TOWARDS ASSESSING INFORMATION QUALITY IN KNOWLEDGE MANAGEMENT IN THE ENTERPRISE 2.0

(Practice-oriented paper)

IQ Practices: Case Studies and Experience Reports

Sven Ahlheid
Siemens AG
sven.ahlheid@c-lab.de

Therese Friberg
University of Paderborn
friberg@cik.upb.de

Gernot Graefe
Siemens AG
gernot.graefe@c-lab.de

Alexander Krebs
Siemens AG
alexander.krebs@c-lab.de

Jan-Philipp Müller
Siemens AG
jan-philipp.mueller@c-lab.de

Dirk Schuster
Siemens AG
dirk.schuster@c-lab.de

Abstract: With regard to the success stories of Web 2.0 based knowledge centers such as the online encyclopedia Wikipedia [54] companies have begun to enrich their corporate knowledge management with Web 2.0 technologies, hoping to benefit from increasing flows of information. Besides information quantity, the quality of information is a key factor determining the return on investment of such Enterprise 2.0 platforms. In this context we will discuss requirements for the concept of information quality, identify important differences to the Web 2.0 environment and also elaborate on the basic design of a system assessing information quality in an Enterprise 2.0 context. We will thereby integrate implicit user feedback and explain the key benefits of this novel approach.

Key Words: Enterprise 2.0, Information Quality, Automated Information Quality Assessment, Implicit Feedback, Explicit User Feedback

1. MOTIVATION

The Internet and its users have evolved rapidly in the last years. While the first users of the World Wide Web were mostly consumers of information, Web 2.0 technologies have stimulated users to become suppliers of content [39]. Be it on personal blogs or within communities, the success of Web 2.0 platforms has changed the Internet fundamentally. Web 2.0 has reduced the barriers that prevented many people from providing content on the Internet to virtually zero and therefore significantly increased the active participation of a world wide community of Internet users. Consequently it has increased the amount of intelligence that can be collected and disseminated [52].

With regard to the success stories of Web 2.0 based social networks such as Facebook [13] and knowledge centers such as the online encyclopedia Wikipedia companies have begun to enrich their corporate knowledge management with Web 2.0 technologies ([28], [33]), hoping to benefit from increasing flows of data.

These so-called Enterprise 2.0 approaches have led to an unprecedented number of contributions to the knowledge management systems in many companies. However, in order for knowledge to constitute an essential competitive factor, the intellectual capital within a company ought to be created, organized and used ([9], [14]). A mere increase of information frequently does not render any additional value. Business intelligence can be overwhelming or, at worst, useless without a way to navigate through, process, interpret and share data [52]. With a growing amount of data the organization of information within a knowledge management system becomes increasingly difficult. Yet it is crucial for the success of a company that all information is being noticed by any employee needing it, prior to any decision taken [44]. Hence, the information in question should be both accessible and findable, the easier the better.

Furthermore, the simplicity of creating and manipulating data has led to an increase in the user's uncertainty about the correctness of the information within the system [15]. This uncertainty may lead to the following situation: although information is generated and collected (creation) as well as sensibly filed and distributed to where it is needed at the correct point in time (organization), employees still decide against using the provided information (usage) ([21],[42]). Evidently, Enterprise 2.0 based knowledge management systems are only successful, if the quality of the knowledge within the system can be determined. Many corporate success stories have illustrated how Enterprise 2.0 approaches can lead to Enterprise 2.0 knowledge centers with an abundance of very reliable information. Moreover, research studies on Web 2.0's prime example Wikipedia have shown that user generated content can develop into highly accurate and valuable information ([18], [23], [53]).

Whereas these research studies on Wikipedia point out the potential of user generated content, launching a corporate wiki does not consequently emerge to be equally successful. Though the easy contribution and usage of information is mandatory, it certainly will not be sufficient. Most noticeable, there is a difference in the number of (potential) contributors. Whilst there are more than 35,000 contributors (with at least five edits per month) active for the English Wikipedia [57], many companies not even have such many employees at all. Another difference stems from the motivation of contributors. While Wikipedia contributors are largely intrinsically motivated, this cannot be expected at a similar high degree for company employees ([17], [34], [43]). Likewise to traditional corporate knowledge management, contributing information lacks quantifiable benefits and thus companies will prefer revenue generating activities.

Generally there will no or very few dedicated knowledge workers be appointed whose sole responsibility is to manage the corporate knowledge pool. Even if there were such employees, they could not supply the information needed as it accrues in the general business processes. After all, it is the advantage of successful Enterprise 2.0 knowledge management approaches that many employees contribute their

personal knowledge, experiences and know-how.

Thus, the benefit of a Web 2.0 based corporate knowledge management platform is highly dependent on the additional amount of workload necessary to obtain a *useful* information resource. One should therefore think twice, before implementing any step that increases the workload on platform users. This applies to the effort necessary to contribute information as well as to making use of any information quality assessment mechanisms incorporating judgments of the users. Currently many Web 2.0 based information platforms that implement some kind of quality assessment rely on explicit user contributed rating. By "explicit user feedback" we refer to input asked from users of the information, with the explicit indication that some kind of information quality assessment is requested. Explicit user feedback is widely used in the Web 2.0. The input is typically requested with phrases like "was this information helpful" or "rate this article", but this term also refers to a full-blown peer review process. We will elaborate more on this in section 3.

