Toward Browser-based Interventions to Tackle Misinformation Online

PRERANA KHATIWADA, University of Delaware, USA
IAN MUMMA, University of Delaware, USA
LUKE HALKO, University of Delaware, USA
ANESEH ALVANPOUR, University of Louisville, USA
MATTHEW LOUIS MAURIELLO, University of Delaware, USA

The spread of misinformation online and its impact on society have become a pressing issue for both the technology industry and democracies around the world. Developing new tools and methods to mitigate this threat is critical. In recent years, "fake news" has received increasing attention because the term covers intentionally false, deceptive stories as well as factual errors, satire, and stories a person does not like. Few users possess the digital media literacy skills necessary to navigate these challenges. In this workshop paper, we propose developing a mixed-initiative, crowd-powered system to (i) semi-automatically explore definitions of fake news online, (ii) curate datasets of new stories for machine learning applications, (iii) power browser-based tools that identify misinformation content *in-situ*, and (iv) deploy digital media literacy interventions and news recommendations that increase a user's hardiness against misinformation threats on online news media platforms. Our objective here is to stimulate discussion around our prototype's design, new mitigation approaches, and their applications.

Additional Key Words and Phrases: fake news, misinformation, crowdsourcing, digital media literacy.

ACM Reference Format:
Prerana Khatiwada, Ian Mumma, Luke Halko, Aneseh Alvanpour, and Matthew Louis Mauriello. 2022. Toward Browser-based Interventions to Tackle Misinformation Online. The Designing for Mis/Disinformation Workshop (CHI '22 Workshop), (April 2022), 6 pages.

1 INTRODUCTION

Searching for information and news online has become commonplace, whether through Google, Twitter, Facebook, newspapers, or local media websites. These platforms have tremendously improved access to information. However, "fake news" and misinformation spread rapidly as well, because individuals who come across misleading news on media platforms may actively spread it through deliberate sharing [8], or simply by "liking" the content, which often triggers the algorithms of media platforms to display more of this content to them and others in their network. Misinformation is often a piece of fabricated information that is presented as fact and intentionally spread to mislead users, despite the fact that it can be proven wrong by fact-checking [15]. Because the phrase "fake news" encompasses both purposefully misleading and deceptive articles as well as factual errors, satire, and things that a person just does not like, it can be difficult to spot [3]. Others have defined six main types of "fake news" as satire, parody, fabrication, manipulation, advertising, and propaganda [2]. As overconsumption of such content can have negative implications for our democracy (e.g., increasing levels of polarization [14], legislative gridlock [16], violence [28], etc), we propose developing a mixed-initiative, crowd-powered system that can help semi-automatically assess online fake news and misinformation as well as deliver interventions to increase user hardiness against these threats. As part of developing our system, this work explores various research questions, including: *What interactions do users have with online misinformation? How might different aspects of aggregated data from browsing sessions (e.g., demographics, content) inform browsing interventions and news recommendations? And, what kinds of interventions are helpful in increasing exposure to diverse, authoritative news content?*

Part of the Designing for Mis/Disinformation Workshop (Virtual)
at CHI '22, April 30 – May 5, 2022, New Orleans, LA, USA

2 RELATED WORK

Misinformation, its spread, and mitigation strategies are being addressed from a range of perspectives, including computer science, social science, journalism, and psychology. Fallis [12, 15] examine the ways people have defined disinformation as opposed to misinformation and conclude that disinformation is false information that misleads and deceives people and online users. In recent years, a number of tools, approaches, and manual fact-checking websites have become available, including Snopes and Politifact, which are aimed at identifying misinformation [13]. However, the speed and scale with which misinformation spreads through the internet limits their ability to manually check news content [24]. Other intervention tools, such as the Ground News Bias Checker [1], give information about publisher bias, but users are unable to provide their own ratings *in-situ* or obtain data about their past consumption. Fake News Alert [4] and B.S. Detector [5] are browser extensions that rely on manually compiled lists of misleading websites. Other examples include CredEye [24] and FactMata4 [21] which require users to read online news to first identify the claims in the articles that might be false and then go to another website to fact-check them. Fake news stories, on the other hand, are unlikely to be tagged since fact-checking takes longer than creating false information [23]. Moreover, tools for fact-checking tend to flag things that are provably untrue, rather than biased or misleading coverage of actual events. Our work focuses on developing a crowdsourcing tool (rather than relying on professional fact-checkers) to assess the credibility of news websites while gathering web usage data that may inform intervention design.

