Bus Inf Syst Eng 57(3):155–166 (2015)
DOI 10.1007/s12599-015-0378-0

RESEARCH PAPER

Managed Wikis

A New Approach for Web 2.0

Thomas Wöhner · Sebastian Köhler · Ralf Peters

Received: 1 July 2014 / Accepted: 6 November 2014 / Published online: 3 March 2015

Abstract Wiki projects can be edited by everyday web users directly within the web browser. Consequently, undesirable contributions like vandalism and spam cannot be ruled out. In this paper, Managed Wikis are introduced as a new approach to avoid such undesirable contributions. Editing rights are assigned according to author reputation, the quality of articles and the occurrence of patterns of suspicious edits. In the paper, the concept of Managed Wikis is evaluated by means of a simulation on the basis of Wikipedia data. The analysis proves that undesirable contributions are blocked effectively. In contrast, desirable contributions are rarely affected by the editing rights restriction. The concept of Managed Wikis addresses open as well as corporate wiki projects where undesirable edits cause significant harm. Furthermore, it can be applied to make traditional websites accessible for the web community.

Keywords Managed Wiki · WCMS · Wiki · Wikipedia · Vandalism · Reputation · Editing rights management · Web 2.0

Accepted after one revision by the editors of the special issue.

Dr. T. Wöhner (✉) · Dipl.-Wirtsch.-Inf. S. Köhler · Prof. Dr. R. Peters
Martin-Luther-University Halle-Wittenberg, Universitätsring 3, 06108 Halle (Saale), Germany
e-mail: thomas.woehner@wiwi.uni-halle.de
URL: http://ebusiness.wiwi.uni-halle.de/

Dipl.-Wirtsch.-Inf. S. Köhler
e-mail: sebastian.koehler@wiwi.uni-halle.de
URL: http://ebusiness.wiwi.uni-halle.de/

Prof. Dr. R. Peters
e-mail: ralf.peters@wiwi.uni-halle.de
URL: http://ebusiness.wiwi.uni-halle.de/

1 Introduction

Wikis are special web content management systems that allow web users to generate content collaboratively via web browsers. Since wiki syntax is simple, users can contribute without possessing any technical background (Cunningham and Leuf 2001; Hippner and Wilde 2005). Wikis are typically used in Web 2.0 to enable the content generation by a large crowd. In contrast, traditional web content management systems (WCMS) can only be accessed by a small group of privileged users (Cunningham and Leuf 2001).

The most famous and most successful wiki is the free online encyclopedia Wikipedia. Wikipedia is available in more than 285 different languages and contains in total more than 34 million articles (Wikimedia 2015). The largest Wikipedia is the English one with more than 4.7 million articles, followed by the Swedish and the German Wikipedia with about 1.8 million articles each (Wikimedia 2015). According to alexa.com, Wikipedia is listed within the top ten most visited websites worldwide and receives about two million page impressions monthly (Alexa 2015).

Due to its importance, Wikipedia demonstrates the benefits but also the drawbacks of wikis in comparison to traditional WCMS. On the one hand, the open editing model attracts a large number of volunteers that maintain and update wiki sites. Therefore, the wiki principle in general leads to a high quality of contents (Giles 2005). On the other hand, undesirable edits like vandalism, edit wars and contributions by opportunistic or inexperienced authors cannot be ruled out (Denning et al. 2005). Usually, such undesirable edits are revised within a very short time interval (Viégas et al. 2004). However, these edits cause significant costs such as additional computer resources and manpower for detection and correction. Besides open wikis

like Wikipedia, corporate wikis have been becoming increasingly important (Arazy et al. 2009). However, especially in corporate context, skepticism concerning the information quality can be a barrier for the adoption of wikis (Bhatti et al. 2011). Furthermore, authors are concerned that other users are able to edit their contributions without any control, and therefore a dynamic access control is suggested for corporate wikis (Holtzblatt et al. 2010).

In contrast to wikis, traditional WCMS are not faced with undesirable edits since they can only be accessed by experienced and trustworthy users. However, due to the restricted number of editors, such websites might be less comprehensive and updated less frequently.

Approaches that combine the benefits of WCMS and wikis are currently not known. To fill this gap, we chose the design science research paradigm for our study (Hevner and Chatterjee 2010). Goals of design science are the identification of practice-relevant problems and the creation of innovative artifacts that contribute to problem solving. Our study addresses the tradeoff between the facts that current systems for web content creation either can only be accessed by a limited number of privileged editors or bear the risk of being damaged by undesirable edits. As an innovative artifact we develop a new concept called Managed Wiki (MaWiki). A MaWiki refers to a kind of wiki that implements rules for assigning editing rights automatically with the goal to prevent undesirable edits. Therefore, on the one hand, MaWikis maintain the open editing model of conventional wikis, and, on the other hand, control the access of editors similar to traditional WCMS. We evaluate our new approach by means of a simulation using Wikipedia data. MaWikis can be applied in open or corporate wiki projects to reduce the costs caused by undesirable edits. Furthermore, MaWikis are also applicable to traditional websites to open the editing process for the internet community, involving only a low risk of damage by undesirable edits.

The paper follows the structure of design science publications suggested by Gregor and Hevner (2013). Section 2 discusses the related work and defines the addressed research gap. Section 3 describes the MaWiki concept in detail. In Sect. 4 we explain our evaluation method and subsequently present and discuss the results of the evaluation. Finally, the conclusion in Sect. 5 summarizes the paper and points out further research directions.

2 Related Work

In the last decade, research on wikis has become a new, highly active research direction. Most of the research is based on data from Wikipedia. Relating to this study, research on approaches to overcome quality problems in wikis is of special interest. To identify relevant literature we queried the scientific databases Google Scholar and ACM Digital Library using the keywords *wiki*, *Wikipedia*, *quality*, *vandalism*, *reputation* and its variants. We applied a backward and forward search to identify further relevant publications. According to our literature review, previous research has introduced three concepts to deal with quality problems in wikis: *automatic vandalism detection*, *automatic quality assessment* and *automatic reputation assessment.*

2.1 Automatic Vandalism Detection

Vandalism refers to edits that deliberately damage wiki sites, for example by inserting senseless text or deleting text passages arbitrarily (West et al. 2010; Mola-Velasco 2011). As Wikipedia demonstrates, vandalism can be a highly relevant phenomenon in wiki projects. According to the study of Potthast (2010), about 7 % of the edits in Wikipedia are classified as vandalism. Due to the open editing model, vandalism can be a relevant issue in all kinds of wikis. However, our literature review did not reveal empirical studies on the relevance of vandalism in wiki projects besides Wikipedia.

