AN EMPIRICAL STUDY ON CRITERIA FOR ASSESSING INFORMATION QUALITY IN CORPORATE WIKIS

(Completed academic paper)

Therese Friberg
University of Paderborn, Germany
Institute for Mechatronics and Design Engineering
friberg@cik.upb.de

Wolfgang Reinhardt
University of Paderborn, Germany
Department of Computer Science
wolle@upb.de

Abstract: Wikis gain more and more attention as tool for corporate knowledge management. The usage of corporate wikis differs from public wikis like the Wikipedia as there are hardly any wiki wars or copyright issues. Nevertheless the quality of the available articles is of high importance in corporate wikis as well as in public ones. This paper presents the results from an empirical study on criteria for assessing information quality of articles in corporate wikis. Therefore existing approaches for assessing information quality are evaluated and a specific wiki-set of criteria is defined. This wiki-set was examined in a study with participants from 21 different German companies using wikis as essential part of their knowledge management toolbox. Furthermore this paper discusses various ways for the automatic and manual rating of information quality and the technical implementation of such an IQ-profile for wikis.

Key Words: Information Quality, Corporate Wikis, Information Quality Criteria, Empirical Study

INTRODUCTION

The largest and most well known representative of wikis is Wikipedia [28], which exists since 2001. Today Wikipedia is biggest multilingual free-content encyclopedia on the Internet with more the 67 million US-visitors as of June 2009 [23], 75,000 active contributors and 13 million articles in 260 languages, thereof more than 2.9 million in English [28]. The success of Wikipedia is founded on its free availability and possibility to collaborate in it [31]. Wikis attain increasing importance in nowadays community and for the knowledge management in enterprises. The popularity of wikis is based on the fact that they enable an effective archiving, organizing and sharing of data to the users. Writing in a wiki is a transparent process for the users on which they can collaborate actively. Furthermore the performed work is directly viewable in the system, which has a positive effect on the user's motivation. Wikis inside of organizations, called corporate wikis, differ from the free accessible ones primarily in the fact that they have explicit access control. That means that the entrance to selected areas can be restricted to certain user groups. In the further process corporate wikis stand in the foreground and hence, they build the context of this examination.

One of the biggest problems concerning the growing content of wikis is the uncertainty of quality of the input. The important advantage of the open access causes simultaneously an enormous problem because each person is able to add, delete, or modify information without a previous review-process independent from the fact if an article gets improved or declined through it. Especially inside of organizations wrong or missing information produce enormous costs. A survey of Information Builders GmbH shows that 54% of 610 asked managers see the biggest barrier against making good decisions is inconsistent, deficient, and incomplete information in organizations [18]. So using a wiki inside an organization is followed by

the question about the quality of information. A credible assessment of the quality of the content on a platform intensifies the trust in knowledge management system and increases the confidence of the user [9]. To satisfactorily answer the question about the information quality is has to be clarified by which criteria it is determined and which factors play a major role concerning the rating of these criteria.

Therefore we have developed a set of information quality criteria adapted for corporate wikis. Afterwards this set has been verified by a survey with authentic users of corporate wikis. With the results from the study we optimized the identified wiki-set to provide an optimal assessment of information quality in corporate wikis.

The remainder of this paper is structured as follows: first the background of the research and related work is referred. Afterwards the wiki-set to assess information quality inside of organizations is derived. After presenting the conducted study and its results, technical approaches for realizing the wiki-set are identified. The paper closes with a conclusion and outlook on future research.

STATE-OF-THE-ART AND BACKGROUND

You can hardly find examinations concerning the analysis of information quality in corporate wikis. Studies focusing the information quality of articles in Wikipedia show that there can evolve articles with high quality content and that these articles can even be gauged with nameable, paper-based, and peer-reviewed encyclopedias [12,15,29]. These findings confirm the existing potential that a wiki can contain high quality articles. In the past a lot of researchers have defined criteria catalogues to define the determination of information quality. Only a few of them focused on wikis and often the criteria are hardly feasible. Hence, to achieve a criteria catalogue for corporate wikis, we had to make adaptations of the existing ones. As [21] emphasize, criteria of information quality vary with the context in which they are used. So we had to regard the corporate context in detail. We present some of the existing criteria catalogues in the following section with the intention to build a base for the definition of the so-called wiki-set to determine information quality in corporate wikis.

The authors in [26] developed a set of 15 information quality criteria under intensive embracing of answers from participants of arranged surveys. The authors have derived a criteria catalogue, which contains the most important criteria to assess information quality from the perspective of users. The framework of [7] also contains 15 criteria, which primarily focus information in databases. The setting plays an important role in the subsequent literature and hence it has been included in the development of the wiki-set. [1] developed a framework with six criteria by focusing Internet pages. The authors created an often-cited catalogue and offer some interesting viewpoints to determine the wiki-set. By analyzing existing frameworks [8] generated a model with 16 criteria for information quality. The analysis was done by literature research and empirical studies. The framework of [4] focuses wikis in detail and contains nine different criteria. The author analyzes the articles of Wikipedia based on criteria catalogues developed for news. [14] have developed a framework with eight criteria, which defines information quality during transactions on the Internet. The author sees the determination of information quality as a process during which a message can become a benefit. The framework of Wikipedia [29] consists of ten criteria, which shall intend the quality of the published articles on the platform. The set of criteria is still used by the community of Wikipedia to assess the information quality of the written output. The strong focusing on a wiki offers huge potential for developing the wiki-set.

APPLIED METHOD

The following section describes the derivation of the wiki-set to assess information quality in corporate wikis.