To date, there have been numerous studies that analyze web usage data. For example, tens of thousands of user browser histories were examined in [14] to understand political polarization in online news consumption. The results indicate that users engage more thoroughly with news articles associated with their own leanings as measured by dwell time (i.e., the period of time during which a user actively engages) and that explicit decisions play an important role in shaping their online news consumption that is not explained by hyperlink clicks alone [Ibid]. Similarly, The DERI Online Behavior Study (DOBBS) [30] makes use of a browser add-on that enables researchers to anonymously and securely follow internet users' browsing habits. However, motivating users to participate is very challenging because, as a passive logging tool, there are no interventions and users do not receive any direct benefits. Our work builds off this prior work by exploring the integration of interventions such as personal informatics dashboards and accuracy nudges.

User-facing interventions into browsing experiences are a relatively new approach to mitigate the influence of misinformation. As a result, it is unclear what approaches might be effective. More interactive approaches have focused on accuracy prompts, which work by shifting attention to the accuracy of a given headline and are typically accompanied by a prompt or "nudge" asking whether the user thinks the following headline is accurate. However, the application of accuracy prompts in other domains, such as biased consumption of online news, as well as the long-term effects of such prompts on deliberation have not been explored [22]. In our work, we plan to combine accuracy nudges with interactive personal informatics and passive logging to allow us to monitor the immediate impacts of interventions on browsing behaviors. For example, *if we report that a user is mostly reading left-wing content, would they explore some right-wing content as well? Or, would they simply ignore the bias of their consumption?* Moreover, our long-term goal is to evaluate whether any observed effects persist over time.

3 CURRENT WORK

Our current work focuses on developing and deploying a prototype Chrome browser plugin that can track the user's web browsing activity, save online news media annotations, display information regarding news media consumption bias, and guide their news-browsing experience through a series of interventions aimed at improving digital media literacy skills. Our tool combines and displays personal consumption data and provides real-time news source bias information. Similar to prior work (e.g.,[30]), we collect domain information and

[1] Ground News Bias Checker: https://chrome.google.com/webstore/detail/ground-news-bias-checker/agleiimpggapjekcdhdjbmegjbbkleie

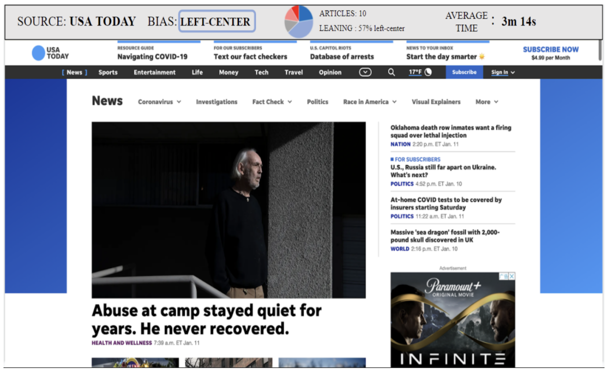

Fig. 1. Our prototype Chrome browser plugin (top) allows users to view the aggregated bias rating of a new site (left), a real-time breakdown of their news media consumption for the day (center), and other metadata associated with their online behavior (right). The tool currently lists about 500 publishers, and this can be expanded further by community annotations.

