordering the list of search results according to estimated quality. This process is usually referred to as the ranking of items, and the appropriate order of a set of search results can be determined using a reference ranking [36]. A reference of this kind is essential in order to evaluate a ranking algorithm. In the case of the proposed models, where such a reference does not exist, it may be appropriate to compose a reference ranking by asking an expert to order the search results according to estimated quality.

1) Spearman's rank correlation coefficient

Rank correlation measures such as Spearman's ρ can be used as a reliable and fairly simple method of testing both the strength and direction (positive or negative) of any correlation between two variables [37]. Spearman's ρ also takes account of problems with ties. The relevant equation is:

$$\rho = 1 - \frac{6 \sum_{i=1}^{n} d_i^2}{n \, (n^2 - 1)} \tag{8}$$

where d_i is the difference between ranking of the reference and ranking of object i, $i \in \{1, 2,..., n\}$. Spearman's ρ is normalized in the interval [−1, 1] (see Table II). When both rankings are identical, $\rho = 1$; while in case one ranking is opposite in order to the other, $\rho = -1$ [38].

TABLE II.   INTERPRETATION OF VALUES OF CORRELATION COEFFICIENT

Correlation coefficient	Dependence between variables
1	absolute
0.9 - 1	very high
0.7 - 0.9	high
0.4 - 0.7	medium
0.2 - 0.4	low
0 - 0.2	very low
0	none

To test whether a perceived value of ρ is significantly different from zero, the t-test is among the most commonly used approaches [39], where

$$t = 1 - \frac{\rho}{\sqrt{\frac{1 - \rho^2}{(n-2)}}} \tag{9}$$

2) Kendall's coefficient of concordance (W)

Where there are more than two rankings of the same domain, Kendall's coefficient of concordance (Kendall's W) can be used to assess agreement between them. This coefficient ranges in value from 0 to 1, where 0 denotes no agreement and 1 denotes complete agreement. Kendall's W is given by [38]:

$$W = \frac{\sum_{i=1}^{n} x_i^2 - \frac{(\sum_{i=1}^{n} x_i)^2}{n}}{\frac{1}{12} \cdot k^2 \cdot (n^3 - n)} \tag{10}$$

where x_i is the sum of ranks for object i, k is the number of rankings, and n is the number of objects. The statistical significance of Kendall's W can be assessed using a χ^2 test with n − 1 degrees of freedom [38]:

$$\chi^2 = W \times k \times (n - 1) \tag{11}$$

C. Experimental Procedure

Users searched for KOs using the Solr search engine. Solr is "an open source enterprise search platform, runs as a standalone full-text search server" [40]. It uses the Lucene Java search library for text indexing and searching. Solr supports advanced customization, using plugin architecture [40]. To query a specific domain, users must first enter keywords. In general, the search results from a query are ranked according to the degree of similarity between keywords. For the purposes of the experiment, we first searched the database using the phrase "technology in teaching." In response to that query, the system retrieved 11 resources. Users can choose how they want search results to be ordered; one of the options is "recommended," which ranks search results in descending order according to the estimated quality score. To begin, we ranked results according to the first model and recoded the order and scores. The same procedure was followed for the second model. A teacher was asked to rank the search results in descending order on the basis of his teaching experience. The ranking strategy should prioritize those KOs that are most valuable to the user. After comparing each approach with the expert list, we compared the two approaches and measured the significance level of scores for each.

D. Experimental Results

1) Spearman's rank correlation coefficient for QUIM and REM

To evaluate the effectiveness of QIEM and REM, we recorded the system's ranking order and the score assigned to each search result for comparison with the reference rank list (Table III).