Development of the Wiki-Set

The development of the potential information quality criteria catalogue was made stepwise. In a first step the overlaps of the before-mentioned frameworks were pointed out and in a second phase the adaptability of the criteria were analyzed. In the beginning there are 50 different labeled criteria derived from the seven frameworks [1,4,7,8,14,26,29] presented above.

To achieve a definition of information quality it is not enough to identify the common items of the existing models. The criteria depend on the used context [21] and furthermore 50 criteria are too much to realize an efficient set of criteria for information quality in wikis. Because of that in a first step common criteria were merged to reduce the quantity. After the first phase of adaptations from the before 50 criteria 19 remained:

- Accuracy
- Accessibility
- Timeliness
- Precision
- Clarity
- Currency
- Reputation
- Interpretability
- Comprehensibility
- Layout and use of multiple media
- Objectivity
- Security
- Completeness
- Usability
- Correctness
- Believability
- Relevancy
- Value-added
- Amount of information

Amongst these remaining criteria there were still very similar criteria regarding their impact and irrelevant criteria for the assignment in corporate wikis. Hence, to continue the optimization of the set of information quality criteria, we ran through a second phase. The remaining 19 criteria were analyzed to their relevance and operationalizability. *Accuracy* is subsumed to the criterion *correctness*. Both criteria intend to measure correct and accurate information, so the reflection of the truth is meant. The criterion *objectivity* is defined to describe clear and true facts without deforming them by personal feelings or other effects. Hence, the criterion was allocated to *believability*. *Reputation* was summed up with *believability* [14] because this criterion aims at the prestige of the creator, which is also contained in the criterion *believability*. But the reputation of a user has no high impact in a corporate wiki because there the employees want to attain a high reliable effect on their colleagues and principals. Accessibility is the duty of a whole system, not that of a wiki article. Thus, the accessibility for disabled people has to be checked during the usability testing in the development of the system. The possibility to access to the information from every desired location is the challenge of network administrators and has nothing to do with the assessment of the content inside a corporate wiki. Especially in organizations the criterion *security* is a very important subject but it already has to be considered during the phase of implementation and testing. *Value-added* was deleted from the wiki-set because the benefit of a wiki inside an organization can be measured through the common activity. Furthermore the value of a single article is very subjective and is better represented by the criterion *relevance*. The two criteria *timeliness* and *currency* were concentrated to one item because both of the criteria deal with the actuality of information [14]. *Amount of information* was merged to *completeness* since both items deal with an adequate quantity of information. The criterion *precision* was subsumed to *accuracy* as both of the criteria examine the exactness of information. *Usability* does not point at the information quality of a single article; the criterion has to be fulfilled by the system itself. *Clarity*, *comprehensibility* and *interpretation* were subsumed to one item because all of them define a clear and articulate language of information. *Layout and use of multiple media* was deleted as a single criterion because it is rather seen as an indicator for the *comprehensibility*.

After this second phase of reduction six criteria remain. This set seem to cover substantial information quality based on the framework of previous research. The remaining criteria are described concerning their context in the next section.

Definition of the criteria

The condensed set of information quality criteria is defined for the context of corporate wikis. Furthermore selected indicators are presented for the six criteria. Here, indicators mean that the identified factor seems to have an impact on the assessment on a certain criterion from the perspective of users.

Believability

Believability complies with the perceived truth of information from the perspective of a receptor [14]. But this truth has not necessarily to go along with the objective reality. To easier assess the criterion, the relationship to the particular main author, a broad and correct list of literature and the previous access to an article were identified as indicators.

Relevance

Relevance summarizes all information being meaningful from the point of view of a user. Hence, the criterion judges, if information has the potential to answer satisfactorily to a request. For example, an idea could be to analyze if the title is chosen apposite and thus the article contains what is promised. Another indicator is the so-called *click popularity*. In a corporate wiki it is interesting and even significant to see popular articles because they could predict a relevant article to a user. Furthermore an available abstract of an article could help the user to assess the relevance of the object. The last identified indicator is the length of an article. We assumed that a longer article could give an advice to a more relevant information object.

Timeliness

Timeliness declares if information is outdated or as up-to-date as required. A potential indicator for this criterion is the creation date in association with the date of last modification. If the last change is only a short time ago, you can assume that the content is relative up-to-date. Another indicator for timeliness is checking availability of hyperlinks within the text and the list of references.

Completeness

Completeness defines, if an issue is covered broadly within an article and possibilities for further reading are given, for example in form of hyperlinks. The list of references offers a good first impression of the completeness. A high number of sources could point towards a more complete article. Furthermore some authors have already shown that checking the length of an article is a good sign for the completeness of a text [2,3,22,24].

Comprehensibility

The correct gathering and processing of information is meant by the criterion comprehensibility. Objective indicators like a confusing structure or a high amount of mistakes in spelling, grammar and punctuation yield first evidences and flag conspicuous articles. Furthermore the existence of figures, tables and other media as well as additional links could support the comprehensibility of an article. The last indicator identified is the declaration of required prior knowledge. With this information the user can easily appraise the difficulty before consuming the article.

Correctness

Correctness defines if given information correctly represents the reality. Hence, this criterion presents the first ultimate ambition of information quality. The assessment of this criterion is a huge problem. Because of that for development a criteria catalogue a lot of authors have left the analyzing of the criterion out (for example [4,20]). Based on this fact, the criterion correctness is eliminated from the wiki-set in order to simplify the assessment of information quality in corporate wikis.

Table 1 lists the identified criteria and important indicators for the respective criterion of information quality.

