

User's Understanding of Reputation Issues in a Community Based Mobile App

Orlando P. Afonso^(✉), Luciana C. de C. Salgado, and José Viterbo

Department of Computer Science, Fluminense Federal University (UFF), Niterói, Brazil
{oafonso, luciana, viterbo}@ic.uff.br

Abstract. With the emergence of the Web 2.0, (digital) services are covering and moving into the digital and mobile world, producing lots of information in real-time, such as traffic conditions, points of interests and so on. With the big amount of collected data from community-based applications, it becomes necessary to know whether such content is trustworthy. When using applications such as Waze, we are supposed to trust in the information provided by unknown users, which act as digital content producers. However, it needs to be transparently clear where this information comes from and how trustable are its providers/endorsers. This paper presents the results of a two-step study to investigate how users recognize (or not) the signs and the reputation model of digital content producers in Waze app. We analysed and found out how the reputation is communicated to the users and the potential impacts on human computer interaction.

Keywords: Reputation · Community based app · Trust · Semiotic engineering · Communicability

1 Introduction

Community based applications, in general, aim at bringing users interested in work together toward a common goal, such as, finding a best route in a city or the best toilet in a town. With these interactive systems, users may collaborate and interact with instant messaging, profiles, forums and other social networking features. Furthermore, each participant may include, edit, exchange, share and evaluate interactive systems' content that may influence community members in decision-making process. Each participant may act as a consumer and/or a producer of digital content.

This work focus on investigating user's understanding of reputation issues in a Community based Mobile App and the potential implications to Human-computer Interaction (HCI). Reputation can be defined as what is said or believed about a person or thing as said by Josang et al. [6]. Josang and co-authors [6] consider reputation as the collective measure of trustworthiness or reliability based on the referrals or ratings from members in a community too. On the Web, the concepts of trust and reputation are applicable in virtual interaction environments through the Reputation systems [6, 8, 10]. These systems collect, distribute and aggregate information based on the behavior of the participants through their interactions. Thus, they help users to decide in whom to trust, to motivate good behavior among them, and to control the participation of those who are considered

dishonest. The concepts of Reputation are considered a long time in our society and now, it extends to the web. Strahilevitz [18] presents that one of the most significant developments during the last decade has been the growing availability of information about individuals, making possible analyze with whom make deals, sells etc.

Donavan and Smith [12] investigated trust models in recommendation systems where trust is estimated by monitoring the accuracy of a profile at making predictions over an extended period. As seen in [19], in many community-based web applications, trust is a very important issue to make users' experience comfortable.

The reputation of a digital content producer, i.e. a user that provides digital content on the Web, involves characteristics like credibility, reliability and so on. If these aspects are not communicated properly or cause breakdowns at interaction time, they can cause misunderstandings and problems for users to complete their tasks successfully. In this research, we focus on understanding the potential breakdowns (in interaction) related to how interactive applications treat the reputation of digital content producers (by whom it is provided and/or endorsed) and communicate it to the end-users.

In this paper, we present and discuss the main results of a two-step study carried out to characterize the reputation model of a mobile application. To understand how the reputation questions affect the human-computer interaction, the first study (Study One) was proposed to investigate how the reputation of the content producers is communicated to users. We chose a mobile application, Waze¹, whose purpose is to promote a smart traffic, where users interact and inform traffic conditions in real-time, aiming at collaborating and helping other users. Applying the Semiotics Inspection Method (SIM) [17], by two researchers, it was possible to: (a) analyze how Waze deal with reputation issues; (b) identify the Waze strategies to classify an information as reliable or not; (c) identify the potential breakdowns in the communication of reputation issues.

In Study Two we conducted an empirical experiment to observe (in practice) how users recognize (or not) the signs and the reputation model of information sources in Waze app. The results show that many users are not aware about reputation of digital content producers in community-based apps. It was also possible to confirm (and find out new ones) some breakdowns in communication strategies of Waze, identified in Study One.

This paper is divided as follows. In Sect. 2, we present an overview of some related work. In Sect. 3 we detail de methodology and findings of our study. Finally, in the Sect. 4, we conclude and discuss the results that we have found.

2 Reputation in HCI

People often base their relationships with others in values like trust, reliability and reputation. When we heard some news (in digital medias or not), for example, it would be appropriate to consider the content producer to decide if that information is reliable or not. In other words, the reputation of a content producer is an important issue to evaluate something as reliable or not.