Though more or less useful on Web 2.0 platforms, we question the usefulness of an employment of explicit user feedback as the sole approach in the Enterprise 2.0 context. Instead we propose a hybrid system consisting of explicit and implicit user feedback as well as an automatic information quality assessment. We will describe this hybrid approach for evaluating information quality in a corporate context in more detail in section 4. The term "implicit user feedback" in this context denotes user input which is not indicated as being used for quality assessment. Implicit user feedback is also widely utilized in Web 2.0, mostly in the form of the number of views of items (e.g., articles, videos, etc.) or the number of references to an item. These numbers might be presented solely as additional information, but often they are used in terms of "word of mouth" – as an indication of information quality and hence are drawn on for ranking purposes. Google's PageRank algorithm is an example of a technology that incorporates implicit user feedback in order to determine information quality as it counts the number of links (at least partly set by humans) to a certain web resource [41].
"Automatic information quality assessment" then again designates any algorithm that determines information quality metrics from attributes of the document, the author and other sources; whether any user has read the article or not. In other words, the algorithm only uses data available at the instant the information item is disseminated. To our knowledge an automatic evaluation of information quality is not widely used in the Internet yet.

As a starting point we will review information quality criteria in the Enterprise 2.0 knowledge management context in the following section.

2. DATA QUALITY CATEGORIES AND DIMENSIONS

In order to describe information quality we will adopt the broad data quality conceptualization of Wang, Strong, and Lee [47]. They define high-quality data as data that is fit for use by data consumers. Using this definition, the characteristics of high-quality consist of four categories as shown in Table 1.

Data Quality Category	Data Quality Dimensions
Intrinsic Data Quality	Accuracy, Objectivity, Believability, Reputation
Accessibility Data Quality	Accessibility, Access Security
Contextual Data Quality	Relevancy, Value-Added, Timeliness, Completeness, Amount of Data
Representational Data Quality	Interpretation, Ease of Understanding, Concise Representation, Consistent Representation

Table 1. Data quality categories and dimensions

We chose to apply this conceptualization, as it is broader than most other conceptualizations that treat information quality merely as an intrinsic concept. Their concept takes into account that the quality of information cannot be assessed independent of the people using them – a view that is well accepted in the literature on information quality [10]. Therefore, this conceptualization is more suitable for our purposes than others that do not consider information quality beyond the intrinsic view.

Comparing information quality in our context to information or data quality in the context of business intelligence, exhibits a number of peculiarities which lead to a different weighing of some of the data quality dimensions. In the context of data mining and business intelligence, information or data quality often refers to quality of data structured in databases, intended for machine processing. In contrast, in our context information typically consists of texts, but also images or even audio or video data intended to be used by human beings. In this situation criteria related to perception of information quality by the data consumer like representational data quality gain importance and have to be modeled.

Furthermore, the kind of information we refer to is typically not structured in a formal manner, as it is e.g., in an address database. Lacking a structure means that the semantic information inherent in the data model is *a priori* not available in our context. To a human data consumer this information is apparent from the context and content of the information, but is often difficult to obtain for a machine. So in analyzing information quality – not limited to the automatic assessment – one either has to dispense with these metadata or generate them e.g., by text analytic methods.

At the same time, due to superior human error correcting capabilities, the weights of certain information quality dimensions might diminish in the overall information quality construct when dealing with unstructured data. Statements like “Vegetarian pizza $8,75” would be considered faulty in the business intelligence context, but often not even recognized as incorrect by humans (decimal sign should be a dot not a comma, see also [36]).

3. CURRENT RATING SYSTEMS ON WEB 2.0 PLATFORMS

Information quality in Web 2.0 is widely addressed. Many of the sites implement some kind of rating for user contributions. To the best of our knowledge, so far, only explicit user feedback is employed. Below, we present a compilation of rating concepts employed on three big knowledge-oriented platforms.

3.1. Wikipedia

For a better understanding of the foundations of Wikipedia's rating system, it is worth to outline its history. Today Wikipedia is well known as a multilingual encyclopedia and one of the prime projects in the participatory evolution, termed Web 2.0. Wikipedia was launched in 2001 as an offshoot of Nupedia, an editorial managed encyclopedia. As articles in Nupedia were written by highly qualified contributors only and had to be reviewed before publishing, the platform grew very slowly. Wikipedia was launched to solve this problem and let the users create content, with the goal of transferring it to Nupedia, when a certain quality standard was reached. This was a revolutionary idea at that time and there were many critics warning about abuse. But Wikipedia was highly accepted by the community and quickly grew to a sufficiently high quality level, so Nupedia was abandoned already in September 2003. Currently Wikipedia is one of the largest websites with 65 million visitors monthly (in 2009), 75,000 active contributors and 13 million articles in 260 languages, thereof more than 2.9 million in English [56]

Considering the historical root of Wikipedia, it is not surprising that rating in Wikipedia is done in completely manual process. The rating system offers three grades of rating: no-rating, 'good article' and 'featured article' and there exist explicit descriptions what comprises a 'good' and a 'featured' article [58]. Only registered users can be involved in the rating process. Registration, however, is open to anybody. Users may nominate articles for either grade starting the respective reviewing process. Any user is allowed to review an article for the 'good' rank, whereas 'featured' articles are reviewed by editors only. When reviewing is finished, an article may be promoted to the 'featured candidate' stage which in a voting process leads to the final decision. The whole process of becoming a featured article is controlled by a 'featured article director' and lasts a few weeks.

3.2. Google Knol

Google Knol was launched in 2007 as a Web 2.0 based knowledge platform. In contrast to Wikipedia, it is not intended as an encyclopedia, but rather a collection of articles by individual authors. Consequently, multiple articles on the same topic are allowed and it is up to the authors to consent to collaborative editing. Google Knol offers a name verification process which may be used by authors to increase credibility. Authors may also decide to include advertisements in their articles (in the form of Google Ads) in order to receive revenues for their writing.