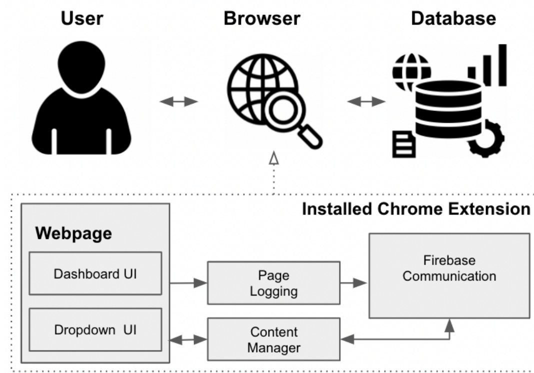

Fig. 2. Our plugin depends on four key components: a Firebase database, the user's browser, a dynamically injected personal informatics dashboard, and a dropdown UI widget. The backend is powered by the Firebase real-time database server. The front-end is built using HTML, Bootstrap, jQuery, D3.js, etc.

the amount of time a user spends online in a privacy-preserving manner. When users access a known news or misinformation domain, a toolbar (Figure 1, top) is dynamically inserted into their page, allowing them to view biased information about the publisher. The primary distinctions between our work and that of existing tools such as Ground News Bias Checker and Media Bias Fact Checker [2] are the addition of functionality that allows users to annotate content with their view of its bias and other journalistic qualities (e.g., use of inflammatory words) as well as users' being able to see personal informatics data about their own daily media consumption.

Our plugin was developed as a web-based application (Figure 2). We used an existing dataset from AllSides [1], to allow our tool to recognize common online news publishers. The dataset contains biased ratings for over 500 online news outlets. We also gather the following information for each user: every domain they visit, the length of time they spend on the domain when they are actively viewing it, and the full URL of every unique news story they read (organized by day). With this data, we can provide the user with information about news sources *in-situ* as well as aggregate and display their personal consumption data. In other words, we can convey to the user the bias of the current publisher they are reading and whether they tend to read content that leans right or left. To make this prototype ready for deployment, we are integrating Firebase authentication and packaging it for Google's Chrome Web Store.

4 PLANNED EVALUATIONS

Our next steps are to conduct a pilot usability study to passively log user browsing history. We plan to use this data to get a better understanding of their experiences with online news media (e.g., how many articles they read per day) and to inform future study designs (e.g., how often should we prompt users about accuracy), as well as what types of other interventions in their browsing sessions might be needed. This could include news recommendations derived from content that similar users find to be of high quality but that comes from sources not frequented by the target user. Alternatively, *would users accept having to answer questions about an article or scroll entirely through the content before engaging in further interactions (such as commenting or sharing), as*

[2]Media Bias Fact Checker: https://chrome.google.com/webstore/detail/official-media-bias-fact/hdcpibgmmcnpjmmenengjgkkfohahegk

proposed in [11]? To evaluate such interventions, we will conduct iterative deployments of future versions of our tool, including deploying our personal informatics dashboard. These latter studies will be complemented by pre-and post-study demographic and media literacy surveys, as well as additional experience sampling surveys embedded in our tool [10, 19, 27]. We will calculate descriptive statistics from our closed-form survey data and open-form responses will be thematically analyzed [6, 7]. Using this data, we will expand the capabilities of our plugin for collecting and displaying personal as well as media informatics data to users.

5 DISCUSSION, LIMITATIONS, & FUTURE WORK

The current work has given rise to several priorities for our future studies. Due to the immense amount of diverse content, it has become very challenging to identify, label, and automate the detection of fake news. Building off our data, our plan is to examine how machine learning and natural language processing (similar to [18, 29]) might help balance data analysis time and classification quality in fake news prediction to improve, speed up, and partially automate elements of the detection and intervention process. To accomplish this, we will tag and annotate one or more categories for each piece of misinformation in the article, such as fake, factual, satire, comedy, and so on. This would assist us in categorizing misinformation at a high level and enable us to explore a semi-automated pipeline for identifying misinformation. We also plan to look at how various recommendation algorithms (e.g., Matrix-Factorization [17], Bayesian Personalized Ranking [25], Hybrid-recommendation [9]) affect the spread of misinformation online as well as debiasing strategies like Inverse Propensity Scoring [26] to combat popularity bias and filter bubbles.