TABLE IV.   SEARCH RESULT RANK AND SCORE USING QIEM AND REM

KO Title	Expert rank	QIEM rank	QIEM Score	REM rank	REM Score
YouTube Launches Site Specifically for Teachers	1	5	2.222	2	3.868
Differentiated Instruction with Technology	2	1	2.714	1	4.000
Pocket Genius Teachers Guide	3	4	2.268	7	3.600
The Power of Documentation	4	3	2.506	5.5	3.668
How can I teach this student?	5	2	2.511	3.5	3.800
Setting expectations in the 21st century	6	8	1.775	10	1.333
Impacts of the Digital Ocean on Education	7	6	1.824	3.5	3.800
ELL Technology Integration and Tips	8	7	1.801	8	3.120
Education's Guide to Mobile Learning Devices	9	9	1.736	5.5	3.668
Airboat Lesson Activity	10	10	1.545	9	2.624
WW technology	11	11	1.536	11	1.226

TABLE V.   SPEARMAN'S COEFFICIENT FOR THE PROPOSED MODELS

Model	Spearman's coefficient	n	Significance level	Critical value	t
QIEM	0.85	11	0.05	0.5273	4.75
REM	0.71	11	0.05	0.5273	3.01

Spearman's ρ ranking coefficient was used to measure similarity between the expert ranking and the system ranking for both models (Table IV).

As noted from Table III, there is a high level of similarity between the rankings of both models and the expert rankings, indicating that the models agree with the expert in most cases.

The next step was to test whether this agreement was accidental. The null hypothesis (H0: "Agreement between both rankings is accidental") can be tested using t-values for both models [41]. For n = 11 and significance level = 0.05, the critical value is 0.5273 [42]. From (11), t = 4.75 for QIEM, and for REM, t = 3.01. As this exceeds the critical value for both models, the null hypothesis was rejected, indicating that agreement between the rankings was statistically significant (i.e., not accidental).

2) Kendall's coefficient (W) for QIEM, REM, and expert rankings

Table V summarizes agreement among QIEM, REM, and expert rankings.

Sum of x_i = 198; sum of x_i^2 = 4402; k = 3; n = 11. From (10):

$$W = \frac{4402 - \frac{39204}{11}}{\frac{1}{12} \cdot 3^2 \cdot (11^3 - 11)} = \frac{838}{990} = 0.846 \qquad (12)$$

This result indicates very high agreement between rankings.

TABLE VI.   EXPERT AND SYSTEM RANKINGS

KO Title	Expert rank	QIEM model	REM model	Sum of x_i	Sum of x_i^2
YouTube Launches Site Specifically for Teachers	1	5	2	8	64
Differentiated Instruction with Technology	2	1	1	4	16
Pocket Genius Teacher's Guide	3	4	7	14	196
The Power of Documentation	4	3	5.5	12.5	156.25
How can I teach this student?	5	2	3.5	10.5	110.25
Setting expectations in the 21st century	6	8	10	24	576
Impacts of the Digital Ocean on Education	7	6	3.5	16.5	272.25
ELL Technology Integration and Tips	8	7	8	23	529
Education's Guide to Mobile Learning Devices	9	9	5.5	23.5	552.25
Airboat Lesson Activity	10	10	9	29	841
WW technology	11	11	11	33	1089

E. Discussion

The main purpose of the experiment was to measure and contrast the efficiency of the suggested quantitative evaluation models in augmenting automatic knowledge sharing and dissemination services in a KMS. The two proposed evaluation models were QIEM (quality indicators-based evaluation model) and REM (reputation-based evaluation model). Fig. 4 compares the ranking performance of both models against the expert ranking, showing that the results from both models align significantly with the expert ranking.

To further assess performance, the rank correlation measure for each model was calculated, in which a higher positive value indicates a more effective model. Table VI summarizes rank correlation values for the proposed models.