Criterion	Indicators
Believability	• Connection to the author • Recognition of the author • Existence of references and additional links • Access rates of an article
Relevance	• Title of an article • Abstract of an article • Access rates of an article • Length of an article
Timeliness	• Date of last modification • Creation date • Outdated or unreachable links
Completeness	• Length of an article • Existence of references and additional links
Comprehensibility	• Mistakes in spelling, grammar or punctuation • Structure of an article • Existence of pictures, tables and other media • Existence of additional links • Declaration of required prior knowledge

Table 1: The identified criteria and its respective indicators

The five identified criteria and associated indicators were examined in a survey with participants from organizations, which utilize corporate wikis. The study and the results are illustrated below.

STUDY CONDUCTED

The criteria and the respective indicators defined have been inspected in a study in [11] within the scope of a master's thesis. The study was conducted amongst users from 21 different German corporate wikis. Altogether we received 206 completely filled out and valid answers over a period from September to November 2008. The survey was conducted as an online survey and the link to the survey was distributed through contact persons at the respective companies.

Figure 1 shows the economic sectors from the participants of the survey. The main part of the participants have recorded to work in the sector IT what can be explained with the fact that wikis are mainly used in companies familiar with the use of information technology. Against this background it is unsurprising that only 17.5% female and 82.5% male participants took part in the survey. Regarding the size of the companies taking part in the survey it can be stated that 51% of the companies were small and medium enterprises with 101-500 employees. More than one-fifth (21.4%) of the companies had more than 1000 employees, whereas another one-fifth had less than 100 employees.
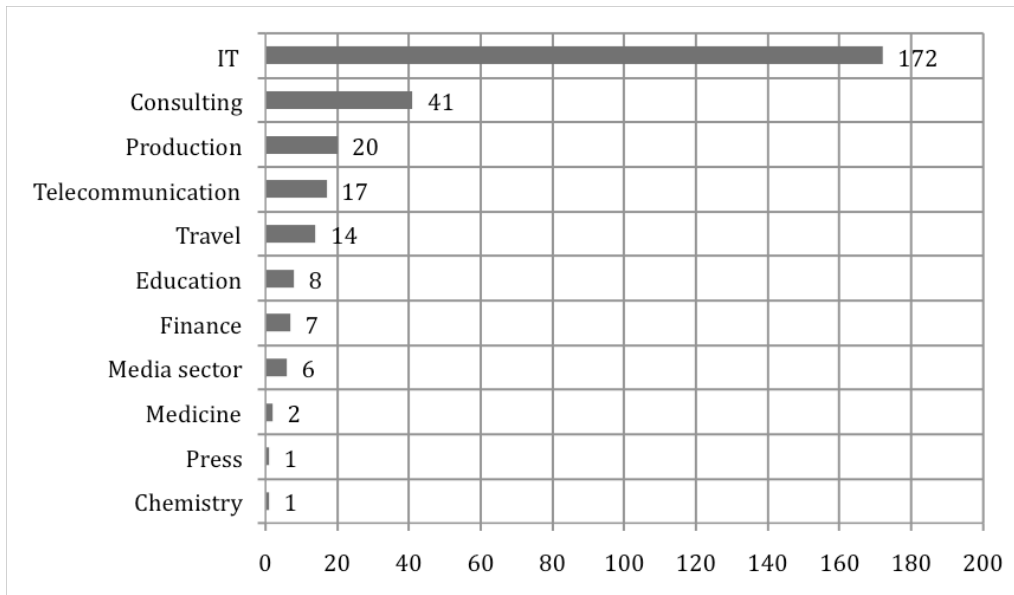
Figure 1: Economic sectors of companies taking part in the survey (multiple selections were allowed)

As shown in Figure 2, the users in the investigated wikis are rather active with 83% of users that at least once created an article and more than 71% of users that ever edited an article from another user.
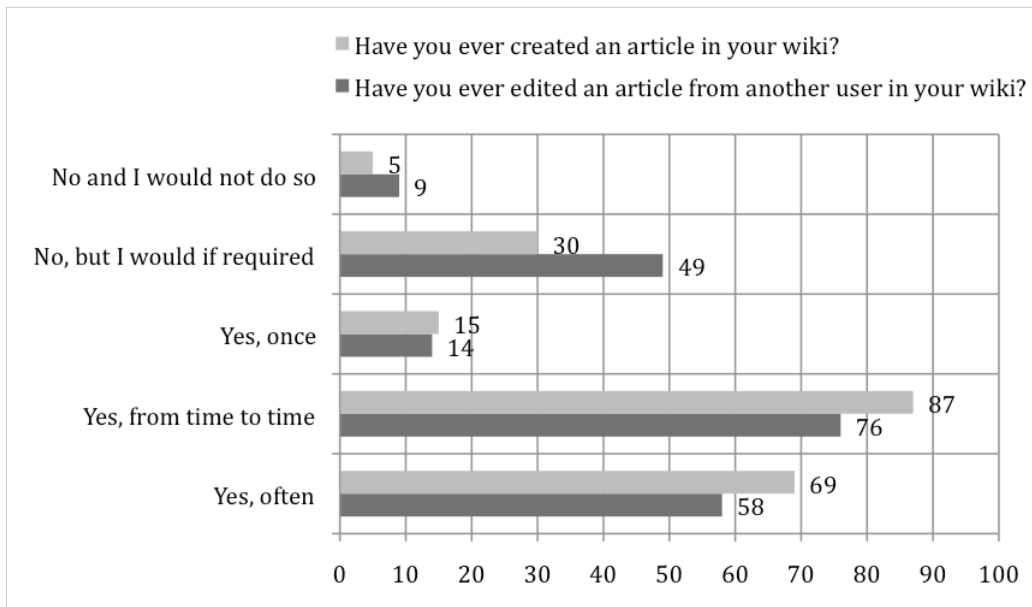

