Privacy-Respecting Reputation for Wiki Users

Benjamin Kellermann[1], Stefanie Pötzsch[1], and Sandra Steinbrecher[1,2]

[1] Technische Universität Dresden, Faculty of Computer Science,
D-01062 Dresden, Germany
{Benjamin.Kellermann,Stefanie.Poetzsch}@tu-dresden.de
[2] SAP Research Dresden,
Chemnitzer Strasse 48, D-01187 Dresden, Germany
steinbrecher@acm.org

Abstract. Wikis are popular tools for creation and sharing of content. Integrated reputation systems allow to assess expertise and reliability of authors and thus to support trust in the wiki content. Yet, results from our empirical study indicate that the disclosure of user reputation evokes privacy issues. As a solution for this conflict between the need to evaluate trustworthiness of users and protecting their privacy, we present a privacy-respecting reputation system for wikis that we realized as OpenSource-Extension for the wiki software MediaWiki.

Keywords: Empirical Study, Privacy, Reputation System, Wiki.

1 Introduction

A wiki gives users the possibility to create a common information basis. Technically it is implemented as a collection of HTML pages, usually called *wiki articles*, that are generated, modified, and reviewed by wiki users. A very popular example is the online encyclopaedia Wikipedia. In January 2011, Wikipedia was available in 276 languages or dialects, among which English (more than 3.5 million articles and more than 13 million registered users) was the largest one.[1] This means, Wikipedia covers more articles than the famous and still in print available Encyclopaedia Britannica (about 65 thousand articles).

Wikis enable every user to contribute to the content with only her web browser. This leads to the drawback that it becomes difficult for readers to assess quality of content, and/or expertise of authors respectively. Wikis are aware of this problem and give users the possibility not only to change others' articles but to tag an article, e.g., with a number of stars and/or adding a written comment. Examples for MediaWiki – the software used to implement Wikipedia – are the extensions ReaderFeedback, Rating or Discussion.[2] This helps other readers to decide on how much they trust in the article.

Furthermore, the background of the users who create and also of those who rate an article are important. Here reputation systems common for other Internet

---

[1] http://en.wikipedia.org/wiki/Wikipedia:About
[2] These extensions are available from http://www.mediawiki.org/wiki/Extension.

applications, e.g., electronic marketplaces like eBay[3] can be used to help wiki users to build up a reputation about their expertise, that they can use as authors and raters of articles. Having a user's reputation publicly available helps others to assess the trustworthiness of this user. However, this reputation is an aggregate of the user's personal properties/expertise/knowledge and conveys personal data, which means that privacy concerns may evoke and should be addressed. In this paper - after discussing background and related work on privacy-respecting reputation systems in Section 2 - we present an empirical study on users' privacy awareness with respect to their reputation in a wiki (Section 3). From the results of the study and existing work we deduce requirements focused on user reputation systems for wikis in Section 4. Based on the study and the derived requirements we propose a reputation system for wikis that allows users to find a balance between their need for privacy and other users' need to assess trustworthiness. In Section 5 we outline the design and implementation of our user reputation system for the wiki system MediaWiki and evaluate how it can fulfill the previously explained requirements in Section 6. The main focus of our work in contrast to previous work is giving the user the choice to balance his wish for privacy with the trustworthiness the content he creates has for others.

2   Background and Related Work

In the scenario of a wiki users may take the following roles:

- *readers* read articles others wrote (possibly with evaluating ratings others gave to the article),
- *authors* write articles or contribute to existing ones, and
- *raters* rate others' articles (usually after reading them).

In this scenario we want to integrate a user reputation system. Reputation systems for an interaction system need to provide the following protocols [1]:

1. Centralized communication protocols that allow users to:
   - provide ratings about other users,
   - obtain reputation of other users from the reputation server.
2. A reputation computation algorithm to derive users' reputation based on received ratings, and possibly other information.

Earlier work about reputation systems focuses on the e-commerce scenario that reflects such an interaction system with buying and selling products, e.g., [2,3,4]. In our scenario where the interactions between users are only implicitly given by reading/writing content an interesting question is whether

1. users get ratings explicitly from other users or
2. users get ratings implicitly from ratings given to the content they wrote.

---

[3] http://www.ebay.com

User reputation and trustworthiness of content alternatively can also be created from system-based observations about the users' behavior. This approach was the one followed for calculating the reputation of content and users in wikis recently: Dondio et al. suggested an algorithm to automatically calculate the overall trustworthiness of a wiki article based on statistical data [5]. Adler and Alfaro [6] present a system for Wikipedia that computes an editor's reputation by observing whether subsequent users preserve the changes that were made. Recommer Systems are slightly different from Reputation Systems by especially recommending content to users. A popular technique used to do so is collaborative filtering, this means using the input from other users in determining what is useful information for a user. The WRS (Wikipedia Recommender System) [7] is a rating-based personalized recommender system that uses trust metrics to give not only general information but personalised recommendations by comparing a user with other users and then calculating a trust profile of their expertise.