Research on vandalism detection develops metrics to automatically decide whether a given edit is vandalism or not. The metrics are based on different types of features:

1. Language-based features (i.e. number of sexual or vulgar words)
2. Text-based features (i.e. percentage of upper-case letters, length of the contribution, repetition of words)
3. Metadata (i.e. daytime of contribution, length of the editing comment)
4. Reputation-based features (i.e. vandalism rate of an author)
5. Article-based features (i.e. time period to the next contribution).

State of the art approaches for vandalism detection combine different features by means of machine learning algorithms (Potthast et al. 2008; West et al. 2010; Adler et al. 2011; Mola-Velasco 2011; West and Lee 2011). Some approaches that automatically detect and reverse vandalism are already used in Wikipedia Bots. The precision of these Bots is already high but the recall is still on a low level (Adler et al. 2011).

2.2 Automatic Quality Assessment

A significant number of research papers investigate approaches for an automatic quality assessment in wikis (e.g. Lih 2004; Dondio and Barrett 2007; Blumenstock 2008; Wöhner and Peters 2009). The subject of quality

assessment is to improve the transparency regarding the quality of an article automatically. In previous research numerous effective metrics are introduced. The metrics are based either on the editing history (Lih 2004; Wöhner and Peters 2009) of a given article (number of edits, number of authors, amount of persistent contributions, etc.) or on content-based features (Blumenstock 2008) of the latest article version (number of images, length of the article, …). In spite of many activities within this research area, approaches for automatic quality assessment are currently not used in Wikipedia.

2.3 Automatic Reputation Assessment

In comparison to the automatic quality assessment, the automatic reputation assessment is a less active research area. The goal of such approaches is to estimate the importance of a given wiki author. In this context, importance refers to the editing intensity of the authors as well as the quality of their contributions. Reputation assessment should motivate authors to contribute to the given wiki project frequently and in high quality. A further intention of reputation systems is to mark contributions of low-reputation authors (Adler et al. 2008). Using this approach, readers obtain advice regarding which facts they should trust.

Effective reputation metrics are introduced in Adler and Alfaro (2007), Javanmardi et al. (2010) and Wöhner et al. (2011). These metrics are based on the quality of edits, which is estimated by means of the survival time of contributions. The published metrics differ in their accuracy and complexity of calculation.

2.4 Research Gap

The approaches described above have two major goals. Firstly, quality as well as reputation assessment try to signal the quality of content. Secondly, vandalism detection is used to automatically detect and return low-quality contributions. However, current approaches do not block undesirable edits in advance. Therefore, temporary impairments of the content quality and effort for the correction of undesirable edits are not avoided. This research gap is addressed in our paper by means of the MaWiki concept. MaWikis block undesirable edits in advance and therefore reduce the costs caused by such contributions.

3 Managed Wiki

In this section we first present the goals and the general MaWiki concept. Subsequently, we introduce an approach to distinguish low- and high-quality contributions. This distinction provides the basis to evaluate the MaWiki concept. Finally, we suggest metrics and the corresponding set of rules for a concrete implementation of our MaWiki concept. We apply this implementation for the evaluation of our concept.

3.1 Goals and Concept

A given edit e of a wiki site typically comprises a number of contributions (c_1, …, c_n). In this context a contribution denotes the deletion or insertion of a word, whereas a word formally defines a sequence of characters between two blanks. MaWikis use a set of rules to decide whether the whole edit e is permitted for the given author a. The rules are designed with the aim of preventing low-quality contributions while high-quality contributions should not be affected. Depending on the real quality of contributions, the classification results listed in Table 1 are possible.

Consequently, goals of the MaWiki concept are a high true positive rate (*TPR*) and a high true negative rate (*TNR*). To achieve these goals, in this paper, we suggest a two-tiered set of rules which consists of the following three components:

- the *quality-based editing rights management* (QRM),
- the *reputation-based editing rights management* (RRM) and
- the *pattern-based editing rights management* (PRM).

The structure of the entire set of rules and the interdependencies of the components is illustrated in Fig. 1. In the following we explain the three components in detail.

3.1.1 Quality-Based Editing Rights Management (QRM)

To ensure that the wisdom of the crowd is not jeopardized and high-quality contributions are not blocked falsely, restriction of editing rights is applied carefully in our approach. Thus, a restriction of editing rights is only employed if low-quality contributions are likely and a significant reduction of the article quality is expected. Since low-quality contributions are performed on high-quality articles frequently (Wöhner and Peters 2009), QRM assesses the quality of articles firstly. If the quality of an

Table 1 Classification in MaWikis

		Quality of contributions	
		High quality	Low quality
Decision of MaWiki	Permit	True Positive	False Positive
	Block	False Negative	True Negative

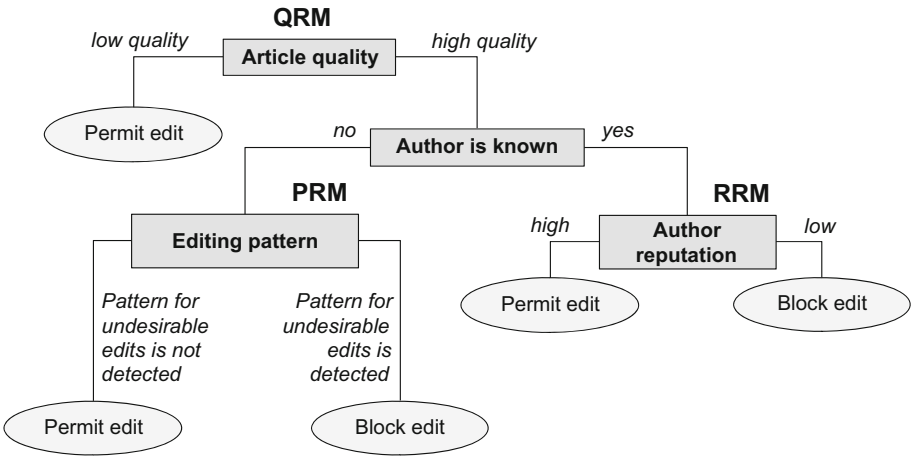


Fig. 1 Components of a MaWiki

article is below a defined threshold value, all edits will be permitted. For this group of articles, low-quality contributions do not lead to a significant quality reduction since the article quality is already on a low level. Moreover, the rule ensures that articles in development are not affected by the editing rights restriction and the open character of wikis is maintained.