Figure 2: Activity index of users in the investigated wikis

Development of the Wiki-Set

The participants had to rate the impact of the five criteria on a four-stage Lickert scale. To make the results more manageable the two highest stages were summarized to *important* and the two lowest to *unimportant*. Figure 3 shows the distribution of the impact of the five asked criteria *believability*, *relevance*, *timeliness*, *completeness* and *comprehensibility*. From participant's point of view, the criterion comprehensibility was outlined as the most important criterion to assess information quality of an article with 96,1%. The results for the criteria believability, relevance and timeliness were directly behind (all

over 91%) and were also defined as important from the perspective of the users. The criterion completeness clearly takes the last position with only 68,5% and seems to be important for the participants but with less impact than the other four criteria.
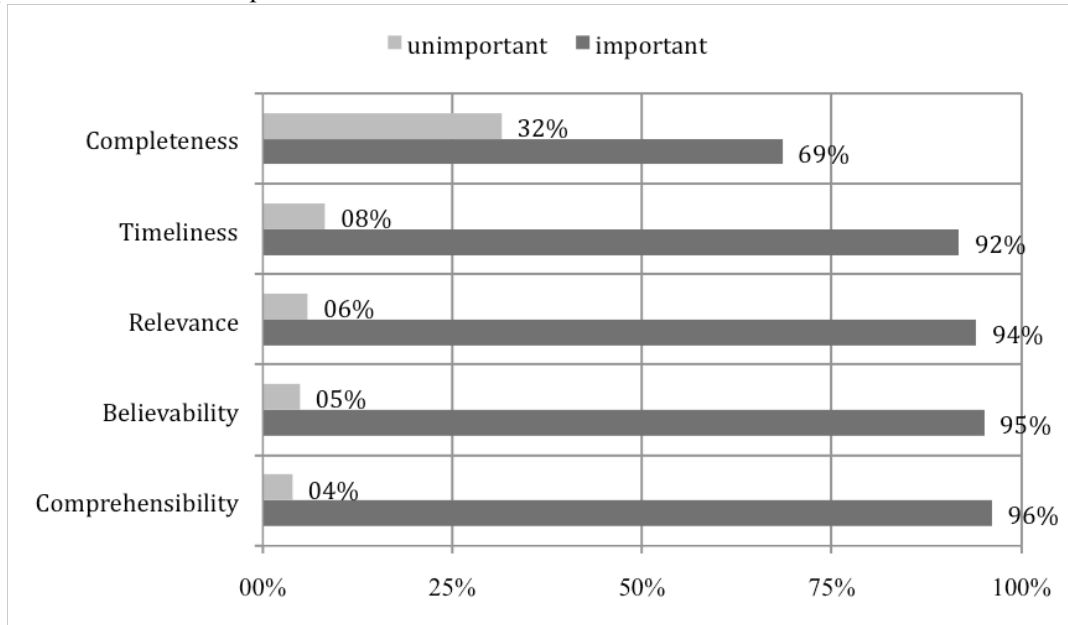

Figure 3: Overall rating of the single criteria of the wiki set

The identified indicators were questioned in the study as well. The rating of the influence on a certain criterion was also done on a four-step Lickert scale. 1 was the weakest value and 4 the strongest value. To make the indicators better comparable, the means for all indicators were calculated (see table 2).

Criterion	Indicators	Mean
Believability	• Connection to the and recognition of the author	2.85
	• Existence of references and additional links	2.38
	• Number of published article of the author	2.16
	• Access rates of an article	2.12
Relevance	• Title of an article	3.66
	• Abstract of an article	2.93
	• Access rates of an article	2.27
	• Length of an article	1.66
Timeliness	• Date of last modification	-
	• Outdated or unreachable links	3.18
	• Creation date	2.40
Completeness	• Existence of references and additional links	2.33
	• Length of an article	1.62
Comprehensibility	• Structure of an article	3.61
	• Existence of pictures, tables and other media	3.12
	• Existence of additional links	3.05
	• Mistakes in spelling, grammar or punctuation	2.96
	• Declaration of required prior knowledge	2.54

Table 2: Arithmetic means for the single indicators of the criteria of the wiki-set

Concerning the criterion believability, the indicator *connection to the author* reached the highest value (with 2.85). The indicators *existence of references and additional links*, *connection to the and recognition of the author* and the *access rates of an article* still had a mean around 2 and thus also impact on the criterion.

The identified indicators for the criteria relevance achieved differing values. Whereas *the title of an article* had a very high value of 3.66 and the *availability of an article abstract* yielded in a mean of 2.93, the *access rates of an article* was in the middle (with 2.27). The *length of an article* got a low mean with 1.66 and this result gets approved in the context with the criterion completeness where the value was only 1.62. This indicator was the only one that got a value under 2.

The highest indicator for the criterion timeliness was *outdated or unreachable links* with a mean of 3.18. The *creation date* reached a value of 2.40. For the last indicator *date of last modification* there was no significant and definite period discovered.

All indicators of the criterion completeness lie behind the other values. The *existence of references and additional links* with a mean of 2.33 is a bit higher than the before-mentioned *length of an article* (1.62).

All indicators of the criterion comprehensibility possessed a high value. The question concerning the influence of the *structure of an article* was very conspicuous with a mean of 3.61. The existence of *multiple media*, *additional links* as well as *mistakes in spelling, grammar or punctuation* all reached a mean around 3 and hence also show a strong impact on the criterion. The last indicator of comprehensibility, the *declaration of required prior knowledge*, had a little lower mean and shows less impact from the user's perspective.

TECHNICAL APPROACHES FOR REALIZING THE WIKI-SET

The results of the study conducted show numerous technical approaches for realizing the wiki-set for information quality. In our opinion, the measurement and rating of the information quality and the respective criteria should be organized in a semi-automatically process. The combination of automatically evaluated and manually rated indicators yields a meaningful information quality profile. For each article in the wiki such an *IQ-profile* should be stored and visualized at a dispositive place within the article. Studies show that such additional information is better placed within the article than on a separate page [5].

