Association for Information Systems
AIS Electronic Library (AISeL)

WISP 2012 Proceedings

Pre-ICIS Workshop on Information Security and Privacy (SIGSEC)

Winter 12-14-2013

Misinformation in Online Health Communities

Srikanth Venkatesan
STATE UNIVERSITY OF NEW YORK AT BUFFALO, svenkate@buffalo.edu

Wencui Han
SUNY Buffalo

Victoria Kisekka
SUNY Buffalo

Raj Sharman
State University of New York at Buffalo, rsharman@buffalo.edu

Vidyadhar Kudumula
SUNY Buffalo

See next page for additional authors

Follow this and additional works at: http://aisel.aisnet.org/wisp2012

Recommended Citation

Authors

Srikanth Venkatesan, Wencui Han, Victoria Kisekka, Raj Sharman, Vidyadhar Kudumula, and Hardeep Singh
Jaswal

Misinformation in Online Health Communities

SrikanthVenkatesan[1]
SUNY Buffalo, United States

Wencui Han
SUNY Buffalo, United States

Victoria Kisekka
SUNY Buffalo, United States

Raj Sharman[2]
SUNY Buffalo, United States

Vidyadhar Kudumula
Deloitte, United States

Hardeep Singh Jaswal
Citigroup, United States

ABSTRACT

Spread of wrong information can be a serious deterrent to information system use especially in case of online community which typically have thousands of end users. However, literature has been weak in linking the prevalence of health misinformation on online social networks to the factors contributing to misinformation. This study seeks to reduce this gap by examining the impact of thread characteristics and user characteristics on the extent of misinformation in online social networking forum related to Parkinson's disease. Our findings show that the correctness of a post is affected by clarity of the thread question, information richness and the user potential for making useful contributions.

Keywords: online health communities, misinformation, thread characteristics, posts characteristics, user Characteristics.

[1]Corresponding author: svenkate@buffalo.edu.

[2]Corresponding author: rsharman@buffalo.edu.

INTRODUCTION

The rising popularity of online social networks has led to an increase in the number of web applications, making it relatively easy for end-users to share information with other internet users. The 2002 Pew Internet & American Life Project estimated that "80% of adult internet users have searched for at least one of 16 major health topics online". The report concluded that "this makes the act of looking for health or medical information one of the most popular activities online" (Fox and Fallows, 2003). In spite of all these studies, existing literature has been weak in identifying the significant factors responsible for misinformation in online healthcare social networks.

Misinformation is an important area of information security research. Spread of wrong information can be a serious deterrent to information system use especially in case of online communities, which typically have thousands of end users. Researchers exploring the inaccuracies of healthcare information on social networks have developed empirical studies that measure and confirm the existence of misinformation on social networks (Scanfeld, et al, 2010; Tsai et al, 2007). Other literature in this area has focused on tools for assessing and measuring the quality of health information published on the internet (Hargrave et al, 2006 and Meric et al., 2002). However, literature has been weak in linking the prevalence of health misinformation on online social networks to the factors contributing to misinformation (G Eysenbach, 2003). This study seeks to reduce this gap by answering the following question: What are the critical factors that influence the extent of misinformation about Parkinson on online social networks? Parkinson's disease is a chronic degenerative disorder of the central nervous system. Thousands of Americans are diagnosed with PD every year, while there is relatively slower progress in researching new drugs and treatments or permanent cure (Dimond, 2010). Parkinson's disease

was chosen for this study for two reasons: firstly, it is chronic in nature and involves many years of slow and steady progression making it suitable for detailed study. Secondly, since it has no permanent cure, patients largely rely on multiple medications and treatments and therefore has more uncertainty involved.

Misinformation in online social networks threatens patient safety even when it does not stem from malicious intent. For example, a user may communicate dosing information that is specific to him/her and potentially unsuitable to the reader. At times the reader of the post maybe allergic to certain medications that he/she may not be aware of, which may negatively impact his/her health. Despite the works such as (Eysenbach 2003), this area of online misinformation remains understudied. To our knowledge, our work represents the first study that examines the impact of thread characteristics and user characteristics on the extent of misinformation. Our findings show that the correctness of a post is affected by clarity of the thread question, information richness and the user potential for making useful contributions. The critical factors identified in this research will provide the necessary information that will prioritize research geared towards improving patient safety at individual and population levels.

RELATED WORK

Despite the growing number of people accessing health information on online social networks, however, the quality of this information is questionable. A large number of researches have studied the quality of health information on social networks. Results from these studies, however, are inconsistent; Several studies have indicated that medical information on the websites and online social networks is sometimes ambiguous, incomplete, misleading, or inaccurate (Corcoran, et al, 2009; G Eysenbach, et al., 2002; Huberman et al, 2009; Meadows-

Oliver and Banasiak, 2010 and Scanfeld, et al., 2010). On the contrary, some studies found the most of the information on the internet was accurate. (Black and Penson, 2006; Murray, et al., 2003). On the other hand, there are empirical findings which concluded that the quality of health information on the internet varies by website (Brewster and Sen, 2011; Scullard, et al., 2010; Tangri and Chande, 2011; van der Marel et al., 2009). These findings reinforce the need to understand online healthcare information exchange and how the information exchanged affects patient safety.