Table VI indicates that both models provide highly accurate quality estimation of KOs. However, the results suggest that the QIEM provides higher accuracy and outperforms the REM. Although the performance of the REM model is 0.71, error arose from insufficient ratings of KOs. Because of the novelty of the system, the resources have not yet gained enough ratings. As the model searches for similar KOs with sufficient ratings, those with more ratings may have a lower similarity score. In addition, users can choose whether to provide their information, resulting in lower similarity scores when comparing the searcher to a set of incomplete users' profiles in the database. Further investigation revealed that the QIEM's superior performance is accounted for by the use of diversity indicators that make up for the absence of other data. Quality estimation can be roughly predicted using only a set of interactions with KOs and characteristics of KOs. However, there is no guarantee that users who interact with the same set of KOs will always return a similar ranking.
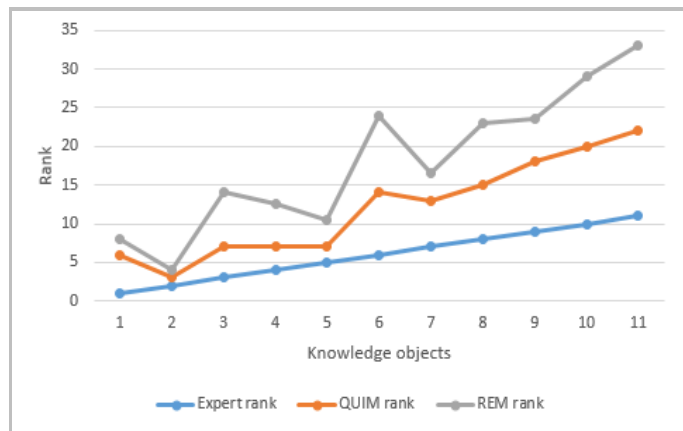

Fig. 4.   Comparison of the two methods of re-ranking with expert ranking.

TABLE VII.   EVALUATION OF THE PROPOSED MODELS

Model	Correlation coefficient	Strength of the correlation
QIEM	0.85	Very strong
REM	0.71	strong

V.   APPLICATION OF THE EVALUATION MODELS

In developing a KMS to support participation in knowledge sharing among university instructors, two problems were encountered. The basic problem in designing the system was how to assess the usefulness of a given KO. In addition, some mechanism was needed to encourage the reuse of knowledge. To address these issues, previous algorithms for recommending KOs were further developed, exploiting reputation scores previously assigned to each user to assess KO quality and encouraging user involvement by awarding extra points for interactions. The KMS was implemented as a knowledge portal using the Drupal content management system. The portal runs on a platform that supports Apache, PHP, and MySQL to store content and settings. The knowledge portal maintains a dynamic graphical user interface running on the client side that handles all user requests and collaborative activities. It facilitates knowledge acquisition, storing, and sharing, enabling users to submit documents, share ideas, work collaboratively, and store knowledge in searchable repositories. Fig. 5 shows the knowledge portal homepage.

To resolve the issue of identifying valuable knowledge, the proposed system integrates QIEM to provide for the automatic evaluation of knowledge, assessing its quality, recommending the qualified experience in terms of various measures, providing a quantitative score for overall rating of knowledge objects, and re-ranking search results based on the quality score.

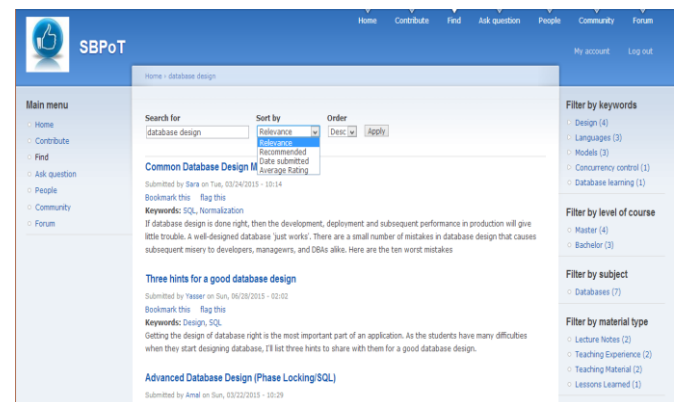Fig. 5.   Knowledge portal homepage.