It has been studied from the legal perspective that user reputation can be seen as personal data [8] and should be protected by technical means [9]. This is contradicting to personalising reputation/recommendation data for users by a system. In [10] a small survey among 13 persons indicated that these participants do not feel comfortable to publish their real names within online communities and that half of them regard their reputation as privacy-sensitive data. Besides, there exists no empirical study on how privacy aware users are about their reputation in the context of a wiki. The other direction on how much content- and author-related information actually influences users' perceived trustworthiness of a Wikipedia article is already studied by [11]. Privacy usually covers three aspects: *anonymity of providing ratings*, *anonymity of obtaining other users' reputation*, and *anonymity of users who are rated and whose reputation is obtained*.

Anonymity services - as suggested in [12] - can only protect the users obtaining reputation. In order to obtain anonymity of users rating and being rated, it needs to be ensured that many users are indistinguishable by an attacker, so that they are in large anonymity sets. For anonymity of users who are rated and whose reputation is obtained, others should not be able to link interactions with the same user. The possibility of recognizing users by reputation is limited if the set of possible reputations is limited [13] or if the reputation is only published as an estimated reputation [14]. Transaction pseudonyms can be used to avoid linkability between transactions [15, 13]. In order to obtain anonymity of raters, interactions and ratings related to these interactions need to be unlinkable. This can be reached by a reputation provider who only calculates a new user reputation after it collected not only one but several ratings [16], or who only publishes an estimation of the actual reputation [14]. Further, a rater can be anonymous against the reputation provider by using convertible credentials [10] or electronic cash [15,17] or involve Trusted Third Parties similarly for separating interactions and ratings while preventing attacks [18, 19].

But to the best of our knowledge there exists no design and implementation of a user reputation system that shows the trade-off between trust and privacy requirements in wikis and allows users to make a decision on which to favor.

3   User Reputation and the Need for Privacy

In the previous section we pointed out that a user's reputation is an aggregate of her personal expertise, skills and reliability based on previous behavior and thus may be classified as personal data. According to [20], it is important what is deemed sensitive or personal data in the perception of the individual rather than if it can be evaluated by third parties (e.g., lawyers, computer specialists). Considering that individuals often claim to have a huge need for privacy but behave differently (see also the *privacy paradox* [21]), we decided to conduct a larger study with experimental part to learn how users actually treat their own reputation value compared to other personal data items. In the following we briefly sketch the design of the study and report key results.

3.1   Study Design

The web-based study consisted of an experiment and a standardized question-naire. Invitations to participate in the study were posted in forums and blogs on the Internet and we also distributed flyers in a university library. All partic-ipants who completed the study were offered the chance to win vouchers for an online shop. For the experiment, all participants were asked to rate the same ar-ticles from a wiki about books and literature according to three given categories. Before participants actually accessed the wiki articles, they did a short litera-ture quiz. By answering four multiple choice questions about famous writers and books, they received between zero and four points. These points are considered as a participant's *reputation*. Participants further were asked to indicate name, age and place of residence. When rating the wiki articles subsequently, each par-ticipant decides whether her name, age, place of residence, and/or reputation should be published together with her rating of the wiki article.

Half of the participants were randomly assigned to the experimental group. Together with each wiki article, privacy-awareness information were displayed to the experimental group (i.e., information about who can see which data). The other half of the participants belonged to the control group and did not receive privacy-awareness information. After finishing this first part of the study, all participants filled in the questionnaire. In this questionnaire we asked about perceived privacy in the wiki, experience with wikis, ratings systems and the Internet in general. We used questions from the applied privacy concerns and protection scale [22] to investigate general caution, privacy concerns and applied protection measures. Finally, we asked about demographic data and whether subjects had given their real name, place of residence and age at the beginning.

We calculated the *Perceived Privacy Index* (PPX)[4] from participants' answers to the questions how public, how private, how anonymous and how identifiable

---

[4] The questionnaire contained the questions: "Please indicate to which extent the following adjectives describe your feelings while using the wiki: 0 % (not at all) — 100 % (very much)?" (originally asked in German). The PPX is composed of the adjectives "public" (scale inverted), "private", "anonymous", "identifiable" (scale inverted).

they have felt in the wiki. Each item was measured on a 0 to 100 % slider scale. The higher the PPX value, the more private a participant has felt.

3.2 Results

After excluding data sets from few participants who admitted not to have seriously participated in the study, 186 valid responses were further analyzed.

30 % of the participants agreed to publish their real name together with their rating of a wiki article. The disclosure of their real age was okay for 57 %, real place of residence for 55 % and 63 % agreed to have their reputation value published. This means, for each data item there was a considerable share of participants who wished to keep this information private. If participants indicated later in the standardized questionnaire that they did not provide true information in one of the first three categories, we treated this data item as not disclosed. Since the reputation value was calculated from answers in the literature quiz, lying was impossible here.