Google Knol has implemented a rating system that offers a variety of interactions. There is a basic five-star rating with an unspecific overall rating of the article as it is widely used in Web 2.0 platforms nowadays. This kind of rating is open to every visitor of the platform. In addition, awards are given to articles with high click rates, to high rated articles as well as to authors with many high rated articles. Users are also randomly confronted with particular information quality related questions (e.g., about authenticity or reputation of the author). These questions are only to be answered by a yes / no / neutral scheme. Another information quality related interaction is the possibility to publish comments which, in addition to the text of the comment, may be classified by '+' or '-'. Even a full review of the article may be done, where the reviewer is asked a number of questions about e.g., the author's reputation or whether the article is considered to a thorough treatment of the topic. However, it is up to the author to allow for comments or unsolicited reviews.

Eighteen months after launch only few articles have more than some tens of ratings. Numerous articles, some of them being on the platform for one year or more and, although frequently viewed since publication, have no ratings or comments at all, even general knowledge articles on science topics (e.g., [6], [16], [32]).

3.3. Helium

Helium is a knowledge platform, where writers are contributing articles and are paid for this based upon advertising revenues. The site started in October 2006 with a concept similar to Yahoo! Answers, where questions were to be answered by a community. In 2007 Helium changed to its current format similar to Google Knol – Helium, however, was first. Content is provided free of charge. According to Wikipedia [55] (site statistics are no longer provided), in February 2008 more than 100,000 users have been registered with Helium.

The platform also relies on a rating system based on explicit user feedback which, however, has a unique feature (at least among the three sites discussed here): users are asked to compare the information value of two documents on the same topic using a six level scale. In addition, a user may flag an article as 'self promotion', 'off-topic', 'plagiarism', 'adult', 'profane', or 'illegal'. The platform also supports direct feedback to the author.

Finally, there is an abundance of ranks and badges a user may obtain. These rewards ('stars') are allocated based on the user's number of contributions, i.e., the number of articles written and/or rated. In contrast to many other sites (e.g., forums), not the total number of contributions is taken into account, but the number of a user's contributions relative to the community standard. Besides, only those articles are counted that are in competition to others. Badges are provided for different activities like selling articles to professional publishers via the platform or winning writing contests, but also for verified members of certain professions (journalists, medical, psychological, educational, or legal professions, financial advisors and architects).

3.4. Comparing Public Review with Dedicated Peer Review

The Web 2.0 websites discussed above employ explicit user feedback in several variations as means to assess the information quality of user generated content. Another well-proven information quality assessment method beyond the Web 2.0 environment is peer review by dedicated reviewers, the standard quality assurance process in the scientific community. Summarizing, we have two basic approaches when it comes to an explicit quality assessment: dedicated reviewers and public review.

Dedicated Reviewers

The dedicated reviewer approach is the one employed by the scientific community and also in some sense by Wikipedia for the featured article selection process. It may be deployed in an open (author knows reviewers and vice versa; typically used for the assessment in formal education), single-blind (typically only the reviewer knows the author) or a double-blind fashion. For the single- and double-blind methods, an editor is needed who takes care of the selection of reviewers, relays the article and the review result. A dedicated reviewer process may be performed in varying levels of profundity from a simple check of formal requirements to a complete reproduction of every detail of the work, analyzing the comprehensibility of the representation, etc. This is dependent on the task assigned to the reviewers as well as on their level of qualification and the resources available. Obviously, the dedicated reviewer approach puts high costs on some individuals (editors and reviewers) and consequently a trade-off between costs and benefits should be conducted.

Public Review

All websites presented above (and many others) employ public user review. For Wikipedia this holds true for the good article selection process. Compared to the dedicated reviewer approach, this method relaxes the requirements for the reviewer, compensating for this by the statistical effects of the "wisdom of crowds" [49]. To achieve a successful compensation it is necessary to consider that any statistical measurement requires a minimum sample size, dependent on the underlying probability distribution and

the statistical test employed. It is important to note that "sample size" in this context refers to the number of estimations of an article's information quality provided by (different) users.

Helium just ranks the articles, but does not give an indication about the underlying number of ratings. Google Knol, in contrast, presents the number of ratings. However, since rating is open to everybody and multiple votes by the same user cannot be excluded, the number of ratings and the sample size might be very different. Even if one assumes that the number of ratings represents the true sample size (i.e., no user rated the article more than once) only few articles have more than 100 ratings.

The number of ratings and sample sizes might also differ because of user misconception or misuse of the rating tool. Without clear indication that a rating of the information quality of the content is requested, users might as well assume that they are asked for their consent or dissent to an article's content in terms of an opinion poll. Even if the aim of the rating request is clear to the user, he or she might choose to rate content in a manipulative way in order to promote or demote its author.

Public review generates the least direct costs, but considering the indirect costs by the work provided by the users as well, it is questionable, whether public review has a better cost-benefit ratio than the dedicated reviewer approach. In an enterprise context, however, it might be hard to obtain management acceptance for a dedicated reviewer process, as it is nearly impossible to quantify the benefits or even compute the return on investment.

4. Assessing Information Quality in an Enterprise 2.0 Knowledge Platform Using a Hybrid Approach

As we have seen, assessing information quality by explicit user feedback in the Web 2.0 context is implemented in a number of places, but suffers from shortcomings. Unfortunately, in an Enterprise 2.0 context we are dealing with a rather small community, with rather little intrinsic motivation, with strong relations between the individuals and last but not least legal restrictions (work constitution, privacy rules). Thus, some important preconditions for an explicit user rating system are only marginally met.