3.1.2 Reputation-Based Editing Rights Management (RRM)

For articles of high quality the editing rights are controlled. The editing rights restriction is based on the author reputation, which is used to predict the quality of further contributions of the given author. However, the author reputation is only measurable for registered users that have performed at least one edit. In that case, RRM calculates the reputation value of the author on the basis of the previous edits. If the reputation values exceed a predefined threshold, the edit is permitted, otherwise it is blocked.

3.1.3 Pattern-Based Editing Rights Management (PRM)

The editing history of anonymous or newly registered authors is unknown, so that it is not possible to estimate the quality of their contributions ex ante. The only information for that kind of users is the intended contribution. Therefore, in case of an unknown reputation, the author enters the contributions and afterwards they are analyzed by means of machining learning algorithms for patterns of undesirable edits. If such a pattern is detected, the edit is not accepted. Theoretically, PRM is also applicable for registered authors. However, this leads to the effect that several contributions already entered are neglected. Therefore, to ensure the usability of the system, we consider PRM only for authors whose reputation is unknown.

3.2 Persistent and Transient Contributions

The MaWiki concept is evaluated by means of a comparison between the actual quality of contributions and the decision of the MaWiki set of rules (edit is permitted or blocked). Based on this comparison the *TNR* and the *TPR* of the system can be calculated. Consequently, the actual quality of contributions has to be known. Because of the extremely large number of edits in wikis, a manual assessment of the quality of contributions is infeasible. However, the editing history of a wiki site comprises implicit user-driven quality ratings. Approaches that quantify these implicit ratings have already been introduced in research papers on automatic quality and reputation assessment (Adler and Alfaro 2007; Javanmardi et al. 2010; Wöhner et al. 2011). The quality is derived from the persistence (survival time) of contributions. It is assumed that a long survival time indicates a high acceptance in the community and therefore a high quality of the contribution. As opposed to this, if the contribution is discarded quickly by the community, a low quality is supposed. The different approaches to assess the quality of contributions (Adler and Alfaro 2007; Javanmardi et al. 2010; Wöhner et al. 2011) vary in their computational complexity. One can suppose that these approaches are suitable to derive editing rights directly. However, since the quality can only be assessed after the contribution has already been processed, editing rights cannot be determined in advance.

In this work, we apply the approach of Wöhner et al. (2011) since the calculation is less complex in comparison to Adler and Alfaro (2007) and Javanmardi et al. (2010). Furthermore, this approach uses a binary definition of quality and is therefore suitable to calculate the *TPR* and *TNR* needed for the evaluation. The approach of Wöhner et al. (2011) distinguishes between persistent and transient contributions. Persistent contributions survive a significant time interval of at least 14 days. These contributions are judged as accepted by the community and therefore

considered as to be of high quality. In contrast, transient contributions are reverted within the time interval of 14 days. These contributions comprise undesirable contributions such as vandalism or spam.

In practical use, an entire edit might consist of a persistent and a transient part. In the following, the amount of the persistent part measured by the number of characters is referred to as *pers*. The metric considers the inserted and the deleted characters as well. In analogy to the persistent part of the edit, the amount of the transient part is denoted as *trans*. Hence, the efficiency of a given edit, which defines the percentage of the persistent part, is calculated as follows:

$$eff = \frac{pers}{pers + trans} \quad (1)$$

3.3 Metrics for MaWiki

To realize the MaWiki concept as described above a wide set of metrics from previous research can be employed (see "Related Work" section). In this section, we introduce a concrete implementation using a metric from Wöhner and Peters (2009) for quality assessment and a metric from Wöhner et al. (2011) to measure the author reputation. For PRM we employ a set of metrics from recent research on vandalism detection (Potthast et al. 2008; West et al. 2010; Adler et al. 2011; Mola-Velasco 2011; West and Lee 2011). The chosen metrics in our study are only examples to demonstrate the potential of MaWikis. For the selection of the metrics we took into account their performance as well as the complexity for the calculation.

To assess the quality of articles we employ the total amount of all previous persistent contributions *Pers* of the given article. Hence, edits will be permitted without any restrictions if *Pers* is lower than a predefined threshold value $Pers^{\tau}$. To calculate an author's reputation we employ the efficiency *Eff* (Wöhner et al. 2011). *Eff* denotes the average efficiency *eff* of all previous edits of the given author. Using *Eff*, RRM permits edits on high-quality articles if the reputation of the author exceeds a predefined efficiency threshold Eff^{τ}. Otherwise, the author is not allowed to edit the article.

PRM identifies undesirable edits on the basis of typical patterns for this kind of edits. This task is similar to the automatic vandalism detection. However, in contrast to vandalism detection, undesirable edits do not only aim at a manipulation of the article, but also cover unintended incorrect information, irrelevant content or edits wars. Regardless of the differences to undesirable edits, metrics from research on vandalism detection are also suitable for PRM. Table 2 provides an overview of the potential metrics for PRM. Besides metrics from the literature, Table 2 considers some new metrics that measure the amount of contributions. We believe that these features might also affect the acceptance of an edit.

The research papers listed in Table 2 discuss some further metrics not mentioned in the table. However, these metrics regard the reputation of an author. Since PRM is only employed for anonymous authors, the reputation is unknown and therefore reputation-based metrics are irrelevant for our study. Moreover, we disregard language-based metrics (e.g. frequency of vulgar and offensive words, frequency of first and second person pronouns) since these features depend on the used language and therefore require domain-specific knowledge.