Further, we used a linear regression model [23] to calculate how the disclosure of these four data items and a few other factors influence user's perceived privacy in the wiki. Results are listed in Table 1 and reveal that availability of privacy-awareness information and application-independent measures, like privacy concerns, general caution and protection measures, did not play a significant role for perceived privacy in the wiki. Yet, there are two factors that significantly decrease perceived privacy: the fact that a user has published her name and the fact that a user has published her reputation value. While it is not surprising that a user feels less private after disclosing her real name, we found that also disclosing the own reputation value has a similar effect on perceived privacy. According to the results it is to say that the reputation value is deemed an even more sensitive data than age or place of residence. Further unreported models show that also including combinations of data items, e.g., age published + reputation value published, have no additional effect.

In line with findings from the small survey in [10], the results of our empirical study underline that a user's reputation (even if only a numerical value) needs to be treated as personal data item, i.e., the user should have control over the disclosure of this information. Therefore we will present a privacy-respecting reputation system based on the ideas proposed in [10] but addressing and showing specifically the trade-off between privacy and trust requirements.

4 Requirements

4.1 Functional Requirements

The goal of a wiki is collecting valid and trustworthy information. The first means to reach this is editing/changing articles by other (possibly more knowledgeable) authors. Therefore many articles in Wikipedia have a long changing history.

---

[5] Intercept indicates the basic value of PPX when all predictors are equal to zero.

Table 1. Regression models for perceived privacy index (n=186)

PPX (dependent var.)	estimate	std. error	p-value	
Application-specific predictors				
Privacy-awareness inform. available	4.57	12.01	0.704	
Name published	−46.66	14.49	0.002	**
Age published	−13.54	16.77	0.420	
Place of residence published	−21.65	16.06	0.179	
Reputation value published	−39.99	14.04	0.005	**
Application-independent predictors				
General caution	0.22	1.79	0.902	
Technical protection	−0.47	1.47	0.750	
Privacy concerns	−1.35	1.10	0.223	
Intercept[5]	288.93	33.47		

sign. levels: $*** p < 0,001$; $** p < 0,01$; $* p < 0,05$

Even information valid at some point in time might change and then the wiki article needs to be changed. But usually many articles will converge to a version containing mostly valid information. Using a content rating system with positive ratings, such acrticles can be marked as such. If the article contains unvalid information the person detecting this can improve or even remove content.

Incentives to improve content can be given by an additional user reputation system. The reputation computation algorithm from the user reputation system uses the content ratings from the content rating system as input to calculate the reputation of users who wrote this article. Users might even receive a monetary compensation for a good reputation. As for many other examples in our society where choices are made certificate-based we assume the following requirement to hold:

Positive integers as ratings and reputation: The rating protocol of the content rating system gets only positive integers as inputs and transfers it to positive reputation of content. The reputation computation algorithm of the user reputation system uses the ratings from the articles a user wrote to calculate her reputation.

Advantages and disadvantages of positive reputation systems depending on the user population are discussed in [24]. The main impact is that it encourages a user population of the wiki that creates valid content of high quality. In the long run all active wiki users would have a maximum reputation value. Low reputation corresponds to not contributing (in a positive way) to the quality of the wiki. To help newcomers [25] to start they can bring in certificates from outside the wiki that are transferred to reputation, e.g., a master in medicine might result in some initial reputation for a medical wiki.

Possibility of non-zero initial reputation: The reputation computation algorithm should be able to transfer a certificate a user got outside the wiki to an initial reputation inside the reputation system

As users have different areas of expertise we assume that wiki, user reputation system and content rating system agree on areas of expertise all articles, ratings, and reputations are linked with:

Area-specific reputation: The reputation system allows to set up different areas of expertise and users can have different reputations in these areas.

We abstract from a concrete classification of areas but refer to [26] for a comparison of several approaches. We only assume an agreement on metadata (tags) for areas of expertise defined for all system components. These areas of expertise should have a tree structure with 'general' as root and certain specific expertise as children. E.g. 'medicine' inherits from 'general' and might have 'gynecology' and 'surgery' as children, which again may have specialized children. All articles in the wiki are tagged with these areas (resp. paths in the tree of areas) as well as user reputation always will be linked to such an area. E.g. an article about breast cancer will be tagged with the expertise path 'general → medicine → gynecology'. Wikipedia already realises this with its Categorys[6], but they are organised as general graph. For reasons of simplicity we assume a tree, but it can easily be extended to a general graph.

The reputation computation algorithm of a user reputation system needs to be accurate and self-correcting. This means that it should consider all ratings given in an appropriate way by considering at the same time various attacks [27]. There exist numerous proposals for such algorithms [1] with respective advantages and disadvantages. For our scenario we assume a rating algorithm of the content rating system and a reputation computation algorithm of the user reputation system with the following features:

Influence of rater on user reputation: The higher the reputation of a rater is the more influence it will usually have on the perceived trustworthiness of the rating for others. For this reason the respecting rater's reputation should be used when calculating a user's reputation.
Influence of time on user reputation: To encourage users to contribute to the wiki continuously their reputation should be aged.
Availability of user reputation with content: As users want to estimate directly the quality of content the reputation of users who edited or rated this article should be stored with the content.

