Sociotechnical Synthesis

STS 4600

Spring 2021

Eric Stoloff

Computer Science

Signature _____Eric Stoloff_____Date_____

Eric Stoloff

_____ Date_____

Richard D. Jacques

Introduction

The proliferation of false information on social media is an incredibly important problem in today's society that needs to be addressed, from its impact on the 2016 Presidential Election to its inflammation of the Covid-19 crisis. For my capstone project in general, I wanted to explore how false information spreads on social media in order to better understand the problem and potentially find the best solution. In my initial research, it became glaringly obvious that this was a human problem. People's apparent inability to recognize false information is due to many factors, including emotions and worldview, and I wanted to address this. In addition, my research revealed that effectiveness and accuracy were the only criteria with which solutions to prevent the dissemination of false information were evaluated on. My tightly-coupled research paper and technical project expanded on these shortcomings, and the two works complement one another. My STS research paper explores the specific psychology-related reasons why false information spreads, and evaluates a diverse set of solutions according to a more extensive list of criteria. My technical project proposed the implementation of short false information awareness activities into the user interface of social media sites, using the psychological principles uncovered in my research. Together, these works provide a different perspective from which to view the issue of misinformation circulating on social media and stress the importance of going in a new direction when it comes to solving this problem.

Evaluation of Existing Solutions and Future Directions

In my STS research, I evaluated solutions that impede the spread of false information on social media on an expanded set of criteria that are vital to consider in the context of this issue. These include effectiveness, but also the time efficiency of the solution, the ethical concerns it may bring about, and practicality and scalability of implementation in real social media platforms. To begin the work, I delved into the psychology behind users sharing and interacting with information on social media because the specific problems are so human-centric. Then, I

evaluated a number of solutions representative of the different classes of false information prevention methods. I divided these solutions into reactive solutions, which respond to specific instances of false information and attempt to correct them, and active solutions, which reshapes the way that information is spread in a network and modifies the user's interaction with it. According to how these solutions match up to the evaluation criteria, it became clear that active methods are the best direction to pursue in the future. In addition, this is supported by the ineffectiveness that corrections or debunking have on users, who could even experience backfire effects where the correction further entrenches the misinformation with the user. In particular, the active solution that showed the most promise was using the idea of active inoculation in order to improve users' ability to recognize false information. Although this solution is not perfect, my research paper concludes that it has great potential to expand and improve in practicality and effectiveness, which are the only two criteria it did not fully satisfy.

Active Inoculation Against False Information Techniques

Further expanding on the idea of active inoculation for users, my technical project produced an approach by which the idea can be integrated into the user interfaces of social media platforms. Active inoculation works by exposing the user to weakened or obvious pieces of false information in order to stimulate critical thinking and educate them about the manipulation techniques that false information employs to facilitate its spread. The implementation of active inoculation that has been researched in the most depth is the game *Bad News*, which has been proven to increase users' competence in identifying false information. My project leverages ideas for this game and spreads them out over the total time that a user spends on social media platforms. It proposes brief, occasional awareness activities to the user in which they only have to answer one question correctly in order to get back to their regular use of the platforms. For these activities, my project included the design of mockups that illustrate how these activities can be implemented in social media user interfaces. One of their

important characteristics is the focus on manipulation *techniques* and nonpartisan nature. This enables these activities to reach a diverse group of users and retain their effectiveness, instead of telling users what to believe on specific issues and alienating others. With more, but scattered repetition possible through my implementation of active inoculation, it improves in both practicality for users and effectiveness, and the project shows real promise for mitigating the proliferation of false information on social media.

Conclusion

From my research on many different false information prevention methods, it is clear that the way in which this problem is currently addressed by social media platforms needs to be amended. Reactive methods of stopping its spread, such as deep-learning and fact-checking approaches, have too many drawbacks, including the time required for them to work and their practicality to implement across the social media landscape. They also fail to take into account that people will accept misinformation that aligns with their beliefs and worldview, regardless of corrections intended to highlight that information is in fact false. Therefore, it is clear that active solutions are the preferred route to go in the fight against false information. My project builds on methods proven with previous research and seeks to expand the inoculation of users against misinformation techniques in order to bring "herd immunity" to social media platforms and make them more informed places. A further benefit to my project is that it does not raise any real ethical concerns towards users, as it does not involve collecting private data and treats all users in the same way. Overall, this project would have a large positive impact on the social media climate, and if implemented by social media platforms, it would make users more informed and much less susceptible to the dangers of false information.