¹ Waze, <https://www.waze.com/>.

As the amount of information produced by community-based apps increases, it is necessary to recognize the importance of reputation issues in this kind of places. Over the last 15 years the HCI research community has been researching this topic [13, 14, 21]. Dwyer and co-authors [4], for instance, studied social networks, where privacy and reliability are barely perceived by the end-users. The authors investigated and compared if the confidence in people in two social networking (Facebook² and MySpace³) affect the desire of sharing information and establishing new relationships. Facebook is supposed to guarantee the authenticity of its members due the associations with physical entities (e.g. university). MySpace, in turn, has a bad reputation in terms of reliability. Despite the study limitation (the veracity of profiles was not considered), the main findings include: subjects from Facebook and MySpace expressed similar levels of concern regarding internet privacy, but the members of Facebook were more trusting of the site and other members. Although the members of MySpace are more active in developing new relationships. These results show that the interaction of trust and privacy concern in social networking sites do not show enough to create a model of behavior and activity in an accurate way.

Ganesh and Sethi [5] presented a work about trust and reputation in Social Networks, showing empirical results from Facebook Reputation System where a score is associated to a person's profile, reflecting their extreme negativity or positivity. In this study, the users are divide into two groups: personal and professional. In the personal group, profile features like predictability, care, expertise, altruism and honesty. In the professional group: leadership, organization, punctuality, reliability and expertise. As limitation, the authors relate that the reputation can change over time and the proposed model does not cover that. Another limitation refers to engagement of the users.

Kittur and co-authors [7] discuss some risks related to trust in online environments such as Wikipedia. In this collaborative based environment any user can include, edit e consume the information available in the articles there. Among the risks, stands out: precision (not knowing if the subject is accurate, frequently perceived for lack of references); reasons (not knowing the reasons' editors, which can deflected for many reasons); expertise (it is not possible to know the level of editors' knowledge; stability (it is not possible to know the number of changes in the paper). The work presents a list of best practices to improve reliability, such as including history of record changes involving subject and their authors. One of the metrics founded in the study is to use percentage of words included by anonymous users. These users are supposed to offer more chances to commit vandalism and spam. Zeng and co-authors [22] use Dynamic Bayesian Networks to calculate the evolution of trust using as input status of editors' writing and inserted texts.

Luca and Zerva [9] relate that most researches about trust present as study case e-commerce situations and is focused on credibility, not trust. To understand how users consider a site as trustable, some aspects were analyzed such as: design look, information structure, information focus, underlying motive, usefulness of information, accuracy of information, name recognition and reputation, advertising, information bias, tone of the

² www.facebook.com.

³ www.myspace.com.

writing, identity of the site operator, functionality of the site, customer service, past experience with the site, information clarity, performance on a test, readability of text and site affiliations. In e-commerce sites, for instance, the most relevant features are name recognition and reputation; in new sites, information bias; in nonprofit organizations sites, less information structure is considered important. In Opinion/Review sites, information sites and information accuracy; in Travel sites, customer service. In web search sites, information design, functionality, advertising.

Massa and Avesani [11] analyze the potential contribution of trust metrics to get better performance in Recommendation Systems, presenting a filtering process that can be informed between users reputation, being propagated by other users through evaluations and the trust model, creating a trust network.

Some studies also were focused on crowdsourcing environments. Alperovich and co-authors [1] propose a system and a method to calculate the reputation of mobile apps. It is done by collecting some attributes of mobile apps and its behavior. It is compared with crowdsourcing data in real-time, changing the reputation score of the application. Accordingly to a score, an application is classified in a level that represents its reputation and a probability to be malicious, identifying some risks.

Varshney and co-authors [20], in turn, suggest through coding, to promote the trust in crowdsourcing environments using control of errors and mathematic models. This work differs from ours because mathematic models are used to indicate when an information can be considered as trust or to identify attacks by malicious crowd workers. Our work uses a different methodology to comprehend these questions: our focus is on the communicability of reputation issues and how it affects interaction.

In HCI research area, we found the work from Shneiderman [15], that claims that the publication of users' past performance patterns and a rich feedback about subjects and authors/users are the best practices to increase the trust in this kind of environments. It presents guidelines to designers involving trust in online experiences. The main ones are disseminate past performance history, indicate references to create reputation systems, obtain certification of third parties as stamps of approval.