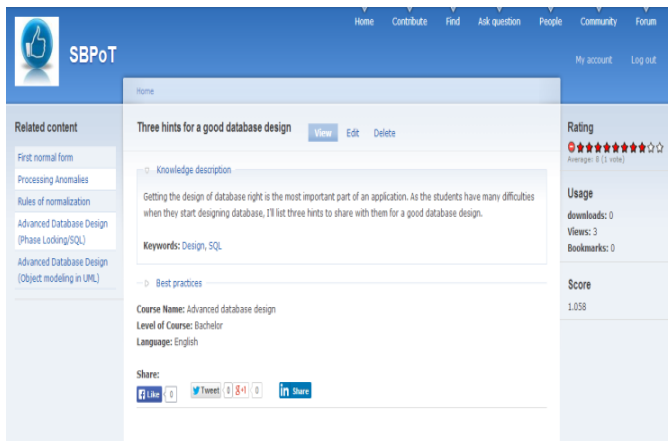Fig. 6.   Ranked list of search results.

Fig. 7. Detailed resource information.

Members can log into the system and search any chosen topic. Once submitted, the query is forwarded to the search engine and database, and search results are compiled and presented. The user receives a wide-ranging set of search results of different types, presented as a ranked list. Additionally, the user can select the ranking method (by relevance, recommended, date submitted, and average rating). Fig. 6 shows an example of a list of matching objects, ranked to help users to find the most valuable KOs.

When the user selects one or several KOs, they can then criticize or rate them to offer the community explicit feedback. Statistical information about users' views, downloads, shares, and bookmarks is recorded for future evaluation of the resource as illustrated in Fig. 7.

VI. CONCLUSION

The present research highlights the importance of assessing the quality of knowledge objects in knowledge management systems and proposes a quantitative model for automatically evaluating that quality, based on a number of metrics. For this purpose, two quality models were introduced. The first of these exploits knowledge quality indicators to recommend quality knowledge objects for online communities, integrating the indicators into a measure and ranking the results according to estimated scores. The second model exploits a content-based recommender technique and user reputation scores for quality estimation. The results show that the proposed models perform well when integrated into the implemented KMS and tested using real data. Additionally, the use of indicators for quality estimation showed better accuracy than the ratings of reputable users.

The findings suggest several directions for future research. As the initial testing was conducted offline using a predefined data set, it is planned to run the experiment online to compare the performance of the two models with real data sets. The reputation-based evaluation model depends on explicit ratings by reputable users, and it is planned to incorporate implicit feedback within the evaluation framework for better performance. It is also planned to adapt the proposed evaluation systems as program modules that can be consolidated into any web-based knowledge management system.

REFERENCES

[1] Al-Rasheed and J. Berri, "Effective Reuse and Sharing of Best Teaching Practices," Computer Applications in Engineering Education, vol. 25, no. 2, pp. 163-178, 2017.

[2] B. Fan, L. Liu, M. Li and Y. Wu, "Knowledge recommendation based on social network theory," in Advanced Management of Information for Globalized Enterprises, 2008. AMIGE 2008, Tianjin, 2009.

[3] W. Zhao, J. Wang and G. Liu, "A Knowledge Recommendation Algorithm Based on Content Syndication," in Fourth International Conference on Computer Sciences and Convergence Information Technology, ICCIT '09, Seoul, 2009.

[4] H. Niu and H. Chen, "An Improved Recommendation Algorithm in Knowledge Metwork," JOURNAL OF NETWORKS, pp. VOL. 8, NO. 6, 1336-1342, 2013.

[5] A. Vizcaíno, J. Portillo-Rodríguez, J. P. Soto, M. Piattini and O. Kusche, "A Recommendation Algorithm for Knowledge Objects based on a Trust Model," in Proceedings of the 3rd International Conference on Research Challenges in Information Scienc, pp. pp. 93-102 , 2009 .