MODEL DEVELOPMENT

The dependent variable of this study is the accuracy of information about PD on online social networks. In online social networks, a discussion thread is a grouping of discussions, responses or comments about a specific topic. In this research, a discussion post is defined as a reply or comment to the question that initiated the discussion thread. The unit of analysis in this paper is post. The accuracy of each post is defined as whether the post correctly answers the initial question. The independent variables are clarity of the thread question, information richness, and user rating. Figure 1 presents the research constructs and model proposed in this research-in-progress.
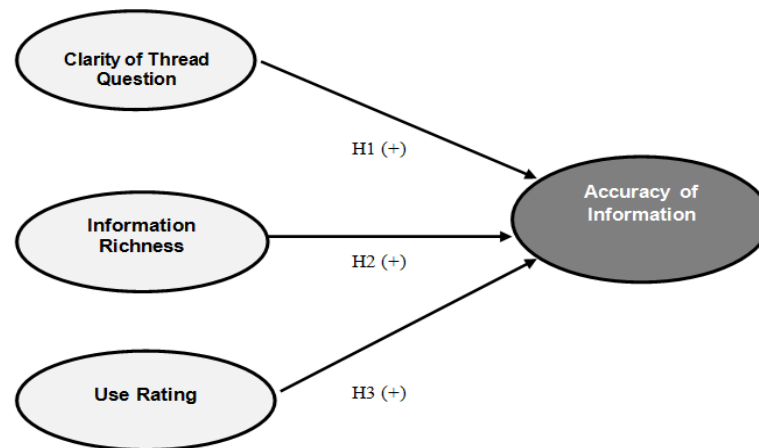
Figure 1.Critical factors affecting information accuracy related to Parkinson's disease.

Clarity of thread question: Clarity of information is a dimension of information quality (Wand & Wang, 1996). Incomplete question(s) in the post initiating the discussion may induce responses that could be classified as erroneous. Often posters seeking advice regardless of whether it is for the purposes of diagnosis, medication, or advice are not aware of what constitutes relevant and appropriate information one needs to convey in order to receive appropriate responses. Habitually the context is intentionally or unintentionally not conveyed. In the event that the information provided is insufficient, there are two possible approaches that other members participating in the thread could take: (a) seek clarification and/or (b) make suitable assumptions and answer the post. Unlike other services that one may purchase where the buyer is more able to express his requirements and evaluate the quality of service, it is difficult to provide information with complete clarity for diagnosis or treatment because the person posting the question or initiating the thread may not be completely aware of what information is needed to provide a useful answer. This leads to the following hypothesis:

H1: *Clarity of thread question positively impacts the accuracy of the information in the posts.*

Information Richness: In our study, Information Richness is defined as the amount of detail contained in a thread. Information Richness is measured by the number of posts the thread already has prior to a given post. There have been influential studies suggesting that the quality of user-driven content improves as the number of users contributing increases (Magnus, 2008). It is also generally argued (particularly in the crowd sourcing literature) that large groups of people have more collective intelligence than small groups (Surowiecki and Silverman, 2007). Based on supporting literature our expectation is that the information richness should positively impact the accuracy of information on an online social network. Therefore,

H2: *Information richness positively impacts the accuracy of information in the posts.*

User rating: Users of the social network may have different levels of expertise with regards to the knowledge of Parkinson's disease and its treatment. They may also have different level of activeness and likeability. In our study, user rating is measured by the number of useful posts a user has posted in the past. For every useful post, users of the website give a "thumbs up" rating. This information is available from the social network and it indicates the knowledge of Parkinson's disease the user has, the active level of the user and the length the user has using the social network. Our expectation is that users with higher level of expertise are more likely to provide accurate information.

H3: *User rating positively impacts the accuracy of information in the posts.*

METHOD

The data was collected from threads and posts related to Parkinson's disease from a very popular social networking. 584 posts were extracted from the social networking site and ported into an application specifically developed for physicians to do the evaluation. The physicians evaluating were neurosurgeons working in the US. The accuracy of the posts is a binary variable based on the physicians' evaluation of whether the post correctly answers the question in the initial post. The clarity of thread question is also a binary variable obtained from physicians' evaluation. The measurement of information richness (how many posts have been posted in the thread) and the measure of user rating (total number of useful posts posted by the user) are both publicly available from the website.

RESULTS

We used logistic regression to test the research model. The following equation represents our logistic regression model, where p is the probability that an event Y occurs and the term $[p/(1 – p)]$ is the odds ratio, noted as Exp (B) in the results (Table 1): Ln $[p/ (1 – p)] = α + β1$ Clarity of thread question $+ β2$ Information richness $+ β3$ User rating $+ e$

Table 1.Logistic Regression Results.