Therefore, we propose to combine both explicit and implicit user feedback as well as an automated evaluation of information quality into a hybrid system. More precisely for each of these three basic approaches we will identify suitable information quality dimensions to be assessed by this respective approach. We do not strive for a mutually exclusive set of quality dimensions. If for some quality dimension more than one approach is deemed to be well suited, the results will be combined in the second step to obtain an overall information quality rating. This combination may either be some kind of mean value or in the form of online training for a classifier.

The remainder of the paper will be dedicated to the discussion of this approach. In particular, we will elaborate in the following two aspects: what potential, with respect to an assessment of information quality, does each of the three constituents of the approach have and how are the parts supposed to merge together in order to provide the user with an estimate of the overall information quality of an item.

4.1. Automatic Information Quality Assessment

There is ongoing research on how to automatically predict certain information quality dimensions based upon formally accessible characteristical traits of a text. The probably oldest approach is the concept of readability indices, e.g. the Flesch-Kincaide Grade Level. Here, structural data (like average syllables per

word, or sentence length) is computed in order to predict the ease of understanding (see [11] for an overview).

A more sophisticated approach has been carried out by Dufty and others ([12], [20]) who measure cohesion in a text to likewise obtain a metric for the ease of understanding. Here, a thorough linguistic analysis including part-of-speech tagging, noun and verb phrase analysis, co-reference-analysis, dictionary lookups for polysemy and hypernyms (using WordNet) etc. is performed resulting in metrics to determine the ease of understanding. A somewhat related approach is the basic idea of Antiqueira et al. ([2], [3]) who pursue a graph theoretic approach utilizing word-adjacency (related to lexical cohesion of a text) and relate this to text quality.

Not surprisingly, Wikipedia has been object of a number of studies to assess information quality from quantitative metrics. Lih [31] proposes the number of edits of an article and the number of different users as a "simple measure for the reputation of the article" and relates these to news outlet citations of Wikipedia articles. Nevertheless, it is not clear, whether the usage of the term "reputation" in Lih's paper is identical to the respective data quality dimension (see Table 1). However, it probably refers to some combination of dimensions from the "Intrinsic Data Quality" category. At about the same time of Lih's publication, Wikipedia started to develop criteria for high-quality articles named "featured articles". Stvilia and others [8] have carried out a factor analysis on featured and random articles to find a set of information quality criteria explaining the quality model of featured articles. The resulting dimensions are: "authority/reputation" (objectivity, believability, reputation and concise representation), "completeness" (completeness), "complexity" (ease of understanding), "informativeness" (relevancy, value-added, completeness), "consistency" (consistent representation, accuracy), "currency" (timeliness) and "volatility" (in parentheses the corresponding data quality dimensions according to Table 1).

A similar analysis, but also somewhat related to the Coh-Metrix research by Dufty [12], was conducted by Blumenstock ([4] and [5]) who extracted more than 100 article features on four different levels: surface features (like words, sentences, syllables, etc.), structural features (e.g., links, images, categories, etc.), readability metrics as well as part-of-speech tags for a sample set consisting of about 11,000 random and featured articles. The results show that word count alone is already a very useful measure for classifying an arbitrary article as 'featured article'. The author indicated a usable discriminant of 1,830 words with an accuracy of more than 96%. However, it seems at least doubtful, whether this result could be applied in contexts other than Wikipedia, since keeping articles moderately sized is a basic Wikipedia rule for writing articles: "*Readers may tire of reading a page much longer than about 30 to 50 KB, which roughly corresponds to 6,000 to 10,000 words of readable prose*" [59] . Longer articles are suggested to be split up. Also among the guidelines for the nomination of good articles ("GA" in Wikipedia abbreviation, a lesser rating grade than 'featured articles'), article length is indicated as a criterion: "*Although there is no set guideline on article length for GAs, it is best for the article not to be too short or so long that there is not enough focus on the topic*" [60]. Short articles of dictionary style will be relocated to Wiktionary. So by these rules a shorter article with high information quality will less probably reach "featured article" status than a longer one.

Another approach, in its method even more Wikipedia related, is the analysis carried out by Hammwöhner [24]. He draws upon consistency of linking of articles and categories in different languages (by exploiting the cross-language links provided in the different language Wikipedias) with the aim to identify quality aspects.

The analysis of Ng, Tang and others ([37], [50]) carried out on news articles indicate two aspects: human perception of document quality is strongly dependent on the individual; document features obtained from a textual analysis will provide enough information to a classifier in order to predict particular quality

features as long as the documents are classified consistently (e.g. by the same judge).

Concluding, though there is no general well-proven automatic quality analysis method available, characteristics of a number of approaches are known in the literature. To build an automatic classifier, one might therefore consider the following guidelines:

- Avoid a "general quality mark", instead concentrate on quality dimensions, which different individuals rate with little variance, following Ng and Tang [37] conclusions.
- Try to identify linking and other structural data which might indicate (possibly only a lack of) certain quality aspects (consistency, relevancy, believability, timeliness), thereby adopting Hammwöhner's approach [24] or Lih's [31] criteria.
- Select document features by statistic or linguistic analysis as input to a classifier in order to identify certain quality aspects, with an emphasis on "ease of understanding" and "concise representation". This suggests the well-established readability metric research as well as research by Dufty and others [12], Stvilia and others [48] and Blumenstock [5].