Beneath the functional requirements the scenario has privacy and security requirements that are partly contradicting and need to be balanced according to the principles of multilateral security [28]. We list the most important ones in the following. But due to the contradiction our system can only fulfill them partly.

4.2   Security Requirements

The following security requirements most important for our scenario of a user reputation system in a wiki are derived from [27, 17]. They need to be met by

---

[6] http://en.wikipedia.org/wiki/Special:Categories

the protocols for providing ratings to the content rating system and obtaining user reputation from the user reputation system:

Integrity of content and ratings: Users want content and ratings given to it to be preserved from manipulations.
Accountability of raters: Raters should be accountable for their ratings.
No self-rating: It should be impossible that users rate themselves [25].

4.3   Privacy Requirements

As we could see from the study in Section 3 both users' real names and reputation should be protected by technical means. For the real name the corresponding requirements are similar to [27,17]:

Anonymity/pseudonymity of raters and authors: Users want to rate and provide content anonymously or under a pseudonym to not necessarily allow others to link this rating to their real name.
Anonymity of readers: Users want to read content and evaluate corresponding user reputation anonymously.

Privacy of reputation can mean keeping the reputation confidential what would contradict the availability of reputation and therefore is no option. The better alternative is to relax the sensitivity of reputation as identifying or linkable attribute of users. This can be described by the following requirements:

Anonymity/pseudonymity of users for showing reputation: Users want to show reputation only pseudonymously or even anonymously.
Unlinkability of users for rating and writing content: Users want to be unlinkable when rating and writing content, also for showing reputation.

It needs to be communicated to users whether a user reputation system has these privacy options to relax their privacy awareness about reputation (cp. Section 3).

5   Infrastructure and Design

Our scenario needs three major system components to be in place: The first component is the wiki itself where we decided to use MediaWiki. Second, for the content rating system, there are already some extensions to MediaWiki available where we chose the ReaderFeedback[7] extension.

As third component for the privacy and security requirements we use the privacy-enhancing identity management system PRIME [29] as it was used in [10] to assist users in controlling their personal data. This data can be certified by third parties in the form of credentials. Credentials can be convertible [30] and might be used with various unlinkable pseudonyms. The implementation used by

---

[7] `http://www.mediawiki.org/wiki/Extension:ReaderFeedback`
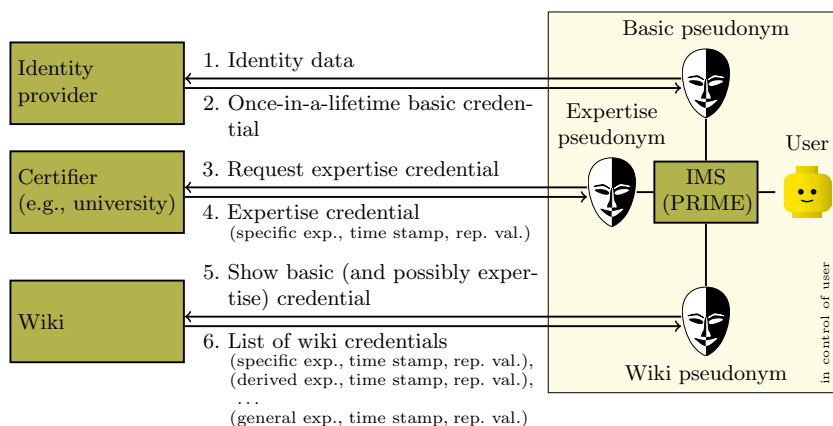
Fig. 1. Registration process enabling unlinkability of a user and her pseudonyms

PRIME for the credential realization is IdentityMixer[8]. This allows for balancing privacy requirements with accountability of users as discussed in [10].

To reach (pseudonymous) accountability of all data and credentials a public-key-infrastructure is needed. For our scenario we assume all actions and ratings to be secured by (pseudonymous) digital signatures. Additionally we assume all communication to be secured by encryption and anonymised by using an anonymity service to reach confidentiality and anonymity of all ratings and actions performed against outsiders.

The user reputation system needs to be integrated in all three system components as we will describe in the following for all actions users can perform. An illustration, how the communication flow of the components is done is shown for the example of editing a wiki article in Fig. 3.

5.1   Registration

A user registers a basic pseudonym with an identity provider by declaration of some identity data (step 1 in Fig. 1). After verifying the data the identity provider issues a basic credential to her (step 2 in Fig. 1). This credential has the meaning, that the identity provider checked the identity data, is willing to disclose it in the case of the user's misbehaviour, and that the credential is issued per user only once in her life.

When the user gains some expertise in an area that can be certified by an independent certifier (e.g., a university) she may ask the certifier (step 3 in Fig. 1) under her expertise pseudonym for a respective expertise credential (step 4).