Which of the metrics listed in Table 2 are relevant depends on the definition of undesirable edits. Hence, in practical use, an efficiency threshold eff^{τ} has to be defined that separates desirable and undesirable edits. Based on this definition, patterns and relevant metrics can be determined using machine learning algorithms.

4 Evaluation

In this section, we first describe our evaluation method and subsequently present and discuss the results.

4.1 Evaluation Method

The MaWiki concept is evaluated by means of a simulation on the basis of Wikipedia data. We chose Wikipedia as example for our analysis since currently known metrics for vandalism detection, quality assessment as well as reputation assessment are also evaluated on Wikipedia data. Due to the lack of studies on other wiki systems it is uncertain how these metrics perform on wikis besides Wikipedia. In the following subsections we first describe our data set and the applied simulation model in detail. Subsequently, we explain the parameterization of the model used for the evaluation.

4.1.1 Data Set

In our simulation we employ the data of the German Wikipedia. The data contains the complete editing history that comprises the source texts of all article versions as well as meta information like the username and the editing time. In the case of an anonymous edit, the IP address is saved instead of the username.

In May 2008 the German Wikipedia modified their editing model by introducing *flagged revisions*. According to this approach, new article versions are only accepted if they are checked by an experienced author (called *Sichter*) to avoid obvious vandalism. However, we are interested in an evaluation of the effectiveness of the MaWiki concept in

Table 2 Metrics for the detection of undesirable edits

Category	Metric	Symbol	Reference
Meta data	Daytime of the edit*	*daytime*	*1, 4, 5*
Text-based metrics	Day-of-week of the edit	*dayofweek*	*1, 4, 5*
	Length of the editing comment	*len_comment*	*2, 3, 4, 5*
	Time interval to the last edit	*time_last_edit*	*1, 4, 5*
	Anonymous author	*anonym*	*1, 2, 3, 5*
	Percentage of upper-case characters	*r_uppercase*	*1, 2, 3*
	Percentage of numerical characters	*r_digits*	*1, 3*
	Length of the longest token	*longest_token*	*1, 2, 3, 5*
	Length of longest consecutive sequence of single character	*same_char*	*1, 2, 3, 5*
Amount of contributions	Length of the deleted text*	*del*	*6*
	Length of the inserted text*	*add*	*6*
	Size difference between previous and current version*	*diff*	*1, 3*
	Ratio of the length of the deleted text (*del*) to the length of the previous article version*	*r_del*	*6*
	Ratio of the length of the inserted text (*add*) to the length of the previous article version	*r_add*	*6*
	Ratio of the size difference (*diff*) to the length of the previous article version*	*r_diff*	*1, 2, 3*

* – Relevant metrics according to Hall and Smith (1999)

(see Parameterization on p. 14)

References:
1 – Adler et al. (2011) 3 – Mola-Velasco (2011) 5 – West and Lee (2011)
2 – Potthast et al. (2008) 4 – West et al. (2010) 6 – New metric

pure wiki systems. The used data set should not be affected by significant modifications of the wiki concept and therefore the data should be captured before May 2008. Hence, we employ the data of the German Wikipedia of 21 January 2008 that is also used by Wöhner and Peters (2009).

Wikipedia is divided into different namespaces. Our study is restricted to the main namespace which includes all encyclopedia articles. We disregard other Wikipedia pages such as discussion pages and user pages because of their special editing process. Overall the data includes 1,023,507 articles and 26,392,081 article versions. 7,602,790 article versions were created by anonymous authors. The remaining article versions were written by 180,488 registered authors.

The available data provides all information needed to calculate our metrics. In order to compute the persistence

of contributions we employed the text comparison algorithm from Hunt and McIlroy (1975) on word level. We developed a set of small java tools to compute the used metrics.

4.1.2 Simulation Model

The simulation model is illustrated in Fig. 2. We implemented the simulation model in an own Java application. In the simulation the edits are processed in chronological order. For each edit the system decides whether the edit is permitted or blocked according to our set of rules. If the edit is permitted, the quality score *Pers* of the given article and the reputation score *Eff* of the given author is adjusted in the simulation. According to this procedure the simulation demonstrates how Wikipedia would have evolved, if the MaWiki concept had been employed. For the interpretation of the simulation results it has to be considered that users might adapt their editing behavior if their editing rights were restricted. Therefore, the simulation evaluates the effectiveness of the MaWiki concept only approximately.