Regarding the automatic analysis of wiki articles in respect to their information quality, different analyses are technically feasible and reasonable to be included in the IQ-profile:

a) Existence of links to other articles in the same wiki,
b) Existence of external links,
c) Availability of external links,
d) Evaluation of hit rates and referrer of an article,
e) Evaluation of changes of an article (number of different authors, size and frequency of changes),
f) Existence of media like tables, figures or videos and
g) Number of spelling mistakes, foreign words and pronouns used.

Besides these generic applicable approaches for the automatic evaluation and manual rating of information quality in wikis we consider specific realizations for the five indentified criteria of the wiki-set subsequently.

As the study pointed out, the criterion *believability* is strongly interconnected with the authors of an article. To support an article's believability all authors and their respective shares in the contents of the article should be depicted clearly in the IQ-profile. [5] suggests using pie charts to visualize the respective shares in an article's modifications (cf. figure 4). Furthermore each wiki should be extended by automatically generated author profiles that shed light on each user's areas of activity within the wiki and

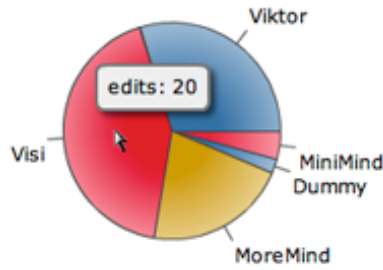allow the direct contacting of the authors.

Figure 4: Pie chart showing the respective shares in editing a wiki article [5]

Regarding this automatically generated author profile, we asked the participants of the study which personal details they would wish to see. Figure 5 shows the answers of the participants. Additionally the participants could enter individual indicators and details that influence the believability of an author. The most common answers were

- Department, Working Field (stated 8 times)
- Name (stated 6 times)
- Position, Situation within the organization, reputation, area of accountability (stated 4 times)
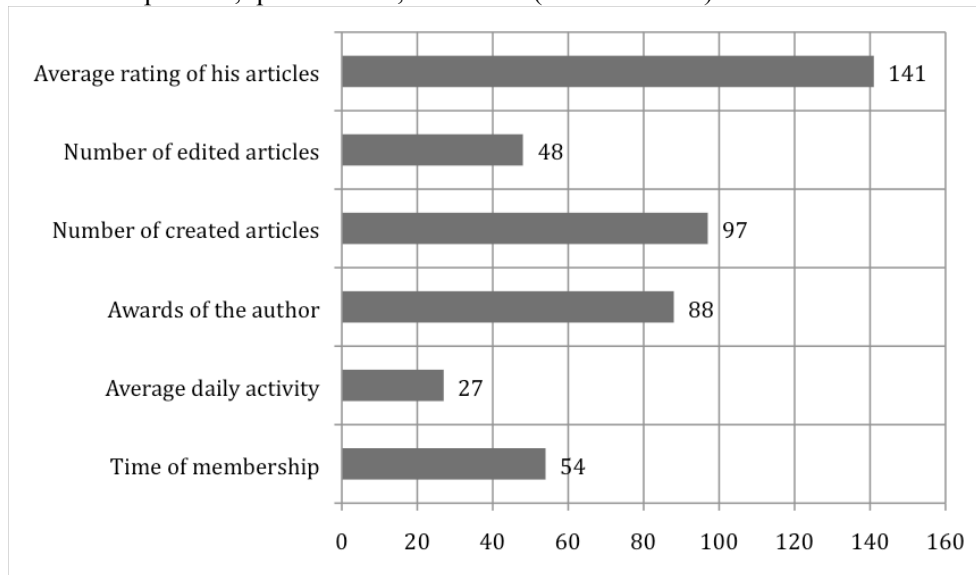- Technical competence, qualification, education (stated 3 times)

Figure 5: Results for the question on what personal details about an author users would wish to have, in order to better asses the information quality of one of his articles (multiple selections were allowed)

Relevance is the catchiest criterion of information quality and moreover the relevance of an article is strongly dependent on the user context of the information seeker. Hence there are hardly any automatic approaches to rate an article's relevance. Besides the usage of semantic toolkits, making use of natural language processing and named entity recognition in order to return facts and entities hidden in the text of a wiki article, tag clouds seem to be a way of simplifying the user's appraisal of an article's relevance. Regarding the manual rating of an article's relevance for stated search terms, light-weighted approaches should be employed. Similar to Google's search wiki [6], wiki users ought to rate the search results and state their appropriateness to the given search terms. For upcoming searches the relevance level and the

respective click popularity for given search terms ease the relevance judgment for users [10]. The survey also showed that the existence of an abstract of an article simplifies the identification on an article's relevance.

The importance of the criterion *timeliness* was affirmed in the study conducted. There are many indicators for timeliness and various ways for an automatic evaluation. Nevertheless timeliness is closely related to the context of a wiki article and its maturity. As of today, there is no way of identifying if an article is out of date, because we cannot automatically judge if information is outdated. If an article has reached a high level of maturity there will not be very much modifications on it. Indicators for timeliness that are automatically rateable are the number of accesses in a specific period of time, reachable links and modification dates. By automatically evaluating the access rates and adding them to the IQ-profile, users can easily meter how often an article was accessed and thus infer the timeliness and temporal relevance of the article. Another indicator for timeliness of wiki articles identified by the study conducted are missing or faulty links. If existing links do not work (anymore) this does negatively affect the timeliness of an article, since the information seems to be outdated. Lastly the IQ-profile should clearly show the date of the last modification and discussion entry.