When the user wants to register in a wiki under a wiki pseudonym she sends the wiki her basic credential (and possibly her expertise credential(s) (step 5 in Fig. 1). The wiki provider creates a list of wiki credentials and sends it back to

---

[8] http://prime.inf.tu-dresden.de/idemix/

the user (step 6). A wiki credential contains an attribute-triple, where the first element is a string for the area of expertise, the second is a time stamp, and the third is a positive integer for the reputation value derived from the expertise, certified in the expertise credential. There might be a list of these triples even if only showing one expertise credential as a user inherits a reputation for all areas lying in the path from 'general' to her specific area of expertise in the tree structure of metadata for the wiki. How the derivation works can be setup up by the wiki administrator depending on her wishes.

The user can show (parts of) a wiki credential with different pseudonyms within the wiki whenever she wants to reach unlinkability of rating or writing content. For considering the type of a rating or reputation, a page must indicate the areas of expertise relevant for it.

5.2   Editing a Wiki-Article

After editing a page (step 1 in Fig. 3), a user is asked which reputation will be shown to others reading the content (step 3). Here it comes clear why we only allow positive ratings because the user would suppress negative one here. There are actually two choices she can make:

First she decides which areas of expertise to choose the reputation from. As the area of the content is fixed, a user may choose the credential, which is specific to the content or an inherited one. A more specific credential will have more impact on the content, a more general one has the benefit of higher anonymity. E.g., if the content is about gynecology, a surgeon might use her credential about medicine or a general knowledge credential.

Second as the reputation is a positive value she can decide on which value smaller or equal to her reputation value in the area she wants to show. For this reason we make use of credentials that allow greater-than-proofs [31]. E.g., when having a reputation value of 63, an author may prove that she has a value greater than 20 or greater than 50 or that she actually has 63. The higher a reputation value is, the more impact it will have on the reputation value of the content but as the number of authors having at least this reputation decreases with increasing reputation value, the anonymity of the author decreases as well.

As every user has to decide on this trade-off on her own, a so called "Send Personal Data Dialog" asks the user for her reputation value and tries to display the trade-off in a graphical way. This dialog is shown in Fig. 2(a). The calculation logics used to show the degree of anonymity is based on the number of credentials with a similar reputation value the wiki server saw recently to calculate the user's anonymity. This is the user's current anonymity set for showing this reputation value. Note that the server does not get the actual reputation value from the user in this dialogue, but only provides the user with the data she got from other users when showing reputation to calculate the user's anonymity. Based on this data the user can calculate her possible anonymity as an estimation and displays it to the user. This problem of only estimating anonymity is inherent existent in privacy-enhanced identity management [32] as there exists no global authority that has all user data and can calculate a user's anonymity exactly.
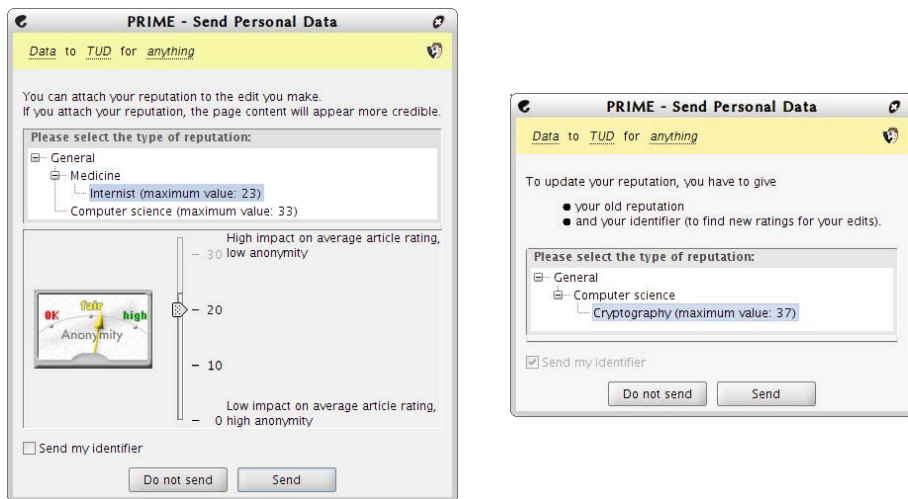
(a) Attaching reputation to the content.   (b) Collecting new ratings

Fig. 2. Two dialogs which show, how personal data is sent

After the user made her choice which reputation to show the wiki server receives the reputation value and can offer it to other users for calculation of anonymity.

In addition to the choices which areas of expertise and which ratings to show, the user can decide on whether she wants to indicate the wiki pseudonym she registered with the wiki to collect reputation. This gives her the possibility to benefit w.r.t. the increase of her reputation value, whenever other raters give a high rating to the page. However, giving the wiki pseudonym makes her linkable. The decision about sending the wiki pseudonym is done with a checkbox shown in Fig. 2(a) ('Send my identifier') on the bottom. This process is driven by the user's own decision whether she wants to collect these ratings or not, both at the time of editing an article by giving a linkable pseudonym and after the article is rated if she includes this rating in her wiki credential.