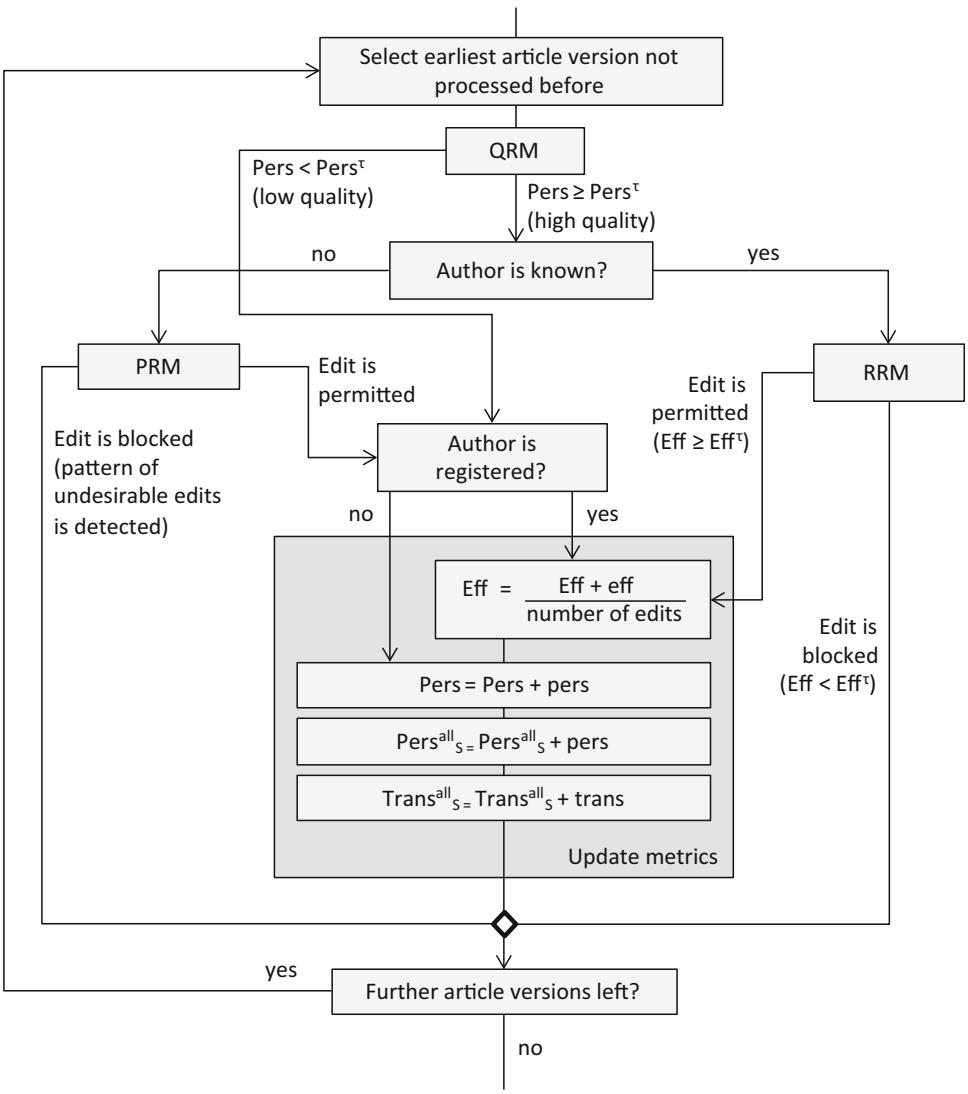
We evaluate the effectiveness of the MaWiki concept on the basis of the total amount of the permitted persistent $Pers_S^{all}$ and transient contributions $Trans_S^{all}$ within the simulation. By means of a comparison to the amount of the original performed persistent $Pers_O^{all}$ and transient contributions $Trans_O^{all}$ the *TNR*

$$TNR = \frac{Trans_S^{all}}{Trans_O^{all}} \tag{2}$$

and the *TPR*

$$TPR = \frac{Pers_S^{all}}{Pers_O^{all}} \tag{3}$$

Fig. 2 Simulation model



are calculated. The accuracy

$$ACC = \frac{Pers_S^{all} + Trans_O^{all} - Trans_S^{all}}{Pers_O^{all} + Trans_O^{all}} \quad (4)$$

refers to the ratio of the amount of correctly classified contributions (permitted persistent and blocked transient contributions) to the amount of all contributions. Therefore, *ACC* quantifies the total effectiveness of the concept.

4.1.3 Parameterization

Since a given edit usually has a persistent and a transient part, a perfect *TPR* and *TNR* are not realizable simultaneously. To manage the trade-off between these two measures, an efficiency threshold eff^{τ} is defined to distinguish between desirable and undesirable edits. Depending on this definition the parameters Eff^{τ} and $Pers^{\tau}$ as well as the patterns for undesirable edits have to be determined.

In our study, we employ a threshold efficiency $eff^{\tau} = 0.5$. Therefore, edits having an efficiency $eff < 0.5$ are defined as undesirable. In the following, we call these edits ineffective edits. On the contrary, edits with an efficiency $eff \geq 0.5$ are assumed as to be desirable; they are referred to as effective edits. In other words, for effective edits the amount of the transient contribution is smaller than the amount of the persistent contribution. Following this definition, the threshold for RRM is defined as $Eff^{\tau} = 0.5$.

For the training of MRM we randomly selected 20,000 ineffective and 20,000 effective edits performed by anonymous or newly registered authors. Using this sample, we first determined which of the metrics listed in Table 2 are relevant. For this task we apply the approach of Hall and Smith (1999), which determines the relevance based on the information value and the intercorrelation between the metrics. The relevant metrics are marked with an asterisk in Table 2. In particular, metrics that regard the amount of contributions are relevant.

Considering the relevant metrics, we employ several popular classifiers to determine relevant patterns. To perform the tests, we make use of the data mining tool Weka.[1] We use two-third of our article sample (13,333 articles) for training and one-third (6667 articles) for testing. The achieved accuracy rates are shown in Table 3. The classifiers are described in detail in Witten et al. (2011) for example.

Table 3 shows the decision tree based and the rule-based classifiers achieve similarly high accuracy rates. In our study, we apply the set of rules calculated by Repeated Incremental Pruning to Produce Error Reduction (RIPPER). In comparison to the other approaches this set of rules is the least complex one. Furthermore, RIPPER was designed to reduce the risk of overfitting. The determined set of rules is described in Table 4. In particular, large deletions are classified as to be ineffective.

Table 3 Comparison of classifiers

Category	Classifier	Accuracy
Bayes	Naive Bayes	52.89%
Functions	SVM	52.83%
Rule-based	RIPPER	63.37%
Decision tree	C4.5	62.92%
	LMT	63.13%
	Random Forest	62.93%
ANN	Multilayer perceptron	52.66%

Table 4 Rules for PRM

	Rule	Decision
1.	$r\_del \leq 0.000047$ & $diff \leq 112$ & $r\_diff \leq 0.057214$ & $diff \geq 35$	Block
2.	$diff \leq -255$	Block
3.	$r\_diff \geq 0.000159$ & $del \leq 18$ & $r\_del \leq 0.018395$	Block
4.	*Else*	Permit

For QRM, $Pers^{\tau}$ does not depend on the efficiency threshold eff^{τ} directly. We determined an appropriate $Pers^{\tau}$ by means of an initial simulation. To reduce the complexity of this initial simulation, we selected all articles of the category E-Business and the corresponding subcategories. The data sample includes 258 articles and 22,040 article versions. The selection of an entire category should ensure that the data sample involves a user group that is as closed as possible. We simulated the evolution of the selected articles by performing several simulation runs. In each run we incremented the parameter $Pers^{\tau}$ by 500, starting from $Pers^{\tau} = 0$ up to $Pers^{\tau} = 100{,}000$. We parameterized RRM and PRM as described above. As a result of this initial study we defined $Pers^{\tau} = 5000$ since we achieved the highest accuracy *ACC* for this value.