Completeness does not seem to have a very high significance to the participants of the study. This may be originating in the fact that completeness like relevance is heavily dependent on the context and previous knowledge of the information seeker. The criterion completeness refers to a complete article, but it is not rateable if an article holds all information. Thus you can only use light-weighted indicators for this criterion. Signs for a complete article are the existence of media, links and the declaration of uses sources and additional material. These factors can be automatically checked and contribute to the IQ-profile.

The study points out *comprehensibility* as elementary part of the information quality of a wiki article. If the comprehensibility of an article is bad or if the article is completely incomprehensible the rating of the other criteria cannot take place. Structure seems to be a good indicator for the comprehensibility of an article; a good structured article lets users easier understand an article. Parts of a good structured article are the existence of sections and subsections, headings and captions, enumerations and lists as well as tables and other media. Unfortunately there are only very little technical possibilities of determine a good structured article because structure is constrained with the context, topic, complexity and length on an article. Mistakes in spelling, grammar or punctuation have a negative impact on the comprehensibility of a written text. Using spell and grammar checker (cf. [30]) and by dint of thesauri, the rate of these mistakes can be computed and the usage of foreign words can be identified. A higher rate of errors and foreign words has a negative effect on the criteria comprehensibility. If an article accounts for the required prior knowledge of the reader this may help the reader in assessing the complexity of the article and thus help understanding the article. The declaration of the required prior knowledge of the reader itself cannot be considered an indicator of comprehensibility because one would need to match the effective knowledge of the readers with the stated required knowledge. On the other hand the study pointed out that this declaration would help assessing the severity on an article. A possible implementation of this indicator information quality can be adopted from the Learning Object Metadata Standard (LOM) [17] that stores the difficulty on a learning object as well as the semantic density and the intended target audience.

Rating and representation of the IQ-profile

The presented technical approaches for manual and automatic rating of the information quality of wiki articles identify first steps towards the extension of wikis to be high-quality media. Contrary to well-known media, the pace of artifact modification in wikis is very fast. This celerity must be taken into account when rating the information quality of articles; the IQ-profile of an article has to be stored separately for every single version of an article. The overall information quality composes as the sum of the single IQ-profiles. Regarding the presentation of the IQ-profile the study asked how to visualize the

value of the information quality in the IQ-profile. Figure 6 shows the results of this question and outlines that a star rating is the most valuable way of representing the information quality in a wiki.
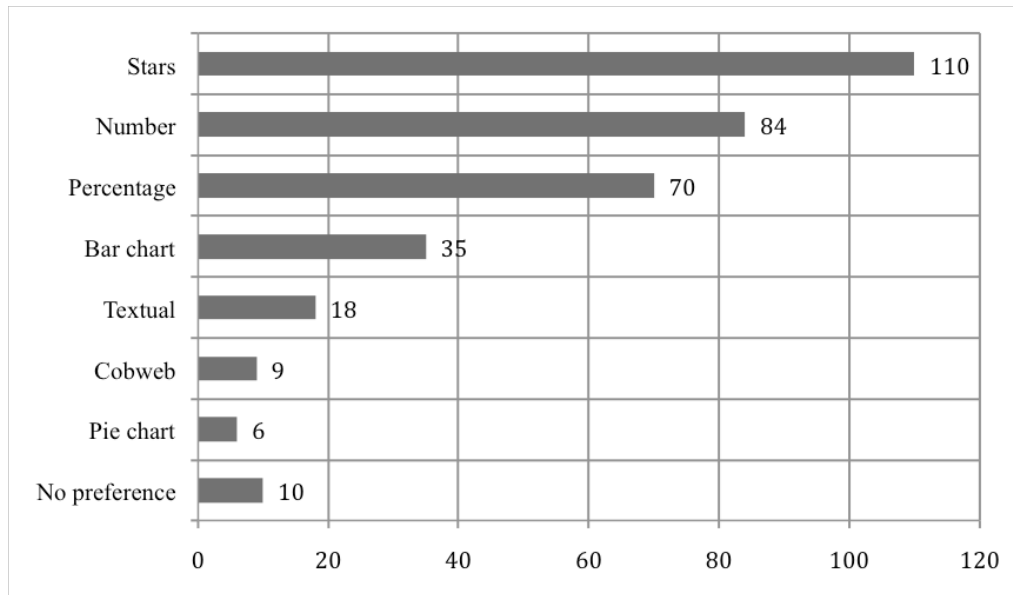

Figure 6: Results for the question on how to visualize the overall information quality in a wiki (multiple selections were allowed)