5.3   Collecting Reputation

If an author provided her identifier, she may collect reputation, given to articles she wrote. She therefore has to show the wiki pseudonym again. The dialog asking for this pseudonym is shown in Fig. 2(b). After showing the old wiki credential, the server issues a new reputation triple. There are a number of possibilities for calculation of reputation [1]. Our reputation system is able to handle ones based on positive ratings with the possibility to weight ratings depending on the rater('s) reputation and to decrease reputation depending on its age indicated in the time stamp of a wiki credential. Both aspects can be configured within our implementation by the wiki administrator.
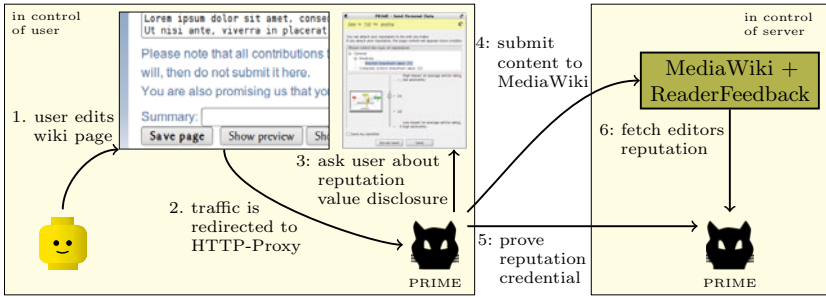
Fig. 3. Communication between the components when editing a wiki article

5.4   Rating Articles in Wiki with Content Rating System

MediaWiki's ReaderFeedback extension allows readers to rate a page in four categories (reliability, completeness, neutrality, presentation) with 1–5 stars. A rating form is shown to the user on the bottom of each page. The aggregated reputation displayed is a daily and selected interval average rating. We modified this extension to collect the user's reputation together with the ratings of a page.

A dialog similar to the one shown in Fig. 2(a) asking for the reputation to show is displayed when a user wants to submit her rating. As for writing content the user has to make a trade-off between showing expertise or remaining anonymous.

5.5   Reading Content in Wiki with User and Content Rating System

Readers have the possibility to view the average rating in the four categories the articles was rated on as in the ReaderFeedback extension. Additionally the detailed list of authors the content had (with their revealed user reputation at this point) and the ratings (with their raters and reputation) can be shown (Fig. 4). The raters' reputation is shown on top of the table below the name or pseudonym of the rater. The different icons represent the reputation type shown (e.g., the syringes represent a certain reputation in medical area). The stars below the raters are the ratings, which were given to a single revision of the page. If an author indicated some reputation together with a page edit, this reputation is shown beneath the authors name or pseudonym.

6   Evaluation

In the following we analyse to which extent the requirements from Section 4 could be met with the prototype.

6.1   Functional Requirements

Positive integers as ratings and reputation: This is given by design of the ReaderFeedback-Extension and our user reputation system based on this.

Possibility of non-zero initial reputation: The expertise credentials allow for indicating prior certified knowledge.

Fig. 4. Example of ratings given to a specific article

Area-specific reputation: The construction of the reputation credential in the form of (area of expertise, time stamp, rating)-triples allows for separating different areas of expertise.

Influence of rater on user reputation: This can be considered by the reputation computation algorithm if the rater agrees to reveal his reputation.

Influence of time on user reputation: This can be considered by the reputation computation algorithm as the wiki credentials contain a timestamp.

Availability of user reputation with content: The users' reputation at the time of editing or rating a wiki article is shown with it. Further updates are not considered.

6.2 Security Requirements

Integrity of content and ratings: As all ratings and content are signed integrity is given.

Accountability of raters: Accountability with the help of the identity provider is given as well as she can reveal the user's identity in the case of misuse. Within the registration process, the user shows her basic credential. Based on this credential, she gets only one wiki credential, which can be converted once by the user to avoid linkability to this registration. This has the advantage that the wiki server can verify if a revision is already rated by an author and update old ratings if necessary. Also self-ratings are prevented and users can collect their ratings to be included in their reputation.

No self-rating: Self-rating is contradicting to the privacy requirements. It therefore is prevented only partially. If users cannot be linked between their areas of expertise, a user can rate her specialized expertise credential with a more general one. However, as the granularity of the ontology has not necessarily to be too fine-grained, a user may not get many ratings (e.g., if the tree from general to specialized is surgery – medicine – general, a user may rate the surgery credential twice) As every user gets only one non-convertible credential from the identity provider, acting with two different wiki credentials which rate one user is prevented.

6.3   Privacy Requirements

Anonymity/pseudonymity of raters and authors: Pseudonymity of users
   is initially given by the infrastructure combined with the identity manage-
   ment.

Anonymity of readers: Anonymity of readers is also assured if reading the
   wiki is possible without registration.

Anonymity/Pseudonymity of users for showing reputation: Users show
   reputation under a pseudonym if they write under a pseudonym. And they
   can determine how much of their reputation to show to remain anonymous.

Unlinkability of users for rating and writing content. For rating and
writing content the users resp. the wiki provider can determine the unlinkability
of users depending on the credentials used and on the reputation shown as we
will outline in the following.