Variable Type	Variable Name	B	Sig.	Exp(B)	95% C.I.for EXP(B)	
					Lower	*Upper*
Thread characteristic	Clarity of thread question	.819	.000	2.267	1.476	3.482
Information characteristic	Information richness	.033	.017	1.033	1.006	1.061
User characteristic	User rating	.001	.001	1.001	1.000	1.001
	Constant	-.695	.091	.499		

Our findings show that misinformation of a post depends on the clarity of the question posted by the first user who created the thread. Our study also finds that latter posts are more likely to correctly answer the thread question. It may be noted that the user characteristic as defined by the number of useful contributions the user has made in the past significantly affects the likelihood of the information of the post being correct.

DISCUSSION AND CONCLUSION

Misinformation is undesirable because when patients receive incorrect advice from the social networking sites and act on it negative health outcomes are highly likely. Our study examines the user characteristics, thread characteristics and their impact on the correctness of posts. This is very valuable to online social network sites providers as they can flag posts by users who have higher likability thus indicating to the user about the higher likelihood the information being correct. Further, social networking sites can employ mechanisms in place to ensure that thread questions are clear. This contributes to decreasing the extent of misinformation in social networking sites, an important area of information security research. This paper therefore makes both theoretical and practical contributions. Our work has several limitations. Future studies can consider a richer set of thread and user characteristics to get a more complete picture on the factors that affect the extent of misinformation. The literature for this work can also benefit from the work has been done in the area of information quality.

REFERENCES

Black, P. C., &Penson, D. F. (2006).Prostate Cancer on the Internet--Information or Misinformation? The Journal of Urology, 175(5), 1836-1842. doi: 10.1016/s0022-5347(05)00996-1

Brewster, L., & Sen, B. (2011). 'Quality signposting': the role of online information prescription in providing patient information. Health Information & Libraries Journal, 28(1), 59-67. doi: 10.1111/j.1471-1842.2010.00912.x

Corcoran, T. B., Haigh, F., Seabrook, A., &Schug, S. A. (2009). The Quality of Internet-sourced Information for Patients With Chronic Pain Is Poor. [Article]. Clinical Journal of Pain, 25(7), 617-623.

Dimond, PF (2010). "No New Parkinson Disease Drug Expected Anytime Soon". GEN news highlights. http://www.genengnews.com/analysis-and-insight/no-new-parkinson-disease-drug-expected-anytime-soon/77899336/.

Eysenbach, G. (2003). The impact of the Internet on cancer outcomes. CA: A Cancer Journal for Clinicians, 53(6), 356-371.

Fox, S., & Fallows, D. (2003). Internet health resources: health searches and email have become more commonplace, but there is room for improvement in searches and overall Internet access: Pew Internet & American Life Project Washington DC.

Hargrave, D. R., Hargrave, U. A., &Bouffet, E. (2006). Quality of health information on the Internet in pediatric neuro-oncology.Neuro-Oncology, 8(2), 175-182. doi: 10.1215/15228517-2005-008

Huberman, B., Romero, D., & Wu, F. (2009). Social networks that matter: Twitter under the microscope. First Monday, 14(1), 8.

Magnus, P. (2008). Early response to false claims in Wikipedia.First Monday, 13(9).

Meric, F., Bernstam, E. V., Mirza, N. Q., Hunt, K. K., Ames, F. C., Ross, M. I., . . . Singletary, S. E. (2002). Breast cancer on the world wide web: cross sectional survey of quality of information and popularity of websites. BMJ, 324(7337), 577-581. doi: 10.1136/bmj.324.7337.577

Meadows-Oliver, M., &Banasiak, N. C. (2010).Accuracy of Asthma Information on the World Wide Web.[Article]. Journal for Specialists in Pediatric Nursing, 15(3), 211-216.

Murray, E., Lo, B., Pollack, L., Donelan, K., Catania, J., Lee, K., . . . Turner, R. (2003). The impact of health information on the Internet on health care and the physician-patient relationship: national US survey among 1.050 US physicians. Journal Of Medical Internet Research, 5(3).

Scullard, P., Peacock, C., & Davies, P. (2010). Googling children's health: reliability of medical advice on the internet. [Article]. Archives of Disease in Childhood, 95(8), 580-582. doi: 10.1136/adc.2009.168856

Surowiecki, J., & Silverman, M. (2007).The wisdom of crowds.American Journal of Physics, 75, 190.

Scanfeld, D., Scanfeld, V., & Larson, E. (2010). Dissemination of health information through social networks: Twitter and antibiotics. American journal of infection control, 38(3), 182-188.

Tangri, V., &Chande, N. (2011).Quality of Internet-based information on gastrointestinal diseases.[Article]. Canadian Journal of Gastroenterology, 25(2), 93-96.

Tsai, C., Tsai, S., Zeng-Treitler, Q., & Liang, B. (2007). Patient-centered consumer health social network websites: a pilot study of quality of user-generated health information.

van derMarel, S., Duijvestein, M., Hardwick, J. C., van den Brink, G. R., Veenendaal, R., Hommes, D. W., &Fidder, H. H. (2009). Quality of Web-based Information on Inflammatory Bowel Diseases.[Article]. Inflammatory Bowel Diseases, 15(12), 1891-1896. doi: 10.1002/ibd.20976

Wand, Y., & Wang, R. Y. (1996). Anchoring data quality dimensions in ontological foundations. Commun. ACM, 39(11), 86-95. doi: 10.1145/240455.240479