Certainly, a classifier will look differently for a corporate wiki than for repositories of project references, customer information or research articles.

4.2. Implicit User Feedback

Another method to assess information quality of a particular dimension is to include implicit user feedback. Implicit feedback is derived from observing the interaction of users with an Internet platform. In the respective literature implicit feedback has been researched in-depth to improve methods to search for objects ([1], [62]) or to replace time-intensive explicit user feedback such as ratings and questionnaires [38]. All user activities belonging to the implicit feedback assume the consumption of an information item and deduce an assessment with respect to its information quality based on the concrete type of activity that a user performs. In general it can be assumed that the more interesting (thus relevant) a piece of information is to a consumer, the longer will he or she spend his or her time with it.

It is suggested to assess the quality of information by means of user activities [40]. The authors align the different actions according to the classification of [38] into four behavior categories:

- **Examination**
  This category contains the direct consuming activities of a user. The authors assign the activities "listening", "selecting" and "viewing" to this category which often is described as "click stream" in the context of the Internet [40]. The activities in the "Examination" category define a direct occupation with an information item.
- **Retaining**
  The second category consists of all activities with the goal of future use of an object. [40] mention the activities "bookmark", "delete", "print", "purchase", "save" and "subscribe". Noticeable is the action "delete" which implicates a negative assessment of the respective information item in contrast to all other activities mentioned before.
- **Reference**
  This category deals with connections between different pieces of information. Oard and Kim assign the actions "cite", "copy-and-paste", "forward", "link", "quote" and "reply" to this category. With these actions one object is connected to another one by the user.
- **Annotation**
  In this category [40] summarize activities through which the users invest additional effort in a certain piece of information; value is added to an object. Thus, one can find the activities "mark up", "organize", "publish" and "rate" in this category.

Clearly, it is possible to include further actions into the four categories. For example, in the "examination category" one could add mouse clicks and scrolling (like [7]); even the reading time is an indicator for the interest in an information item as [8] and [35] have found out. [27] point out that by recording the printing of a document it is possible to detect more relevant items, because most of them would not have been identified when taking reading time into account. Furthermore, the analysis of clicking an object is a good prediction of the relevance of the document [25]. In the same way the "annotation" category could be amended by tagging (like [19]) because in the literature it is suggested that assigned tags assist the user to manage bookmarks, links, etc. more effectively and efficiently. In general, recommendation systems relying on the invested time of users can almost be as correct as systems that are based on explicit user feedback [29]. This awareness fosters research in the area of implicit feedback even more.

[30] utilize the connection between information items for an assessment of the data quality dimension relevancy. Also the above mentioned PageRank algorithm allocates the usefulness of such an analysis [41]. Here, a source is defined as valuable when it points to a lot of valuable targets. A good target is recognized if a lot of sources utilize it. Both approaches are assignable to the "reference category".

One type of activity needs further and more detailed explanation: bookmarking. [45] have developed a recommendation system based on bookmarks of the users. No other action of the implicit feedback correlates as highly with the interest of a user as bookmarking [46]. [22] show that search engines based on recommendation of bookmarks are an alternative to search engines based on algorithms.

The rating activity in the "Annotation" category is a summary of all kinds of explicit feedback. In the context of the implicit feedback, solely the fact that a user is willing to rate an item is of interest. By rating a piece of information it is assumed that the user has engaged in it before ([51] and [26]). The detailed examination of the rating (the so-called explicit user feedback) is explored in the following paragraph as the third constituent of our proposed hybrid approach to assess information quality in an Enterprise 2.0 knowledge platform.

4.3. Explicit User Feedback

At first sight it might look trivial to establish an explicit user feedback or rating system for an Enterprise 2.0 type knowledge center. However, there are a number of constraints in this context which have to be obeyed in order to build a successful system:

- **Return on investment**
  Web 2.0 based knowledge platforms are sometimes (like Web 2.0's prime example Wikipedia) non-profit charity projects which are funded by donations and supported by volunteers, while others base their economies upon advertisement. Both models do not work in a corporate setting. At least in the long-run any such project has to prove its return on investment to the management. While the benefits and revenues generated by such a platform are hard to measure directly, it is important to keep the costs for running such a platform low. This includes the direct running costs, but also the indirect costs like expenditure of time for contributing to the platform (writing articles and/or giving feedback on the quality of the content). If not limited by explicit management decisions, every contributing employee has to justify the time dedicated to the platform with respect to other direct revenue generating activities.

- **Law and regulations**
  User feedback might be considered as a kind of employee appraisal. Depending on the local laws, labor management or union contracts, there might be restrictions on collection, storage and presentation of such ratings. These restrictions also apply to tracking employees' click-streams in order to derive implicit user feedback. This requires additional considerations by global enterprises. Not being experts in this field, we will not consider regulatory constraints subsequently.
- **Restricted number of users**
  Compared to the successful Web 2.0 projects like Wikipedia, Facebook or YouTube the number of potential users in a corporate context is rather limited, even in the biggest companies. Especially if one takes into account that the information contained in such a knowledge repository is highly partitioned, thus every part is only relevant to a smaller fraction of the staff. A procurement manager might not be able to assess the quality of development-related information. Field service workers might not really evaluate information relevant to operations.
- **Reluctance, harassment**
  On the one hand, being in a tight and economically important relationship, employees might be reluctant to express their opinion freely. The relationship to their superiors, to co-workers or other employees might be affected by such action. On the other hand, information quality ratings might also serve as a medium for harassment at work.

However, there are also a number of factors in favor of the Enterprise 2.0 context. In particular, blatant vandalism [61] is probably less relevant.