[6] F. Dai, X. Gu, L. Zeng and Y. Ni, "An Enterprise Knowledge Management System (EKMS) Based on Knowledge Evaluation by the Public," Knowledge Engineering and Management, Advances in Intelligent and Soft Computing, pp. vol. 123, pp. 267-272, 2011.

[7] A. Al-Rasheed and J. Berri, "Knowledge Management of Best Practices in a Collaborative Environment," International Journal of Advanced Computer Science and Applications (IJACSA), vol. 7, no. 3, 2016.

[8] G. A. Wang, J. Jiao, A. S. Abrahams, W. Fan and Z. Zhang, "Expert rank: A topic-aware expert finding algorithm for online knowledge communities," Decision Support Systems, pp. vol. 54, no. 3, pp.1442-1451, 2013.

[9] "American Heritage dictionary of the English Language," 2015. [Online]. Available: https://ahdictionary.com/word/search.html?q=reputation. [Accessed 6 12 2015].

[10] B.-C. Chen, B. Tseng, J. Yang and J. Guo, "User Reputation in a Comment Rating Environment," in Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining, 2011.

[11] B.-C. Chen, J. Yang, J. Guo and B. Tseng, "A survey of trust and reputation systems for online service provision.," Decision Support Systems, vol. v. 43, pp. pp. 618-644, 2007.

[12] B. Adler, A. Kulshreshtha, L. d. Alfaro and I. Pye, "Reputation Systems for Open Collaboration," Communications of the ACM , vol. Volume 54, no. Issue 8, pp. pp. 81-87, 2011.

[13] K. Fujimura and N. Tanimoto, "The EigenRumor Algorithm for Calculating Contributions in Cyberspace Communities," in Trusting Agents for Trusting Electronic Societies, Berlin Heidelberg, Springer, 2005, pp. pp. 59-74.

[14] M. Chen and J. P. Singh, "Computing and using reputations for internet ratings," in Proceedings of the 3rd ACM conference on Electronic Commerce, EC '01, New York, NY, USA, 2001.

[15] X. Ochoa and E. Duval, "Quality Metrics for Learning Object Metadata," in in World Conference on Educational Multimedia, Hypermedia and Telecommunications, 2006.

[16] X. Ochoa and E. Duval, "Relevance Ranking Metrics for Learning Objects," in IEEE Transactions on Learning Technologies, pp. vol. 1, no. 1, pp. 34–48, 2008.

[17] A. Zapata, V. Menendez, Y. Eguigure and M. Prieto, "Quality Evaluation Model for Learning Objects from Pedagogical Perspective. A Case of Study," in ICERI2009 Proceedings, 2009.

[18] Eurostat, "Handbook on improving quality by analysis of process variables," produced by ONS-UK, INE Portugal, NSS of Greece and Statistics Sweden, 2004.

[19] W. Mansfield and P. Grunewald, "The use of Indicators for the Monitoring and Evaluation of Knowledge Management and Knowledge Brokering in International Development," Report of a workshop held at the Institute for Development Studies 8th March, 2013.

[20] E. R. N. Valdez, J. M. C. Lovelle, O. S. Martínez, C. E. M. Marín, G. I. Hernández and S. Verma, "Social Voting Techniques: A Comparison of

the Methods Used for Explicit Feedback in Recommendation Systems," International Journal of Interactive Multimedia and Artificial Intelligence, pp. vol. 1, no. 4. p. 61, 2011.

[21] J. Bian, Y. Liu, E. Agichtein and H. Zha, "Finding the right facts in the crowd: Factoid Question Answering over Social Media," in in Proceeding of the 17th international conference on World Wide Web - WWW '08, 2008.

[22] E. Agichtein, C. Castillo, D. Donato, A. Gionis and G. Mishne, "Finding high-quality content in social media," in in Proceedings of the International Conference on Web Search and Web Data Mining, New York, NY, 2008.