Our work presents an exploratory research to understand better the communication of the reputation questions of digital producers.

3 Investigating Reputation Issues in Waze App

The purpose of this study was to investigate user's understanding of reputation issues in a Community based Mobile App and the potential implications to HCI. The general research question we were asking was: How users recognize (or not) the signs and the reputation model of information sources in community-based apps?

This study is part of a broader research to investigate reputation issues in crowdsourcing apps. In our study, we chose the app Waze (shown in Fig. 1), which is about a smart traffic through information in real-time posted by users. In Study One, we conducted a semiotic [16] inspection and found out that Waze has some strategies to classify whether an information posted by a user can be considered trustworthy or not.

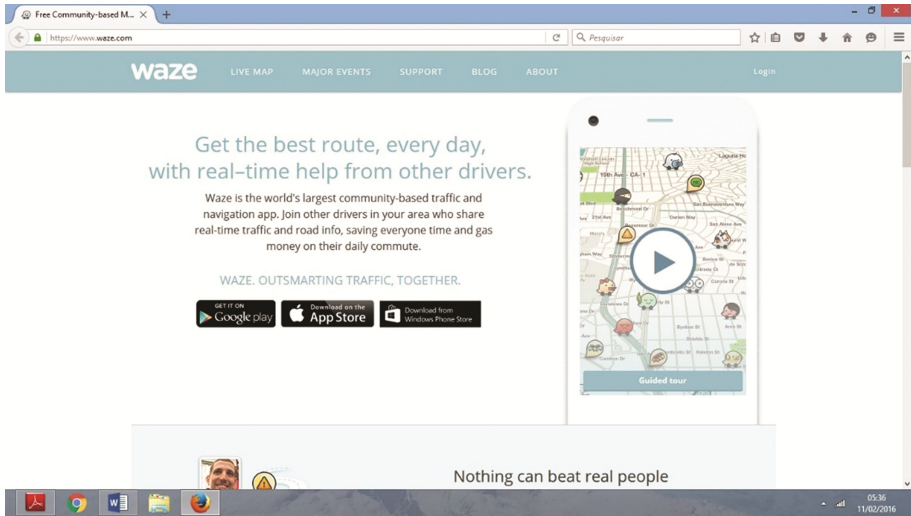


Fig. 1. Waze website (<https://www.waze.com/>)

Other findings from Step One show that there are potential breakdowns in the communication of reputation issues. Waze has some strategies to classify an information as trustworthy or reliable as said before, but these forms to help the users are now well communicated to the users, causing some problems to understand these questions. In summary, with the application of the Semiotics Inspection Method (SIM) [11], it was possible to: (a) analyze how Waze deal with reputation issues; (b) identify the Waze strategies to classify an information as reliable or not; (c) identify the potential breakdowns in the communication of reputation issues. With respect to category (b), Waze offers verification mechanisms to try to ensure the integrity and accuracy of information that are: alerts and number of thanks and comments; threshold alerts per day; option “does not exist”. To prevent cheating and abuse of function “alert”, Waze put a limit on the number of alerts that each person can make in one day or even in one hour. If a user informs something that is incorrect or wrong, he may lose points of his classification level or be prohibited to post information in the app for a period. However, it happens only if the user report that the warning does not exist to the app. If a person decides to block the street for personal proposals, for instance, and a group of friends confirms that information, it can cause incorrect decisions to other drivers. Some users could assign that this message was incorrect, but what could happen to other drivers while the information was still visible? The number of comments associated to a post may influence the Waze reputation model, because it keeps the information available in the app for more time than usual. Our study also revealed that the number of “thanks” in the app suggests that the information is reliable. But, sometimes the sign of “thanks” does not represent that the users agree with that information. The users usually use that sign only to thank to the information posted.

As to the category (c), Waze is not clear about how the user is informed of the reliability of information. The communication lies with the interpretation of each user,

and is therefore a point where the application should be improved. The track interdictions can be communicated with the hazard classification, works or events. Another limiting factor is that the information be available until the other user to enter its non-existence, leaving in the interval available information. It is not clear to the user when a source is reliable or not. For those users who know the application documentation, information that has a higher number of required become more reliable. Each user has an avatar that can change its level before some tasks such as driving a specific distance. Some problems were identified in this kind of situation. This resource (the avatar) is to promote engagement of the users, but in other situations, it could be the unique form that the user likes to customize his profile. At the map, when a user click on an avatar, some information is shown like: number of points, a classification of the user, how long the user is on Waze, etc. However, it is difficult to get these information on the fly, when the user is driving. Furthermore, many taxi drivers drive all the time during the week and get many points only for driving. So, big scores in Waze, sometimes, do not correspond to the idea of trust.