When considering explicit user feedback one has a set of questions related to the different quality dimensions (see Table 1) to ask and expects the users to answer these questions. The rating will be performed by some form of statistical evaluation. However, when employing statistics upon user feedback, one should be aware of the restricted number of users. Consequently also the small number of expected replies in an Enterprise 2.0 context should be considered. Thus, the questions should be limited in their amount and easy to understand and answer. This also helps reducing the variance in the answers which in turn decreases the necessary sample size. Questions about facts ("Does the article contain enough references?") rather than about impressions ("Is the article plausible to you?") serve the same goal.

Considering what has been said about harassment and reluctance to answer a carefully designed presentation of the results might help to reduce this impact.

Finally, in order to guarantee a high usage of the explicit rating mechanism by the employees, it is necessary for the company to implement adequate incentive mechanisms for participation. According to the theory of rational choice, incentives have to pertain to the individual's self-interest to be effective. For instance, such incentive might be the fact that employees participating in the rating process receive detailed statistics on their overall engagement within the knowledge platform, which otherwise is only known to the organizational unit operating the platform. The employees might then decide, whether they reveal this data to their superiors, e.g., in the context of an annual evaluation talk, as means to illustrate their efforts towards the company. Another benefit for the individual employee of actively rating content on an enterprise 2.0 knowledge platform might be seen in the organization of the individual knowledge space that is accompanied by providing explicit feedback to other users: by evaluating the information quality of different articles on the platform, employees simultaneously provide structure to the different pieces of information relevant to their daily work. To secure the effectivity of this incentive, it is necessary that users of the platform are offered the option to bookmark or add links to different pieces of content to their individual user profile.

4.4. Interplay of the three Different Approaches

In the discussion of the research on automatic, implicit and explicit user feedback approaches and their respective limitations in an enterprise context, for each of the three mechanisms we observed a focus on some specific subset of the complete information quality model (see Table 1). As we have seen, the automatic assessment approach emphasizes the quality dimensions “ease of understanding”, “consistent representation”, and possibly “timeliness”. Hence, primarily the representational and contextual quality categories are covered. The implicit user feedback component then again with its “examination” category reflects the quality dimensions “interpretation” and “ease of understanding”. Regarding its “retaining”, “reference”, and “annotation” categories the dimensions “believability”, “reputation”, “relevancy”, and “value-added” are addressed. The explicit user feedback approach conceptually could yield ratings for any quality dimension. However, the limitations we have discussed above suggest reducing the number of quality dimensions is necessary.

Both the automatic and implicit user feedback components yield attributes of the information item, but do not directly produce metrics for the different information quality dimensions. This will be the task for the corresponding classifiers using these attributes as input data. For such classifiers to provide valid and thus useful metrics, besides input data we also need training and feedback. While certain training can occur during the setup process, it is conceivable that this training possibly does not suffice to cope with changes occurring in the enterprise. Here we need the explicit user feedback. The explicit feedback component is thus supposed to (a) produce quality ratings for otherwise unavailable dimensions and to (b) provide the mentioned feedback data for the classifiers of the other components.

The design of such a hybrid approach to assess information quality in an Enterprise 2.0 based knowledge platform thus requires a thorough analysis of the specific information quality requirements for the particular repository, be it a wiki or a platform for project references, customer information or research articles.

5. SUMMARY AND MANAGEMENT IMPLICATIONS

We discussed the necessities, requirements and constraints as well as the resulting possible approaches to assess information quality in an Enterprise 2.0 based corporate knowledge platform. Our conclusion is as follows: we propose a hybrid approach consisting of an automatic information quality assessment that is complemented by implicit as well as explicit user feedback. We believe that such a hybrid approach is likely to result in the highest return on investment in a corporate context. For each of the three constituents of this hybrid approach we reviewed the literature with respect to possible implementation approaches and discussed these by taking the corporate context into account.

In our opinion, (a) automatic information quality assessment that is enriched, in particular, by implicit user feedback as well as (b) cost-benefit analyses of an assessment of information quality in corporate contexts (including the indirect costs resulting from time spent rating content) are fruitful avenues for empirical studies on Enterprise 2.0.

Acknowledgements
Research has been supported by Bundesministerium fuer Wirtschaft und Technologie (German Federal Ministry of Economics and Technology) as part of the THESEUS research program (grant number 01MQ07014).

REFERENCES

[1] Agichtein, E., Brill, E., Dumais, S., "Improving web search ranking by incorporating user behavior information" *Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval, August 06-11*, 2006. Seattle, Washington, 2006.

[2] Antiqueira, L.; Graças, M.; Nunesm V.; Oliveira O. N.; Da F. Costa, L. "Strong correlations between text quality and complex networks features*" Physica. A*, 373, 2007. pp. 811-820

[3] Antiqueira, L.; Graças, M.; Nunesm V.; Oliveira O. N.; Da F. Costa, L., "Complex networks in the assessment of text quality". arXiv:physics/0504033v1 [physics.soc-ph]. 2008.

[4] 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*, 2008. pp. 1095-1096.

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

[6] Gary "Chemical Reaction", http://knol.google.com/k/gary/chemical-reaction/ow5jbvr76bz9/89# (viewed on 2009-06-26)

[7] Claypool, M., Le, P., Wased, M., Brown, D. "Implicit interest indicators" In: *Proceedings of the 6th international Conference on intelligent User interfaces (Santa Fe, New Mexico, United States, January 14 - 17, 2001). IUI '01*. ACM. New York, NY, 2001. pp 33-40.