[23] M. Claypool, P. W. M. Le and D. Brown, "Implicit interest indicators," in Proceedings of the 6th International Conference on Intelligent User Interfaces, 2001.

[24] D. M. Nichols, "Implicit ratings and filtering," in Proceedings of the 5th DELOS Workshop on Filtering and Collaborative Filtering, 1997.

[25] D. Oard and J. Kim, "Modeling information content using observable behaviors," in Proceedings of the 64th Annual Meeting of the American Society for Information Science and Technology, 2001.

[26] D. Kelly and J. Teevan, "Implicit feedback for inferring user preference. , 37 (2),," SIGIR Forum, pp. vol. 37, no.2, pp. 8-28, 2003.

[27] S. Fox, K. Karnawat, M. Mydland, S. Dumais and T. White, "Evaluating Implicit Measures to Improve Web Search," ACM Transactions on Information Systems, pp. Vol. 23, No 2, pp. 147-168, 2005.

[28] X. Ochoa and E. Duval, "Automatic evaluation of metadata quality in digital repositories," International Journal of Digital Libraries, pp. vol. 10, no. 2, pp. 67–91, 2009.

[29] A.-R. Alfarez. and S. Hailes, "Supporting Trust in Virtual Communities," in in 33rd Hawai'i International Conference on System Sciences (HICSS 33), 2000.

[30] A. El-korany, "Integrated Expert Recommendation Model for Online Communities," International Journal of Web & Semantic Technology (IJWesT), pp. Vol.4, No.4, 2013.

[31] M. J. Blooma, D. H. Goh and A. Y. Chua, "Predictors of high-quality answers," Online Information Review, pp. vol. 36, no. 3. pp. 383–400, 2012.

[32] h.-h. j. Gang xu, "Overview on the structure and information diffusion of the online social networks," Journal of Theoretical and Applied Information Technology , p. Vol. 51. No. 3 , 2013.

[33] N. Zardari, K. Ahmed, S. Shirazi and Z. Yusop, Weighting Methods and their Effects on Multi-Criteria Decision Making Model Outcomes in Water Resources Management, Springer, 2015.

[34] Y. Shih and D. Liu, "Product Recommendation Approaches: Collaborative Filtering via Customer Lifetime Value and Customer Demands," Expert Systems with Applications, ,, pp. vol. 35, nos. 1/2, pp. 350-360, 2008.

[35] M. Eleni and K. John, "Utilizing vector space models for user modeling within a learning environments," Computers & Education , pp. 51(2), 493–505, 2008.

[36] G. Shani and A. Gunawardana, "Evaluating Recommendation Systems," in Recommender Systems Handbook, Springer US, 2011, pp. pp 257-297.

[37] G. A. Fredricks and R. B. Nelsen, "On the relationship between spear spearman's rho and kendall's tau for pairs of continuous random variables," Journal of Statistical Planning and Inference, p. 137(7):2143–2150, 2007.

[38] J. Mazurek and K. slova, "EVALUATION OF RANKING SIMILARITY IN ORDINAL RANKING," Acta academica karviniensia, pp. pp. 119-128, 2011.

[39] M. G. Kendall and A. Stuart, The Advanced Theory of Statistics, Volume 2: Inference and Relationship, Griffin. ISBN 0-85264-215-6. (Sections 31.19, 31.21), 1973.

[40] D. Smiley, E. Pugh, K. Parisa and M. Mitchell, Apache Solr 4 Enterprise Search Server (1st ed.), Birmingham: Packt Publishing, p. 451, 2014.

[41] H. Abdi, "The Kendall Rank Correlation Coefficient," in In: Neil Salkind (Ed.). Encyclopedia of Measurment and Statistics, 2007.

[42] J. White, A. Yeats and G. Skipworth, Tables for statisticians, United Kingdom: Nelson Thornes Ltd, 1979.