5 Results and Discussion

In this section, we first present the results of our evaluation and subsequently discuss the practical implications of our findings.

[1] http://www.cs.waikato.ac.nz/ml/weka/.

Table 5 Evaluation of the MaWiki concept

	Hypothetically perfect classification	Simulation
TPR	*98.1%*	***95.5%***
TNR	*86.4%*	***52.9%***
ACC	*95.1%*	***84.6%***

5.1 Results

Table 5 presents the results of our evaluation. The simulation shows that the suggested MaWiki concept is able to assign editing rights with a high accuracy *ACC* of 84.6 %. The persistent contributions are hardly affected by the editing rights restriction. Thus, 95.5 % (*TPR*) of the performed persistent contributions are permitted according to our set of rules. In contrast, 52.9 % (*TNR*) of the transient contributions are blocked. Even if some ineffective edits are permitted falsely, the editing rights restriction prevents a considerable amount of the transient contributions and therefore leads to a significant improvement. As discussed above, edits in wikis typically include transient as well as persistent contributions. If the system blocked all ineffective edits perfectly, 98.1 % of the persistent contributions would be permitted and 86.4 % of the transient contributions would be blocked. In comparison to this perfect classification, our set of rules is noticeably effective.

Table 6 shows how the contributions are distributed among the three components of the MaWiki concept. Hence, the table indicates the relevance of the three components.

As assumed above, transient contributions are rarely made towards low-quality articles (19.8 % of the total transient contributions). Nevertheless, one could claim that low-quality articles should also be controlled by RRM and PRM. However, this procedure would affect the development of Wikipedia negatively. To investigate this effect in more detail, we carried out a simulation without the QRM component. In that case a significant percentage of 18.6 % of the persistent contributions would be blocked falsely. Some misclassifications are caused by the fact that RRM blocks all edits of low-reputation authors. These authors have no possibility to improve their reputation and therefore are permanently excluded from editing. This analysis proves the necessity of the QRM component that enables authors to rehabilitate their reputation by editing low-quality articles. Hence, with regard to usability and fairness of the system, we suggest making use of QRM in a practical implementation.

RRM deals with the main portion of the persistent contributions (50.2 %), but only 17.6 % of the transient contributions. Since the majority of registered authors have a high reputation ($Eff \geq 0.5$), RRM works less restrictive so that on the one hand about 99.13 % of the persistent contributions are permitted and on the other hand only 12.3 % of the transient contributions are blocked. But even if RRM permits the majority of the edits, MaWikis without

Table 6 Relevance of the components of a MaWiki

		QRM *Low quality article*	RRM	PRM
$Pers^{all}_{o}$	Absolute	2,997,732,622	3,719,759,744	692,741,440
	Overall percentage	*40.5%*	*50.2%*	*9.3%*
$Pers^{all}_{s}$	Absolute	2,997,732,622	3,694,293,363	382,958,760
	Percentage permitted	*100%*	*99.3%*	*55.3%*
$Trans^{all}_{o}$	Absolute	502,887,053	447,281,564	1,587,624,147
	Overall percentage	*19.8%*	*17.6%*	*62.6%*
$Trans^{all}_{s}$	Absolute	502,887,053	392,015,616	299,845,819
	Percentage permitted	*100%*	*87.6%*	*18.9%*

RRM will be ineffective. If all edits of registered authors were permitted, authors intending vandalism could register and perform edits without any restrictions.

The main part of the transient contributions (62.6 %) is covered by PRM since most of the transient contributions are performed by anonymous and newly registered authors. The rules of PRM are very restrictive and therefore about 81.1 % (*TNR*) of the transient contributions are blocked. As a drawback of the restrictive assignment of editing rights only 55.3 % (*TPR*) of the persistent contributions are permitted by PRM. However, since the amount of persistent contributions is small in comparison to the two other components, the restriction hardly affects the overall performance of the MaWiki concept. According to the relatively low *TPR*, it seems that anonymous and newly registered authors are strongly affected by the MaWiki concept. However, QRM allows these authors to edit low-quality articles without any restrictions. Considering the entire Wikipedia, 75.6 % of the contributions of anonymous and newly registered authors are permitted in the simulation. Consequently, an introduction of the MaWiki concept would not lead to a drastic impact on authors with an unknown reputation.

5.2 Practical Implications and Discussion

The evaluation on the basis of Wikipedia shows that the proposed MaWiki concept is able to block undesirable contribution with a high accuracy. Our set of rules can be employed in different use cases.

A first use case is the application of MaWikis in existing open as well as corporate wiki projects to reduce the damage caused by undesirable edits. However, especially in the case of corporate wikis, the motivation of employees to contribute is often a critical success factor (Paroutis and Saleh 2009). Hence, restriction of editing rights might be risky since it may prevent users from participating in the wiki project. Otherwise, uncertainty regarding the information quality (Bhatti et al. 2011) and uncontrolled access (Holtzblatt et al. 2010) are also identified as barriers for the adoption of wikis. In these cases, MaWikis can improve the prospects of success of wiki projects. Hence, it depends on the particular conditions whether MaWikis or traditional wikis are better suited.

In order to use MaWikis in existing wiki projects, effective metrics for quality assessment, reputation assessment and vandalism detection have to be identified first. For the example of Wikipedia, the metrics explained in Sect. 3.3 are suitable. For wikis other than Wikipedia, metrics from Wikipedia research are potentially applicable but their effectiveness has to be evaluated in advance. On the basis of the selected metrics, the set of rules can be parameterized in a final step as described in Sect. 4.1.3.

Furthermore, the MaWiki concept can also be used to open corporate administrated web projects for the web community. Such a strategy can be employed for example in crowd sourcing projects to participate customers in the writing of product reviews or product descriptions. However, in these scenarios only limited information exists about the author's reputation and the patterns of undesirable edits. Therefore, the access should be open at system launch, and the set of rules can be trained and adapted incrementally while using the platform.