[8] Cooper, M. D. "Predicting the relevance of a library catalog search" *J. Am. Soc. Inf. Sci. Technol.* 52, 10. 2001. pp 813-827

[9] De Long, D.W., Fahey, L. "Diagnosing cultural barriers to knowledge management" *The Academy of Management Executives* 4. 2000, pp. 113-127

[10] Deming, E. W. *Out of the Crisis*. MIT Center for Advanced Engineering Study. Cambridge., Mass., 1986.

[11] DuBay, W.H.: *The Principles of Readability*. Impact Information, Costa Mesa, Calif., 2004.

[12] Dufty, D., McNamara D., Louwerse M., Cai Z., Graesser A. "Automatic Evaluation of Aspects of Document Quality" In: *Proceedings of the 22nd annual international conference on Design of communication: The engineering of quality documentation.* 2004.

[13] Facebook, Inc.: http://www.facebook.com

[14] Frey, B.S., Osterloh, M.: *Managing Motivation: Wie Sie die neue Motivationsforschung für Ihr Unternehmen nutzen können*.. Gabler, Wiesbaden , Germany, 2002

[15] Friberg, T., Reinhard, W.. „Ergebnisse einer empirischen Studie zu Informationsqualitätskriterien in Corporate Wikis." Preprint, 2009.

[16] Frish, Petr: Newton's Laws - The Concepts.http://knol.google.com/k/petr-frish/newtons-laws-the-concepts/3trm10yysobsi/2# (reviewed on 2009-06-26)

[17] Garnefeld, I., Iseke, A.: Krebs A. "(Not) Just for the Money – Effects of Incentives on Member Participation in Online Communities." *AMA 2009 Summer Marketing Educators' Conference, August 7-10,* Chicago, Illinois. 2009 (accepted)

[18] Giles, J. "Internet encyclopaedias go head to head." *Nature,* 43,.2005. pp.900–901

[19] Golder, S., Huberman, B. "The Structure of Collaborative Tagging Systems" *Journal of Information Science* 32(2). 2006. pp.198-208

[20] Graesser, A.C., McNamara, D.S., Louwerse, M.M., Cai, Z. "Coh-metrix: analysis of text on cohesion and language." *Behav Res Methods Instrum Comput*. 36(2). 2004. pp.193-202

[21] Graefe, Gernot: Incredible Information on the Internet: Biased Information Provision and a Lack of Credibility as a cause of Insufficient Information Quality, in: Eppler, Martin J. / Helfert, Markus (Eds.): Proceedings of the 8th International Conference on Information Quality at MIT, 2003, pp. 133-146.

[22] Graefe, Gernot / Maaß, Christian / Heß, Andreas: Alternative Searching Services: Seven Theses on the Importance of "Social Bookmarking", in: Auer, Sören / Bizer, Christian / Müller, Claudia / Zhdanova, Anna V. (Eds.): The Social Semantic Web 2007: Proceedings of the 1st Conference on Social Semantic Web (CSSW), GI Edition, Lecture Notes in Informatics, 2007, pp. 11-21.

[23] Hammwöhner, R., Fuchs, K.-P., Kattenbeck, M., und Sax, C. „Qualität der Wikipedia. Eine vergleichende Studie." *Internationales Symposium Informationswissenschaft, ISI 200,* Köln. 2007. (accepted)

[24] Hammwöhner, R. „Interlingual Aspects if Wikipedia's Quality".In *12th International Conference on Information Quality*. 2007.

[25] Joachims, T., Granka, L., Pan, B., Hembrooke, H., Radlinkski, F., Gay, G. "Evaluating the accuracy of implicit feedback from clicks and query reformulations in Web search." *ACM Trans. Inf. Syst*. 25,), 7. 2007.

[26] Kelly, J. D. *Understanding Implicit Feedback and Document Preference: a Naturalistic User Study*. Doctoral Thesis. UMI Order Number: AAI3117613., Rutgers University. 2004.

[27] Kim, J., Oard, D. W., Romanik, K. *Using implicit feedback for user modeling in internet and intranet searching.* University of Maryland CLIS Technical Report 00-01. 2000.

[28] Koch, M., Richter, A.: *Enterprise 2.0 - Planung, Einführung und erfolgreicher Einsatz von Social Software in Unternehmen.* Oldenburg Wissenschaftsverlag, München, Germany. 2008.

[29] Konstan, J., Miller, B., Maltz, D., Herlocker, J., Gordon, L., Riedl, J. "GroupLens: applying collaborative filtering to Usenet news" *Communications of the ACM*, 40 (3). 1997. pp.77-87

[30] Lawrence, S., Bollacker, K., Giles, C. L. 1999. "Indexing and retrieval of scientific literature." In *Proceedings of the Eighth international Conference on information and Knowledge Management (Kansas City, Missouri, United States, November 02 - 06, 1999)*. S. Gauch, Ed. CIKM '99. ACM, New York, NY. 1999. pp.139-146

[31] Lih, A "Wikipedia as Participatory Journalism: Reliable Sources? Metrics for evaluating collaborative media as a news resource." *5th International Symposium on Online Journalism. April 16-17, 2004.*

[32] Lo, Glenn: "The Elements" http://knol.google.com/k/glenn-lo/the-elements/28prei6i5vxn0/9# (reviewed on 2009-06-26)

[33] McAfee, A.: "Enterprise 2.0 – The Dawn of Emergent Collaboration." In: *MIT Sloan Management Review* 47 (3). 2006. pp. 21-28