To a better comprehension about how and whether it affects the users' interaction, Study Two focused on the reception of users.

3.1 Methodology

We used a qualitative approach because it is especially appropriate for studies like ours [2, 3], which explores intensively and at greater depth a specific research question. Our primary data was produced by six participants (P1, P2, P3, P4, P5 and P6). The main empirical evidences were collected in post-evaluation interviews: the participants' discourse about the experiment and answer some questions about the test and reputation issues in Waze. Secondary empirical data was collected from the questionnaire sent before the tests including question about profile of participants and questions about reputation. The participants' interviews were analyzed separately, using discourse analysis techniques. This analysis consisted of a systematic exploration to find out major meaning categories in discourse relating intra-participant analysis and inter-participant analysis. In order to run the empirical study, a real route was selected. To conduct the tests, a scenario of use was elaborated, to motivate the participants and guide them about the tasks they should achieve during the study. This scenario presented a situation where a worker used the Waze to get the best route to a specific destiny, considering trust and the information reliability. The researcher also included a fake warning (in the Wazes' route) to observe the participants reaction. In order to avoid distractions we elaborate a questionnaire that prepared our participants to comprehend a little about the question of reputation. Beyond this, we introduced Waze and explain the objectives around the tests and characteristics about the route selected. It was possible, at any time of the test, that the participant interact with the interface of the app.

Participants and Procedures. The 6 (six) participants had little experience with Waze app in real experiences such as using the app while driving a car. The entire experiment had a duration of about 30 min.

Firstly, a pilot test was conducted to verify the viability of the test and making some adjusts. The participants were invited to answer a questionnaire and after that, inside the car, some aspects of Waze were presented and the scenario of use. Each participant should act like the character of the scenario. Then, after this, an interview was conducted to comprehend better some situations during the test and to understand how the participants lead with the reputation questions.

4 Experimental Results

The empirical data collected in the tests was examined using discourse analysis technique. It generated categories which are part of the broader results of the research. We arrived on evidences provided by six specific subcategories of meanings: (i) the participants are not worried about reputation issues in this kind of app; (ii) the participants do not recognize a reputation model in Waze; (iii) the participants believe that to trust in the information posted in the app is necessary another kind of knowledge; (iv) the participants credit to the app the provided information reliability; (v) the participant trust in information provided by friend and family.

As evidence for (i), we see the following excerpts:

P1: "If it was easy to post, I posted". The criterion that I used to select a route was the short time".

P2: "To help the Waze working well, I put the information, but I don't have this habit". "I used as criterion to select the route, the smaller traffic".

P3: "The selected criterion to choice a route is the smaller distance (if it shows)". "I posted the information, because it is an app which depends on that to help other users".

P4: "If I know how to post the information, I posted. Nevertheless, if I was driving, I do not. I don't have the habit to do this".

During the interview, a question was included: If you are driving and heard about an accident in the road, would you post this information? The point here is identify if the participants are worried about their own reputation and if they look for more information before post something. It demonstrate that participants do not matter about the information sources.

As evidence for (ii), we see in the following excerpts some breakdowns in communication of Waze and the reception of users:

P1: "I didn't see any element in the interface different from the others to give best routes. It presented the same icons etc.".

P2: "When we pass around the mall, it show me the traffic. It is important to help the drivers".

P3: "The app show the traffic condition (heavy or low)".

P4: "I think that trust in an information posted is responsibility of each user because it is not possible to verify by the app".

P5: "I guess if the app show the number of evaluations about an information, it would be a better model to trust in that information".

Many elements present in the interface of Waze are not well communicated to the users. The participants did not see aspects like information sources, accidents and other conditions of traffic during the tests. Therefore, we identify some ruptures in communicating these elements.

As evidence for (iii), we see in the following excerpts that the participants think that is necessary to have some knowledge to compare with what is showed by Waze.