-   **Unlinkability for different articles:** The wiki provider issues one creden-
    tial per user and article. The user can convert the credentials once to assure
    unlinkability. The wiki provider learns from issuing the credentials the pages
    a user is interested in (except a user fetches credentials for all pages by de-
    fault or some dummy-credentials in addition to the real ones). Still this has
    the advantage that the wiki server can verify, if a revision is already rated by
    an author and update old ratings if necessary. Also self-ratings are prevented
    and users can collect their ratings to be included in their reputation. Ratings
    received under different pseudonyms can be aggregated.
-   **Unlinkability for versions of an article:** As wiki articles change fre-
    quently users might also want to be unlinkable for different versions. To
    achieve this the wiki provider issues additionally a number of one-show cre-
    dentials (containing the user's reputation) in a specific time period. With
    these credentials the user may rate a number revisions of an article in the
    specific period, proving her reputation anonymously. This makes the user
    unlinkable but she is not able to collect her ratings afterwards and whatever
    she does under these one-time pseudonyms does not influence her reputation.
-   **Unlinkability with respect to reputation:** Even if users are unlinkable
    for different (versions of) articles or ratings with respect to showing creden-
    tials they still might be linkable by their reputation if always showing the
    same one or one following deterministic rules. According to [13] the set of
    pseudonyms $\mathcal{P}$ at time t in a system is split into several subsets regarding
    the relation "has the same reputation". The set $\mathcal{P}_{t,r} = \{p_{t,i}|rep(t, p_{t,i}) = r.\}$
    with the pseudonyms having the reputation $r \in R$ is the anonymity set a user
    has for re-using a pseudonym if we assume an attacker who knows the repu-
    tation of all users at a time (like the wiki provider does). This anonymity set
    can be increased in our system if a user shows a smaller than her actual rep-
    utation value. Then her anonymity set will be $\mathcal{P}_{t,\geq r} = \{p_{t,i}|rep(t, p_{t,i}) \geq r.\}$
    with $|\mathcal{P}_{t,\geq r}| \geq |\mathcal{P}_{t,r}|$ as all reputations only increase.

7   Conclusion and Future Work

In this paper we have investigated the privacy needs of users with respect to their reputation in a wiki in comparison to other personal data. As results of our empirical study shew that users' reputation value should be treated as personal data item we built a framework that allows users to balance their need for privacy and others' wish for trustworthiness of content by the usage of metadata to indicate users' areas of expertise. Still the design of the framework uses the assumption the platform MediaWiki makes for the reputation computation of content. Future studies need to show which set of possible reputation values users like most and how areas of expertise can be related.

References

1. Jøsang, A., Ismail, R., Boyd, C.: A survey of trust and reputation systems for online service provision. Decision Support Systems 43(2), 618–644 (2007)
2. Schafer, J.B., Konstan, J.A., Riedl, J.: E-commerce recommendation applications. Data Min. Knowl. Discov. 5(1-2), 115–153 (2001)
3. Resnick, P., Zeckhauser, R.: Trust among strangers in internet transactions: Empirical analysis of ebay's reputation system. The economics of the internet and e-commerce. Advances in Applied Microeconomics 11, 127–157 (2002)
4. Livingston, J.A.: How valuable is a good reputation? A sample selection model of internet auctions. The Review of Economics and Statistics 87(3), 453–465 (2005)
5. Dondio, P., Barrett, S., Weber, S., Seigneur, J.-M.: Extracting trust from domain analysis: A case study on the wikipedia project. In: Yang, L.T., Jin, H., Ma, J., Ungerer, T. (eds.) ATC 2006. LNCS, vol. 4158, pp. 362–373. Springer, Heidelberg (2006)
6. Adler, B.T., de Alfaro, L.: A content-driven reputation system for the wikipedia. In: WWW 2007: Proceedings of the 16th international conference on World Wide Web, pp. 261–270. ACM, New York (2007)
7. Lefévre, T., Jensen, C.D., Korsgaard, T.R.: Wrs: The wikipedia recommender system. In: Ferrari, E., Li, N., Bertino, E., Karabulut, Y. (eds.) IFIPTM 2009. IFIP Advances in Information and Communication Technology, vol. 300, pp. 298–301. Springer, Heidelberg (2009)
8. Bygrave, L.: Data Protection Law, Approaching Its Rationale, Logic and Limits. Kluwer Law International, The Hague (2002)
9. Mahler, T., Olsen, T.: Reputation systems and data protection law. In: eAdoption and the Knowledge Economy: Issues, Applications, Case Studies, pp. 180–187. IOS Press, Amsterdam (2004)
10. Pingel, F., Steinbrecher, S.: Multilateral secure cross-community reputation systems. In: Furnell, S.M., Katsikas, S.K., Lioy, A. (eds.) TrustBus 2008. LNCS, vol. 5185, pp. 69–78. Springer, Heidelberg (2008)
11. Kittur, A., Suh, B., Chi, E.H.: Can you ever trust a wiki?: impacting perceived trustworthiness in wikipedia. In: CSCW, pp. 477–480. ACM, New York (2008)