[34] Mergel, I., Reimann, M. „Anreizsysteme für Wissensmanagement in Unternehmensberatungen." *Wissensmanagement*, 2(4). 2000. pp.15-19

[35] Morita, M., Shinoda, Y. Information filtering based on user behavior analysis and best match text retrieval. In Proceedings of the 17th Annual international ACM SIGIR Conference on Research and Development in information Retrieval (Dublin, Ireland, July 03 - 06, 1994). W. B. Croft and C. J. van Rijsbergen, Eds. Annual ACM Conference on Research and Development in Information Retrieval. Springer-Verlag New York, New York, NY. 1994. pp.272-281

[36] Naumann, F. , Sattler, K.-U.: "Information Quality: Fundamentals, Techniques, and Use". 2006. p.7. Avaliable at http://www.hpi.uni-potsdam.de/fileadmin/hpi/FG_Naumann/publications/EDBT06Tutorial_IQ.pdf (rechecked on 2009-06-18)

[37] Ng, K. B., Kantor, P., Strzalkowski, T., Wacholder, N., Tang, R., Bai, B., Rittman, R., Song, P., and Sun, Y. "Automated judgment of document qualities: Research Articles." *J. Am. Soc. Inf. Sci. Technol.* 57, 9 .2006. pp.1155-1164

[38] Nichols, D. "Implicit ratings and filtering." In *Proceedings of the 5th DELOS Workshop on Filtering and Collaborative Filtering, Budapest, Hungary 10-12, ERCIM*. 1997.

[39] O'Reilly, T. What is Web 2.0 - Design Patterns and Business Models for the Next Generation of Software. 2005. Available at http://oreilly.com/pub/a/web2/archive/what-is-web-20.html?page=1 (rechecked on 2009-06-28)

[40] Oard, D., Kim, J. "Modeling information content using observable behavior." In: *Proceedings of the ASIST Annual Meeting*. 2001. pp.481-488

[41] Page, L., Brin, S., Motwani, R., Winograd, T. *The page rank citation ranking: Bringing order to the Web.* Technical report, Stanford Digital Library Technologies, Project. 1998.

[42] Rese, Mario / Graefe, Gernot / Herter, Valerie: Relevance as an Information Quality Problem in Corporate Decision-Making Processes, in: Eppler, Martin J. / Helfert, Markus / Pernici, Barbara (Eds.): DIQ'04 International Workshop on Data and Information Quality at the 16th Conference on Advanced Information Systems Engineering, Riga (Latvia), 2004, pp. 25-36.

[43] Rheinberg, F.: Intrinsische Motivation und Flow-Erleben, in: Heckhausen, J., Heckhausen, H. (ed.). *Motivation und Handeln*, 3rd ed., Heidelberg, German. 2006. pp.331 -354

[44] Ritsch, K. *Wissensorientierte Gestaltung von Wertschöpfungsnetzwerken*, Dissertation, Graz. 2004.

[45] Rucker, J., Polanco, M. J. „Siteseer: personalized navigation for the Web." *Commun. ACM* 40, 3,. pp.73-76

[46] Seo, Y., Zhang, B. "Learning user's preferences by analyzing Web-browsing behaviors." In *Proceedings of the Fourth international Conference on Autonomous Agents (Barcelona, Spain, June 03 - 07, 2000). AGENTS '00*. ACM, New York, NY. 2000. pp.381-387

[47] Strong, D., Lee, Y., Wang, R."Data Quality in Context" *Comm. ACM*, 40, (5).1997. pp. 103-110

[48] 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*. 2005.. pp. 442-454

[49] Surowiecki, James. *The Wisdom of Crowds*. Garden City: Doubleday, 2004

[50] 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.

[51] White, R. W. "Implicit feedback for interactive information retrieval." *SIGIR Forum* 39, 1. p.70

[52] Whitney, M. "Enterprise 2.0: What, Why and How" In *Enterprise 2.0 Conference, May 2009*. (to appear)

[53] Wiegand, D. „Entdeckungsreise: Digitale Enzyklopädien erklären die Welt." *c't*, 6/2007. 2007. pp.136–145

[54] Wikimedia Foundation, Inc.: http://www.wikipedia.org

[55] Wikimedia Foundation. http://en.wikipedia.org/wiki/Helium.com (last modification date 25 April 2009 at 02:29)

[56] Wikimedia Foundation. http://en.wikipedia.org/wiki/Wikipedia:About (last modification date 13 June 2009 at 12:35)

[57] Wikimedia Foundation. http://stats.wikimedia.org/EN/TablesWikipediansEditsGt5.htm, data for May 2006 (rechecked on 2009-06-17)

[58] Wikimedia Foundation. http://en.wikipedia.org/wiki/Wikipedia:Good_article_criteria and http://en.wikipedia.org/wiki/Wikipedia:Featured_article_criteria

[59] Wikimedia Foundation. http://en.wikipedia.org/wiki/Wikipedia:Article_size#Splitting_an_article (last modification date 17 June 2009 at 10:53)

[60] Wikimedia Foundation. http://en.wikipedia.org/wiki/Wikipedia:Guide_for_nominating_good_articles (last modification date 20 June 2009 at 00:58)

[61] Wikimedia Foundation. http://en.wikipedia.org/wiki/Wikipedia:Vandalism (last modification date 22 June 2009 at 15:14)

[62] Yanbe, Y., Jatowt, A., Nakamura, S., Tanaka, K. "Can social bookmarking enhance search in the web?" In *Proceedings of the 2007 conference on Digital libraries, June 18-23, 2007*,. Vancouver, BC, Canada. 2007. pp.107-116