We also plan to explore engaging different classes of users with the project itself. To engage users online, our personal informatics and visualization dashboard is rendered in the user's internet browser and updates depending on the articles they read. This raises some interesting questions: *Do different visualizations affect user's behavior? If so, which visualization would be most suitable for our work? And, how does visualization literacy impact our results?* Other future expansions of our application include the modularization of our prompting interface, which might allow other researchers to use our tool as a platform for their studies with similar infrastructure requirements. Gamification may also be used to make the tool more engaging and help to develop a community of citizen fact-checkers (similar to or within citizen science communities, e.g.[20]).

In this workshop paper, we have discussed possible interventions that might help limit internet users' exposure to low-quality news content and misinformation. However, it is important to consider the ethical implications of our work and the need to avoid passing on the biases and values of our system designers to users. Additionally, while AllSides' media bias rating data is based on perspectives from people across the political bias spectrum, these media bias ratings are updated regularly and shift over time. As a result, future work will have to evaluate the best way to synchronize with these updates as we explore synergistic opportunities created by collecting additional data about and from internet users.

6 CONCLUSION

This project focuses on building a Chrome browser plugin to display online news media consumption data and intervene in browsing experiences to tackle the challenges online misinformation presents for users and, more generally, democracies around the world. Our goal is to assist users in becoming more thoughtful and critical of the information they consume, thus attempting to boost digital media literacy and empower users to make their own judgments about the content they consume. We expect to explore the usefulness of *in-situ* accuracy prompts and look at how they interact with partisan social identity to improve the intervention's efficacy. Interactions with online news media, as well as intervention efficacy over time, seem to be relatively understudied. Our work aims to address this gap by examining how prolonged exposure to prompts and interventions impacts learned behavior when seeking to make users more selective in terms of their online news consumption.

7 ACKNOWLEDGMENTS

We thank AllSides for their public data on news publisher bias ratings, which we used to bootstrap our preliminary work. Sourced from the Noun Project, we also would like to credit the following creators for the icons used in Figure 2, including Profile by The Icon Z, Browser by DinosoftLab, and Database by Eucalyp.