12. Pavlov, E., Rosenschein, J.S., Topol, Z.: Supporting privacy in decentralized additive reputation systems. In: The Second International Conference on Trust Management, Oxford, United Kingdom, pp. 108–119 (March 2004)
13. Steinbrecher, S.: Design options for privacy-respecting reputation systems within centralised internet communities. In: 21st IFIP International Information Security Conference. IFIP, vol. 201, pp. 123–134. Springer, Heidelberg (2006)
14. Dellarocas, C.: Immunizing online reputation reporting systems against unfair ratings and discriminatory behavior. In: Proceedings of the 2nd ACM conference on Electronic commerce, pp. 150–157. ACM Press, New York (2000)
15. Androulaki, E., Choi, S.G., Bellovin, S.M., Malkin, T.: Reputation systems for anonymous networks. In: Borisov, N., Goldberg, I. (eds.) PETS 2008. LNCS, vol. 5134, pp. 202–218. Springer, Heidelberg (2008)
16. Dellarocas, C.N., Dini, F., Spagnolo, G.: Designing reputation (feedback) mechanisms. In: Dimitri, N., Piga, G., Spagnolo, G. (eds.) Handbook of Procurement, Cambridge University Press, Cambridge (2006)
17. Schiffner, S., Clauß, S., Steinbrecher, S.: Privacy and liveliness for reputation systems. In: Martinelli, F., Preneel, B. (eds.) EuroPKI 2009. LNCS, vol. 6391, pp. 209–224. Springer, Heidelberg (2010)
18. Kerschbaum, F.: A verifiable, centralized, coercion-free reputation system. In: WPES, pp. 61–70. ACM, New York (2009)
19. Schiffner, S., Clauß, S., Steinbrecher, S.: Privacy, liveliness and fairness for reputation. In: Černá, I., Gyimóthy, T., Hromkovič, J., Jefferey, K., Královič, R., Vukolić, M., Wolf, S. (eds.) SOFSEM 2011. LNCS, vol. 6543, pp. 506–519. Springer, Heidelberg (2011)
20. Adams, A.: The implications of users' privacy perception on communication and information privacy policies. In: Proceedings of Telecommunications Policy Research Conference, Washington, DC (1999)
21. Pötzsch, S.: Privacy Awareness: A Means to Solve the Privacy Paradox? In: Matyáš, V., Fischer-Hübner, S., Cvrček, D., Švenda, P. (eds.) IFIP WG 9.2, 9.6/11.6, 11.7/FIDIS. IFIP Advances in Information and Communication Technology, vol. 298, pp. 226–236. Springer, Heidelberg (2009)
22. Taddicken, M.: Measuring Online Privacy Concern and Protection in the (Social) Web: Development of the APCP and APCP-18 Scale. In: 60th Annual ICA Conference (International Communication Association), Singapur (June 2010)
23. Wirtz, M., Nachtigall, C.: Deskriptive Statistik. Statistische Methoden für Psychologen, vol. 1. Juventa-Verlag (2008)
24. Whitmeyer, J.M.: Effects of positive reputation systems. Social Science Research 29(2), 188–207 (2000)
25. Friedman, E., Resnick, P.: The social cost of cheap pseudonyms. Journal of Economics and Management Strategy 10, 173–199 (1999)
26. Jensen, C.D., Lefévre, T.: Classifying areas of expertise in the wikipedia recommender system. Online Paper at IFIPTM (2010), http://www.ifip-tm2010.org/lib/exe/fetch.php?media=shortpaper08.pdf
27. ENISA: Position Paper. Reputation-based Systems: a security analysis (2007), http://www.enisa.europa.eu/act/it/oar/reputation-systems/reputation-based-systems-a-security-analysis/at_download/fullReport
28. Rannenberg, K., Pfitzmann, A., Müller, G.: IT security and multilateral security. In: Müller, G., Rannenberg, K. (eds.) Multilateral Security in Communications (Technology, Infrastructure, Economy), vol. 3, pp. 21–29. Addison-Wesley, München (1999)

29. Casassa-Mont, M., Crosta, S., Kriegelstein, T., Sommer, D.: Architecture v2. Deliverable D14.2.c, PRIME (March 2007),
https://www.prime-project.eu/prime_products/reports/arch/
pub_del_D14.2.c_ec_WP14.2_v1_Final.pdf
30. Chaum, D.: Showing credentials without identification. In: Pichler, F. (ed.) EUROCRYPT 1985. LNCS, vol. 219, pp. 241–244. Springer, Heidelberg (1986)
31. Camenisch, J., Van Herreweghen, E.: Design and implementation of the idemix anonymous credential system. In: CCS, pp. 21–30. ACM, New York (2002)
32. Clauß, S.: A framework for quantification of linkability within a privacy-enhancing identity management system. In: Müller, G. (ed.) ETRICS 2006. LNCS, vol. 3995, pp. 191–205. Springer, Heidelberg (2006)