The MaWiki concept as suggested in this paper is fully automated. Alternatively, the concept can be adapted and applied in a semi-automated way. In some use cases it is hard to assess the quality of wiki sites automatically. In this scenario a human-driven quality rating may be employed to decide if the editing rights should be controlled by RRM and PRM. Such human-driven ratings additionally allow the separate assessment of particular parts of wiki sites. This procedure provides the benefit that editing rights might be restricted for high-quality content only whereas low-quality parts of the wiki site can be edited without restrictions. This approach is not applicable in a fully automated setting since currently known metrics for quality assessment affect an entire wiki site.

Furthermore, instead of blocking edits automatically, our set of rules can also be applied to flag potentially questionable contributions similar to the approach from Adler et al. (2008). This procedure might help users to decide whether they can trust the given fact or not. Moreover, the marking could motivate users to validate and revise the text where necessary. By using the set of rules for marking, transient contributions may contribute to improve the quality of wiki sites in an indirect way.

The concept of MaWikis comprises the idea of restricting editing rights based on the reputation of authors (RRM) and the occurrence of suspicious editing patterns (PRM). This basic concept is not limited to wikis only. Moreover, undesirable contributions are also a relevant issue in other Web 2.0 applications like blogs, photo, and video sharing platforms as well as social networks. These platforms allow users to post information or comment on published contents, which leads to the arising problem of online harassment and cyberbullying (Yin et al. 2009). The development of automatic reputation systems and the machine-learning based detection of suspicious editing patters will enable the adoption of our MaWiki concept to overcome the mentioned problems.

6 Conclusion

Traditional WCMS can only be accessed by a small number of privileged authors. Because of the restrictive

access, the content is generally less comprehensive and less up to date. In comparison, wikis like Wikipedia employ an open access model where daily web users are allowed to edit the content directly within the browser without any restrictions. This procedure usually leads to a high article quality, but undesirable edits like vandalism or spam cannot be ruled out. Previous research introduced three approaches to overcome the quality problem in wikis. The automatic quality assessment (e.g. Lih 2004; Dondio and Barrett 2007; Blumenstock 2008; Wöhner and Peters 2009) as well as the reputation assessment (Adler and Alfaro 2007; Javanmardi et al. 2010; Wöhner et al. 2011) try to signal the quality of content. Goal of the automatic vandalism detection (Potthast et al. 2008; West et al. 2010; Adler et al. 2011; Mola-Velasco 2011; West and Lee 2011) is the identification and the reverting of low-quality contributions. However, previous approaches do not block undesirable edits in advance. In this paper, we develop a completely new concept called Managed Wikis (MaWikis) that combines the benefits of traditional WCMS and wikis. In comparison to previous approaches, MaWikis block undesirable edits in advance and therefore reduce the costs caused by such contributions. Furthermore, we develop a new evaluation method to judge the effectiveness of MaWikis.

According to the MaWiki concept, editing rights are controlled by a set of rules that involves three components. The quality-based editing rights management (QRM) permits edits on low-quality articles without any restrictions so that the open character typical for wikis is maintained. For high-quality articles, editing rights are controlled by the reputation-based editing rights management (RRM) and the pattern-based editing rights management (PRM). For these articles, low-quality contributions lead to a significant loss of article quality. RRM regulates the editing rights of registered authors and permits edits if the author's reputation exceeds a predefined minimum reputation score. PRM controls the edits of anonymous and newly registered authors and allows edits if no typical patterns for undesirable edits are detected.

To implement the MaWiki concept, a variety of metrics presented in previous research can be used. In this paper, we introduce an implementation that is based on the detection of persistent and transient contributions. Persistent contributions survive at least a significant time interval of 14 days. They are accepted by the community and therefore assumed as to be of high quality. In contrast, transient contributions are discarded by the community quickly and therefore judged to be of low quality. Based on the detection of persistent contributions, QRM assesses the quality of wiki sites by the total amount of persistent contributions to the given site. For RRM, the reputation of authors is measured by the *efficiency*, which denotes the percentage of the persistent contributions of a given author. Patterns of undesirable edits are detected by PRM, in particular by means of the amount of deleted text.

We have evaluated the proposed implementation of our MaWiki concept on the basis of a simulation using Wikipedia data. The evaluation shows that our set of rules assigns editing rights effectively. Transient contributions are blocked significantly (52.9 %), whereas persistent contributions are hardly affected by the editing rights restriction (95.5 % are permitted). Our proposed MaWiki concept is applicable to existing wikis like Wikipedia to reduce damage caused by undesirable edits. Moreover, MaWikis can be employed in new wikis or in web projects that use traditional WCMS. With MaWikis the editing process can be opened for the web community with only a limited risk that the content is tampered with vandalism or spam.

This research is a typical design science work according to Hevner et al. (2004). The MaWiki concept as new artifact is evaluated on the basis of Wikipedia and the evaluation demonstrates the utility of this approach. The final set of rules is the result of a multi-step search process. In each step we tested different variants of the rule tree and judged them by their performance. We interpret the blocking of persistent contributions as type I error. Considering this definition, the goal of the search process was to find a configuration that maximizes the amount of blocked undesirable contributions under the condition of a usually accepted error rate of 5 % at most ($\alpha \leq 0.05$ respectively $TPR \geq 95$ %). This rate ensures that the main part of desirable contributions is permitted.

In detail, we tested the following further variants of the set of rules:

- set of rules without QRM
- non-binary notions of author reputation and article quality
- application of the PRM for registered authors

We could achieve the best performance by using the final rule tree described above. Our search process provides a starting point, and in future work further variants of the set of rules can be evaluated.

As Hevner et al. (2004) state, in a next step the artifact should be employed in practical conditions and be evaluated in more detail by behavioral research work. Such research will validate the effectiveness of the system, the user acceptance, possible changes in the user behavior as well as the computational complexity of MaWikis. Based on such behavioral research, the concept can be brought forward and alternative use cases like the marking of questionable contributions instead of blocking edits can be tested. Finally, a further interesting research goal is the adoption of the idea of automatic editing rights management to other application types in Web 2.0 such as blogs,

social networks and photo and video sharing platforms. This method might be employed to overcome current problems in Web 2.0 like online harassment and cyber-bullying (Yin et al. 2009).