As the results of the survey show, users are willing to invest five clicks or 35 seconds to rate the information quality of an article. Thus the manual rating should be easy and viable with little expenditure of time. This is realizable using a simple star rating of the main criteria – optional a detailed rating of all single factors should be selectable to the users.
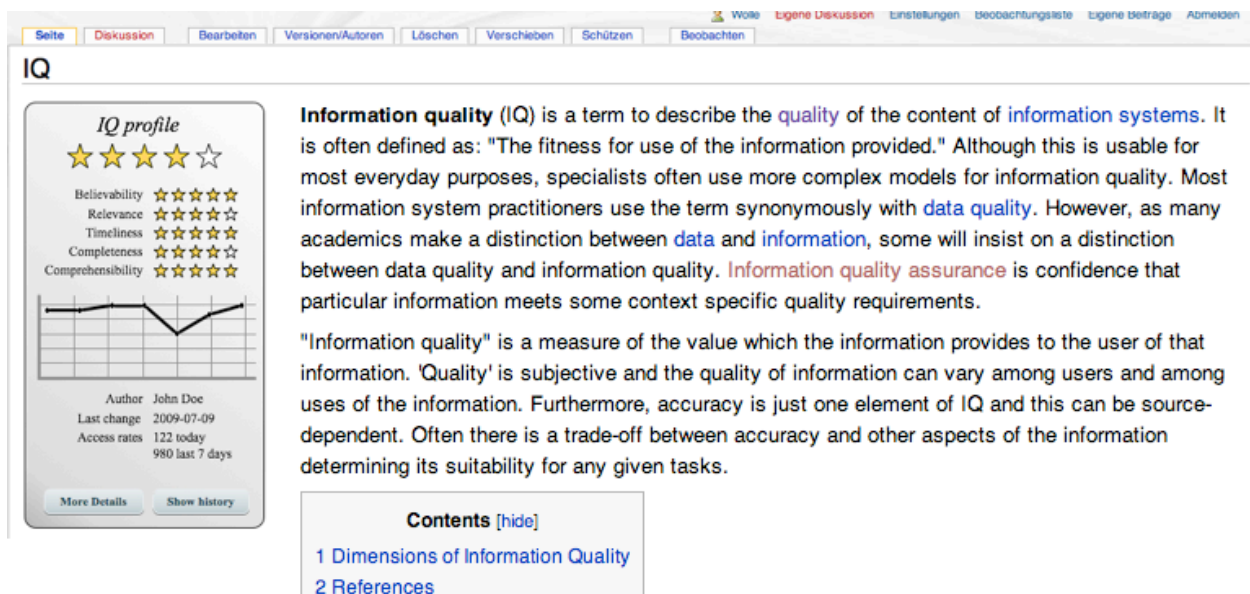

Figure 7: Prototypical integration of the IQ-profile in a wiki article

Figure 7 shows a current prototype of an extension to MediaWiki allowing the rating and visualization of the IQ-profile of an article. The extension is currently under development within the scope of a master's

thesis and will be verified for its practical application in a field study later this year. The IQ-profile depicts the overall value of information quality, based on automatically and manually ratings on a five stars scale. The values of each criterion of the wiki-set are also displayed. If a user has not already rated the latest version of the article he can do so by selecting the desired number of stars for each criterion. By clicking in the button "more details" the user can rate the single indicators for each criterion as they are stated in section *Definition of the criteria*. The line chart shows the dynamics of the information quality of the article over a specific period of time. By clicking the button "show history" the IQ-profile shows how all indicators and the access rates performed over time.

CONCLUSION AND OUTLOOK

In this paper we have presented the body of knowledge in the field of information quality in textual documents and derived a wiki-set of five criteria for assessing information quality in corporate wikis. We presented the results of an empirical study conducted with participants from 21 different German companies using wikis as essential part of their knowledge management toolbox.

The results presented in this paper possess high potential to initiate a dedicated assessment of information quality in corporate wikis. A future approach could be to examine further indicators for each of the criteria to reach a more automated rating to disburden the user more and more. Another task for further research is to detect an appropriate and flexible weighting of the criteria. For example, could the weighting be adapted to an individual user. Furthermore, a similar study in foreign countries would be an interesting approach to compare the results with the here presented results from German companies. During the evaluation of the conducted study no significant differences between economic sectors, age of gender of the participants could be revealed. To make more generic claims from a future study, the considered economic sectors as well as the gender of the participants should be more balanced.

The drafted technical approaches to realize the wiki-set are under development in an ongoing master's thesis and will be tested with selected companies regarding their practical use. The exploration of the dynamics eventually will lead to a new understanding of the maturing of wiki articles. [25] show approaches for identifying semantic convergence of wiki articles in order to make claims about the maturity – eventually the convergence of the IQ-profiles will contribute in making these claims.

ACKNOWLEDGEMENT

REFERENCES

[1]   Alexander, J. E. and Tate, M. A.: *Web Wisdom; How to Evaluate and Create Information Quality on the Web*. L. Erlbaum Associates Inc., Hillsdale, NJ, USA. 1999.

[2]   Blumenstock, J.E.: Size matters: word count as a measure of quality on Wikipedia. In: *WWW '08: Proceeding of the 17th international conference on World Wide Web*. pp. 1095-1096. 2008.

[3]   Blumenstock, J.E.: *Automatically Assessing the Quality of Wikipedia Articles*. School of Information. Paper 2008-021. 2008.

[4]   Brändle, A.: *Zu wenige Köche verderben den Brei: Eine Inhaltsanalyse der Wikipedia aus Perspektive der journalistischen Qualität, des Netzeffekts und der Ökonomie der Aufmerksamkeit [Too less cooks spoil the broth: A content analysis of the Wikipedia from the view of journalistic quality and the economy of attention]*. Lizentiatsarbeit, University of Zürich, Philosophische Fakultät. 2005.

[5]   Dite, V.: *Visualisierung komplexer Daten und Strukturen in Wikis [Visualizing complex data and structures*

in wikis]. Masters thesis, University of Paderborn, 2009.

[6]   Dupont, C., Anderson, C.: *SearchWiki: make your own search*. Available online at http://googleblog.blogspot.com/2008/11/searchwiki-make-search-your-own.html. 2008. (last viewed on 2009-05-11)

[7]   English, L. P.: *Improving Data Warehouse and Business Information Quality: Methods for Reducing Costs and Increasing Profits*. John Wiley & Sons, Inc., New York, NY, USA. 1999.

[8]   Eppler, M. J.: *Managing Information Quality: Increasing the Value of Information in Knowledge-intensive Products and Processes*. Berlin: Springer. 2003.

[9]   Eppler, M.J., Gasser, U., and Helfert, M.: Information Quality: Organizational, Technological, and Legal Perspectives. *Studies in Communication Sciences*. 2(4): 1-16. 2005.