P2: "In an unknown city I trusted based on following example: if in the cities that I know it works well, it will work well there too".

P3: "I trusted only if I can compare it with other app similar".

P4: "If the app suggest a different place from I know, I don't select the suggested route".

P5: Places that we know by magazines, tv programs are shown with intense traffic and I pass there and the traffic is that way, I can trust".

Waze presents as one of its strategies to identify an information the number of likes. However, the participants associated this to the Facebook, not associating to the idea of confirmation of trust.

As evidence for (iv), we see in the following excerpts that sometimes users credit to the app the function of analyze if an information is reliable.

P3: "The app show the traffic condition (heavy or low)".

P5: "This information posted on Waze means that someone inform that? Or not? Because only one or two can post something and it will be considered reliable"?

The choice of a direct observation was due to the difficulties in reproducing traffic situations on the laboratory. Participants were recruited after answering a questionnaire about their technological profile and experience about social media, especially mobile community based apps. During the study, their task was to use the Waze app to find a route to a specific address in the city, in a typical scenario. In this route, each participant analyzed the traffic conditions and the content confident. We then interviewed participants, individually, to find out what meanings and use they associated with reputation while using Waze app.

As evidence for (v), we see in following excerpt that a known information by the participants is easier to trust in.

P5: "One of criterion adopted by me to consider an information as trust is its origin. If the information source is my friends or my family, I can trust".

The following evidences was identified: the participants have not provided no care about their own reputation, posting information listened by third parties without confirming its veracity and validity, attributing ease of use as criteria as task. The participants did not perceive elements of interface used by Waze to communicate features reputation and trust. In this point, the users did not demonstrate preoccupation in knowing information sources to analyze and make their decisions. So, the reputation is not considered when selecting a route. Waze bring many information about information sources like avatars, time of use, etc. However, the users recognized none of them during

the tests. Another aspect that was identified is: many times the participants consider localization or previous knowledge to classify an information as trust forgetting the role of the participants that are subject producers. This way, the responsibility and veracity of information was attributed to the app. Aspects like reliability and credibility only are considered when the participant compare the results demonstrated by the app with other one on same function. The doubt of information and reputation only appear when errors are detected (wrong routes etc.). Some participants attribute totally to the app the information that are posted and others understand the role of users (producing and consuming information in real time).

The participants presented a profile that use the app many times only to know a route in an offline way, not demonstrating so much interesting in real time traffic. As seen in Study One, some breaks in communicating strategies of confirm the information and its authors happen. The Reputation Model adopted by Waze is not comprehended so well by the users.

5 Conclusions

This research intended to understand how users deals with reputation issues, i.e., given some content, how they recognize (or not) aspects related with the credibility and reliability of the content producers. Another aspect of reputation analyzed in this study is how the users see their self-reputation and if they are worried about this. We conducted 2 studies: one focused on how the message is communicated to the users, other has focused on reception of this message, by the users.

Our findings show that reputation is something that the apps needs to work better. The reputation models of apps are weak and communicate not so well about it. It should exist more resources to help the users to make decisions based on information posted by other users.

During the test was possible to see how difficult is to drive and use the app to get more information. Many decisions are made in a question of seconds and how better it is communicated to the users, more important it becomes.

It is important to develop in an easy way to communicate this kind of information, because it is not possible to stop the car in any part of a road to verify what is happen. The fake warning created to see the reaction of the participants was not communicated by Waze and the participants didn't see it.

The evidences collected in the Study One have improved our understanding of the interactive strategies to communicate reputation issues in crowdsourcing apps. The empirical evidences from Study Two showed to us that novice users are not aware about the app mechanisms to deal or classify the reputation of a content and they believe that content reliability is provided by the app. With a characterization of the reputation model of a community based mobile application we aim at give a contributing to HCI design process.

In the results we can also understand see that at many times the users are not aware of their own reputation beyond to attribute to the app itself the function of communication traffic conditions, ignoring the collaborative aspects of the app. The model reputation of Waze is not perceived correctly by the users as the study demonstrates.

To help the users to decide better the best route and other situations of traffic, the designers should making easier to understand the reputation model of their apps. In a technology age, many users do not think about their digital attitudes and consequences yet.