References

Adler BT, Alfaro L (2007) A content-driven reputation system for the Wikipedia. In: Proceedings 16th int conf world wide web, ACM, New York, pp 261–270. doi:10.1145/1242572.1242608

Adler BT, Chatterjee K, Alfaro L, Faella M, Pye I, Raman V (2008) Assigning trust to Wikipedia content. In: Proceedings international symposium wikis (Wikisym 2008), ACM, New York

Adler BT, Alfaro L, Mola-Velasco SM, Rosso P, West AG (2011) Wikipedia vandalism detection: Combining natural language, metadata, and reputation features. In: Proceedings 12th international conference computational linguistics intelligent text processing (CIC Ling'11), Tokyo, pp 277–288

Alexa (ed) (2015) Top sites. The top 500 sites on the web. http://www.alexa.com/topsites. Accessed 2 Nov 2015

Arazy O, Gellatly I, Soobaek J, Patterson R (2009) Wiki deployment in corporate settings. Technol Soc Mag IEEE 28(2):57–64

Bhatti ZA, Baile S, Yasin HM (2011) The success of corporate wiki systems: an end user perspective. In: Proceedings 7th international symposium wikis open collaboration (WikiSym 2011), ACM, New York, pp 134–143

Blumenstock JE (2008) Size matters: word count as a measure of quality on Wikipedia. In: Proceedings 17th international conference world wide web (WWW08), ACM, New York, pp 1095–1096. doi:10.1145/1367497.1367673

Cunningham W, Leuf B (2001) The wiki way quick collaboration on the web. Addison-Wesley, Boston

Denning P, Horning J, Parnas D, Weinstein L (2005) Wikipedia risks. Commun ACM 48(12):152. doi:10.1145/1101779.1101804

Dondio P, Barrett S (2007) Computational trust in web content quality: a comparative evaluation on the Wikipedia project. Informatica 31(2):151–160

Giles J (2005) Internet encyclopedias go head to head. Nature 438(15):900–901. doi:10.1038/438900a

Gregor S, Hevner AR (2013) Positioning and presenting design science research for maximum impact. MIS Q 37(2):337–355

Hall MA, Smith LA (1999) Feature selection for machine learning: Comparing a correlation based filter approach to the wrapper. In: Proceedings Florida artificial intelligence symposium, AAAI, pp 235–239

Hevner A, Chatterjee S (2010) Design research in information systems. Theory and practice. Springer, New York

Hevner AR, March ST, Park J, Ram S (2004) Design science in information systems research. MIS Q 28(1):75–105

Hippner H, Wilde T (2005) Social Software. WIRTSCHAFTSINFORMATIK 47(6): 441–444

Holtzblatt LJ, Damianos LE, Weiss D (2010) Factors impeding wiki use in the enterprise: a case study. In: Proceedings CHI'10, ACM, pp 4661–4676

Hunt J, McIlroy M (1975) An algorithm for differential file comparison. Comput Sci Tech Rep Bell Lab

Javanmardi S, Lopes C, Baldi P (2010) Modeling user reputation in Wikipedia. J Stat Anal Data Min 3(2):126–139. doi:10.1002/sam.10070

Lih A (2004) Wikipedia as participatory journalism: Reliable sources? Metrics for evaluating collaborative media as a news resource. In: Proceedings 5th international symposium online Journal, Austin. doi:10.1.1.117.9104

Mola-Velasco SM (2011) Wikipedia vandalism detection. In: Proceedings 20th international conference companion world wide web (WWW'11), ACM, New York, pp 391–306. doi:10.1145/1963192.1963349

Paroutis S, Saleh AA (2009) Determinants of knowledge sharing using Web 2.0 technologies. J Knowl Manag 13 (4): 52–63

Potthast M (2010) Crowdsourcing a Wikipedia vandalism corpus. In: Proceedings 33rd international ACM SIGIR conference research development information retrieval ACM, New York, pp 789–790

Potthast M, Stein B, Gerling R (2008) Automatic vandalism detection in Wikipedia. In: Proceedings adv information retrieval 30th European conference IR research, Springer, New York pp 663–668. doi:10.1007/978-3-540-78646-7_75

Viégas F, Wattenberg M, Dave K (2004) Studying cooperation and conflict between authors with history flow visualizations. In: Proceedings SIGCHI Conference human factors computational system, Vienna, pp 575–582. doi:10.1145/985692.985765

West AG, Lee I (2011) Multilingual vandalism detection using language-independent and ex post facto evidence. In: PAN-CLEF'11 – Notebook Papers on Uncovering Plagiarism, Authorship, and Social Software Misuse, Amsterdam

West AG, Kannany S, Leez I (2010) Detecting Wikipedia vandalism via spatio-temporal analysis of revision metadata. In: Proceedings 3rd European workshop system security (EUROSEC'10), New York, pp 22–28. doi:10.1145/1752046.1752050

Wikimedia (ed) (2015) List of Wikipedias. http://meta.wikimedia.org/wiki/List_of_Wikipedias. Accessed 2 Nov 2015

Witten IH, Frank E, Hall MA (2011) Data mining: Practical machine learning tools and techniques, 3rd edn. Morgan Kaufmann, Burlington

Wöhner T, Peters R (2009) Assessing the quality of wikipedia articles with lifecycle based metrics. In: Proceedings 5th international symposium wikis and open collaboration (WikiSym 2009), Orlando. doi:10.1145/1641309.1641333

Wöhner T, Köhler S, Peters R (2011) Automatic reputation assessment in Wikipedia. In: Proceedings international conference information system (ICIS 2011), Shanghai

Yin D, Xue Z, Hong L, Davison B, Kontostathis A, Edwards L (2009) Detection of harassment on web 2.0. In: Proceedings content anal web 2.0 (CAW2.0) Workshop at WWW 2009, Madrid