[10]   Freyne, J., Farzan, R., Brusilovsky, P., Smyth, B., Coyle, M.: Collecting community wisdom: integrating social search & social navigation. In: IUI '07: *Proceedings of the 12th international conference on Intelligent user interfaces.* New York, NY, USA : ACM. – ISBN 1–59593–481–2, S. 52–61. 2007.

[11]   Friberg, T.: *Bewertung der Informationsqualität in Unternehmenswikis. Empirische Studie zu Kriterien der Informationsqualität [Rating the information quality in corporate wikis. An empirical study on the criteria of information quality]*. Masters thesis, University of Paderborn, 2009.

[12]   Giles, J.: Internet encyclopaedias go head to head. *Nature*, 438: pp. 900–901. Available online at http://www.nature.com/nature/journal/v438/n7070/full/438900a.html. 2005. (last viewed on 2009-07-08)

[13]   Goh, C. H., Bressan, N., Madnick, S.E., Siegel, M.D.: Context Interchange: New Features and Formalisms for the Intelligent Integration of Information. *ACM Transactions on Office Information Systems*, 1999.

[14]   Gräfe, G.: *Informationsqualität bei Transaktionen im Internet - Eine informationsökonomische Analyse der Bereitstellung und Verwendung von Informationen im Internet [Information quality at transactions on the Internet – An information economics analysis of the supply and usage of information on the Internet]*. Wiesbaden: Deutscher Universitäts-Verlag. 2005.

[15]   Hammwöhner, R., Fuchs, K.-P., Kattenbeck, M., Sax, C.: *Qualität der Wikipedia. Eine vergleichende Studie [Quality of the Wikipedia. A comparative study]*. In Oßwald, A.; Stempfhuber, M; Wolff, C. (eds.) Open Innovation. Proc. 10th Int. Symposium on Information Science in Cologne. UVK, pp. 77-90. 2007.

[16]   Huang, K., Lee, Y., Wang, R.: *Quality Information and Knowledge*. Prentice Hall, Upper Saddle River: N.J., 1999.

[17]   IEEE, Learning Technology Standards Committee Learning Standards Committee: Draft Standard for Learning Object Metadata. Available online at http://ltsc.ieee.org/wg12/files/LOM_1484_12_1_v1_Final_Draft.pdf. 2002. (last viewed on 2009-05-19)

[18]   Information Builders GmbH: *Studie: Millionenschaden durch mangelnde Informationssysteme in Unternehmen [Study: Loss of millions by lack of information systems in enterprises]*. Available online at http://informationbuilders.de/presse/PR_2007/WF_No_barriers.html. 2007. (last viewed on 2009-05-10)

[19]   Kotler, P.: *Marketing Management: Analysis, Planning, Implementation, and Control*. 9th Edition ed. Prentice Hall, 1997.

[20]   Rössler, P.: Qualität aus transaktionaler Perspektive. Zur gemeinsamen Modellierung von "User Quality" und "Sender Quality": Kriterien für Onlinezeitungen [Quality from a transactional point of view. On a common modeling of user quality and sender quality: Criteria in online newspapers]. In Beck, K., Schweiger, W., and Wirth, W. (Eds.): *Gute Seiten - schlechte Seiten. Qualität in der Online-Kommunikation [Good pages – bad pages. Quality in the online communication]*, pp. 127–145. München: Verlag Reinhard Fischer. 2004.

[21]   Shankar, G., Watts, S.: A Relevant, Believable Approach for Data Quality Assessment. In *IQ*, pp 178–189, 2003.

[22]   Stvilia B., Twidale, M.B., Smith, L.C., Gasser, L. "Assessing information quality of a community-based encyclopedia." In: *Proceedings of the International Conference on Information Quality - ICIQ 2005*. pp. 442-454. 2005.

[23]   Quantcast Coporation. *Quantcast Audience Profile for wikipedia.org*. Available online at http://www.quantcast.com/wikipedia.org. 2009. (last viewed 2009-07-12)

[24]   Tang, R., Ng, K. B., Strzalkowski, T., Kantor, P. B.: Automatically Predicting Information Quality in News Documents. In: *Proceedings of Human Language Technology - North American Chapter of the Association for Computational Linguistics*. 2003.

[25]   Thomas, C., Sheth, A. P.: *Semantic Convergence of Wikipedia Articles*. In: WI '07: Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence. pp. 600-606. 2007

[26]   Wang, R. Y., Strong, D. M.: Beyond Accuracy: What Data Quality Means to Data Consumers. In: J. Manage (Ed.) *Information Systems*, 12(4):5–33. 1996.

[27]   Wiegand, D.: *Entdeckungsreise: Digitale Enzyklopädien erklären die Welt [Expedition: Digital encyclopedias explain the world]*. In: *c't*, 6/2007:S.136–145. 2007.

[28]   Wikimedia Foundation. Available online at http://en.wikipedia.org/w/index.php?title=Wikipedia:About&oldid=299569761, version as of 2009-07-10, 2009.

[29]   Wikimedia Foundation: *Featured Article Criteria*. Available online at http://en.wikipedia.org/w/index.php?title=Wikipedia:Featured_article_criteria&oldid=299782911, version as of 2009-07-10, 2009.

[30]   Wikimediaspellchecker. A spellchecker for MediaWiki. Available online at http://code.google.com/p/wikimediaspellchecker/, 2009. (last viewed on 2009-07-11)

[31]   Wolkensteiner, G.: *Die Qualität von Wikipedia im Vergleich zu traditionellen Enzyklopädien [The quality of Wikipedia compared to traditional encyclopedias]*. Seminar paper, Wirtschaftsuniversität Wien, Institut für Informationswirtschaft. 2006.