This work is our first attempt in trying to characterize the reputation model of a community based mobile application. As next steps, we intend to compare the reputation model of another community-based app to confirm (or not) what was discovered in this study. This experiment was made on Sunday, at the city Campos dos Goytacazes, Rio de Janeiro. In this day of week, the traffic is less than days like Monday, Tuesday etc. Maybe it be interesting to apply the test in these days and verify if the users recognize the same signs or have the same behavior.

References

1. Alperovitch, D., Krasser, S., Brinkley, M.: U.S. Patent Application No. 13/426, 363 (2012)
2. Creswell, J.W.: *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. Sage Publications, Beverly Hills (2013)
3. Dezin, N.K., Lincoln, Y.S. (eds.): *The Landscape of Qualitative Research*. Sage Publications, Thousand Oaks (2008)
4. Dwyer, C., Hiltz, S., Passerini, K.: Trust and privacy concern within social networking sites: a comparison of Facebook and myspace. In: *AMCIS 2007 Proceedings*, p. 339 (2007)
5. Ganesh, J., Sethi, P.: Reputation and trust in Social networks: empirical results from a Facebook reputation system. In: *19th Americas Conference on Information Systems, AMCIS 2013 Chicago, Illinois, USA, 15–17 August 2013*, Association for Information Systems (2013)
6. Josang, A., Ismail, R., Boyd, C.: A survey of trust and reputation systems for online service provision. *Decis. Support Syst.* **43**(2), 618–644 (2007)
7. Kittur, A., Suh, B., Chi, E.H.: Can you ever trust a wiki? Impacting perceived trustworthiness in wikipedia. In: *Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work*, pp. 477–480. ACM (2008)
8. Kraut, R.E., Resnick, P.: *Building Successful Online Communities: Evidence-Based Social Design*. MIT Press, Cambridge (2012)
9. Luca, M., Zervas, G.: *Fake It Till You Make It: Reputation, Competition, and Yelp Review Fraud*. Harvard Business School NOM Unit Working Paper No. 14-006 (2013)
10. Malaga, R.A.: Web-based reputation management systems: problems and suggested solutions. *Electron. Commer. Res.* **1**(4), 403–417 (2001)
11. Massa, P., Avesani, P.: Trust-aware collaborative filtering for recommender systems. In: *the move to meaningful internet systems*. In: Meersman, R., Tari, Z. (eds.) *CoopIS, DOA, and ODBASE*, pp. 492–508. Springer, Berlin (2004)
12. O'Donovan, J., Smyth, B.: Trust in recommender systems. In: *Proceedings of the 10th International Conference on Intelligent User Interfaces*, pp. 167–174. ACM (2005)
13. Resnick, P., Kuwabara, K., Zeckhauser, R., Friedman, E.: Reputation systems. *Commun. ACM* **43**(12), 45–48 (2000)

14. Rowley, J., Johnson, F.: Understanding trust formation in digital information sources: the case of Wikipedia. *J. Inf. Sci.* (2013). 0165551513477820
15. Shneiderman, B.: Designing trust into online experiences. *Commun. ACM* **43**(12), 57–59 (2000)
16. De Souza, C.S.: *The Semiotic Engineering of Human-Computer Interaction*. The MIT Press, Cambridge (2005)
17. De Souza, C.S., De Leitão, C.F.: Semiotic engineering methods for scientific research in HCI. *Synthesis Lectures on Human-Centered Informatics*, vol. 2(1), pp. 1–122. Morgan & Claypool, Princeton (2009)
18. Strahilevitz, L.J.: Reputation nation: law in an era of ubiquitous personal information. *Nw. UL Rev.* **102**, 1667 (2008)
19. Teruel, M.A., Navarro, E., López-Jaquero, V., Simarro, F.M., González, P.: A comparative of goal-oriented approaches to modelling requirements for collaborative systems. In: ENASE, pp. 131–142 (2011)
20. Varshney, L.R., Vempaty, A., Varshney, P.K.: Assuring privacy and reliability in crowdsourcing with coding. In: *Information Theory and Applications Workshop (ITA) 2014*, pp. 1–6. IEEE (2014)
21. Woodruff, A.: Necessary, unpleasant, and disempowering: reputation management in the internet age. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 149–158. ACM (2014)
22. Zeng, H., Alhossaini, M.A., Ding, L., Fikes, R., McGuinness, D.L.: Computing trust from revision history. *Stanford Univ Ca Knowledge Systems LAB* (2006)