REFERENCES

[1] 2021. Balanced News via media bias ratings for an unbiased news perspective. https://www.allsides.com/unbiased-balanced-news
[2] Jae-wook Ahn, Krist Wongsuphasawat, and Peter Brusilovsky. 2011. Analyzing user behavior patterns in adaptive exploratory search systems with lifeflow. (2011).
[3] Michael Barthel, Amy Mitchell, and Jesse Holcomb. 2016. Many Americans believe fake news is sowing confusion. *Pew Research Center* 15 (2016), 12.
[4] Christina Boididou, Symeon Papadopoulos, Yiannis Kompatsiaris, Steve Schifferes, and Nic Newman. 2014. Challenges of computational verification in social multimedia. In *Proceedings of the 23rd International Conference on World Wide Web*. 743–748.
[5] Yazan Boshmaf, Ildar Muslukhov, Konstantin Beznosov, and Matei Ripeanu. 2011. The socialbot network: when bots socialize for fame and money. In *Proceedings of the 27th annual computer security applications conference*. 93–102.
[6] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology* 3, 2 (2006), 77–101.
[7] Virginia Braun and Victoria Clarke. 2012. Thematic analysis. (2012).
[8] Tom Buchanan. 2020. Why do people spread false information online? The effects of message and viewer characteristics on self-reported likelihood of sharing social media disinformation. *Plos one* 15, 10 (2020), e0239666.
[9] Robin Burke. 2002. Hybrid recommender systems: Survey and experiments. *User modeling and user-adapted interaction* 12, 4 (2002), 331–370.
[10] Mihaly Csikszentmihalyi and Reed Larson. 2014. Validity and reliability of the experience-sampling method. In *Flow and the foundations of positive psychology*. Springer, 35–54.
[11] Maria D. Molina, S Shyam Sundar, Md Main Uddin Rony, Naeemul Hassan, Thai Le, and Dongwon Lee. 2021. Does clickbait actually attract more clicks? Three clickbait studies you must read. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–19.
[12] Don Fallis. 2014. A functional analysis of disinformation. *IConference 2014 Proceedings* (2014).
[13] Miriam Fernandez and Harith Alani. 2018. Online misinformation: Challenges and future directions. In *Companion Proceedings of the The Web Conference 2018*. 595–602.
[14] Kiran Garimella, Tim Smith, Rebecca Weiss, and Robert West. 2021. Political polarization in online news consumption. *ICWSM* (2021), 152–162.
[15] Jennifer Golbeck, Matthew Mauriello, Brooke Auxier, Keval H Bhanushali, Christopher Bonk, Mohamed Amine Bouzaghrane, Cody Buntain, Riya Chanduka, Paul Cheakalos, Jennine B Everett, et al. 2018. Fake news vs satire: A dataset and analysis. In *Proceedings of the 10th ACM Conference on Web Science*. 17–21.
[16] Paul T Jaeger and Natalie Greene Taylor. 2021. Arsenals of lifelong information literacy: Educating users to navigate political and current events information in world of ever-evolving misinformation. *The Library Quarterly* 91, 1 (2021), 19–31.
[17] Yehuda Koren, Robert Bell, and Chris Volinsky. 2009. Matrix factorization techniques for recommender systems. *Computer* 42, 8 (2009), 30–37.
[18] Chun-Ming Lai, Mei-Hua Chen, Endah Kristiani, Vinod Kumar Verma, and Chao-Tung Yang. 2022. Fake News Classification Based on Content Level Features. *Applied Sciences* 12, 3 (2022), 1116.
[19] Reed Larson and Mihaly Csikszentmihalyi. 2014. The experience sampling method. In *Flow and the foundations of positive psychology*. Springer, 21–34.
[20] Roman Lukyanenko, Andrea Wiggins, and Holly K Rosser. 2020. Citizen science: An information quality research frontier. *Information Systems Frontiers* 22, 4 (2020), 961–983.
[21] Sebastião Miranda, David Nogueira, Afonso Mendes, Andreas Vlachos, Andrew Secker, Rebecca Garrett, Jeff Mitchel, and Zita Marinho. 2019. Automated fact checking in the news room. In *The World Wide Web Conference*. 3579–3583.
[22] Gordon Pennycook and David Rand. 2021. Nudging social media sharing towards accuracy. (2021).
[23] Gordon Pennycook and David G Rand. 2019. Fighting misinformation on social media using crowdsourced judgments of news source quality. *Proceedings of the National Academy of Sciences* 116, 7 (2019), 2521–2526.
[24] Kashyap Popat, Subhabrata Mukherjee, Jannik Strötgen, and Gerhard Weikum. 2018. CredEye: A credibility lens for analyzing and explaining misinformation. In *Companion Proceedings of the The Web Conference 2018*. 155–158.
[25] Steffen Rendle, Christoph Freudenthaler, Zeno Gantner, and Lars Schmidt-Thieme. 2012. BPR: Bayesian personalized ranking from implicit feedback. *arXiv preprint arXiv:1205.2618* (2012).

[26] Tobias Schnabel, Adith Swaminathan, Ashudeep Singh, Navin Chandak, and Thorsten Joachims. 2016. Recommendations as treatments: Debiasing learning and evaluation. In *international conference on machine learning*. PMLR, 1670–1679.
[27] Saul Shiffman, Arthur A Stone, and Michael R Hufford. 2008. Ecological momentary assessment. *Annu. Rev. Clin. Psychol.* 4 (2008), 1–32.
[28] Marianna Spring. 2020. Coronavirus: The human cost of virus misinformation. *BBC News* 27 (2020).
[29] Aswini Thota, Priyanka Tilak, Simrat Ahluwalia, and Nibrat Lohia. 2018. Fake news detection: a deep learning approach. *SMU Data Science Review* 1, 3 (2018), 10.
[30] Christian von der Weth and Manfred Hauswirth. 2013. Dobbs: Towards a comprehensive dataset to study the browsing behavior of online users. In *2013 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT)*, Vol. 1. IEEE, 51–56.