

Empowering Users on Social Media for Better Content Credibility

Farnaz Jahanbakhsh

Problem:

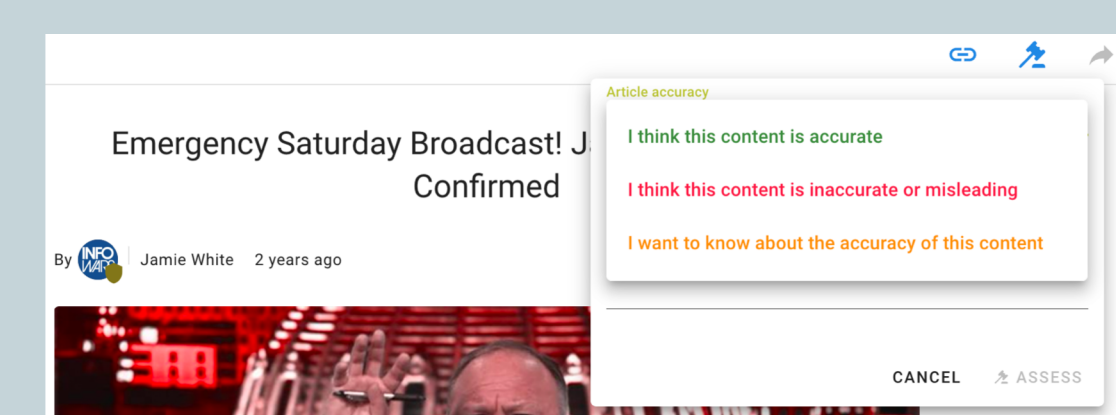
Social media platforms are the authority on what content is considered misinforming and what should be done with such content, e.g., whether to flag, downrank, or remove it. This centralized governance:

- 1 Treats users as passive consumers, without agency.
- 2 Assumes that users trust the platforms' competence and goodwill.
- 3 Does not address the needs of every user. Some users e.g., want to see the misinformation that their friends post to know how to talk to them.

Solution:

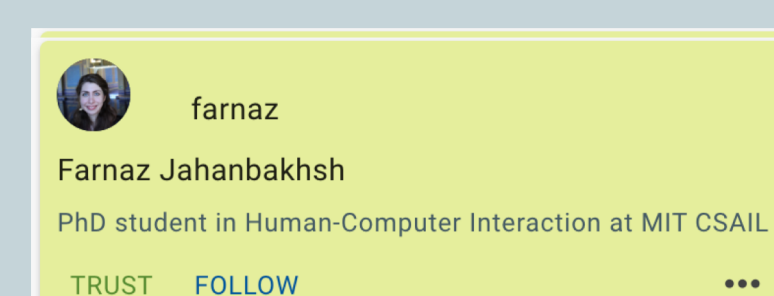
Democratizing misinformation moderation, by providing system affordances to empower users.

Trustnet: a social media platform that enables users to

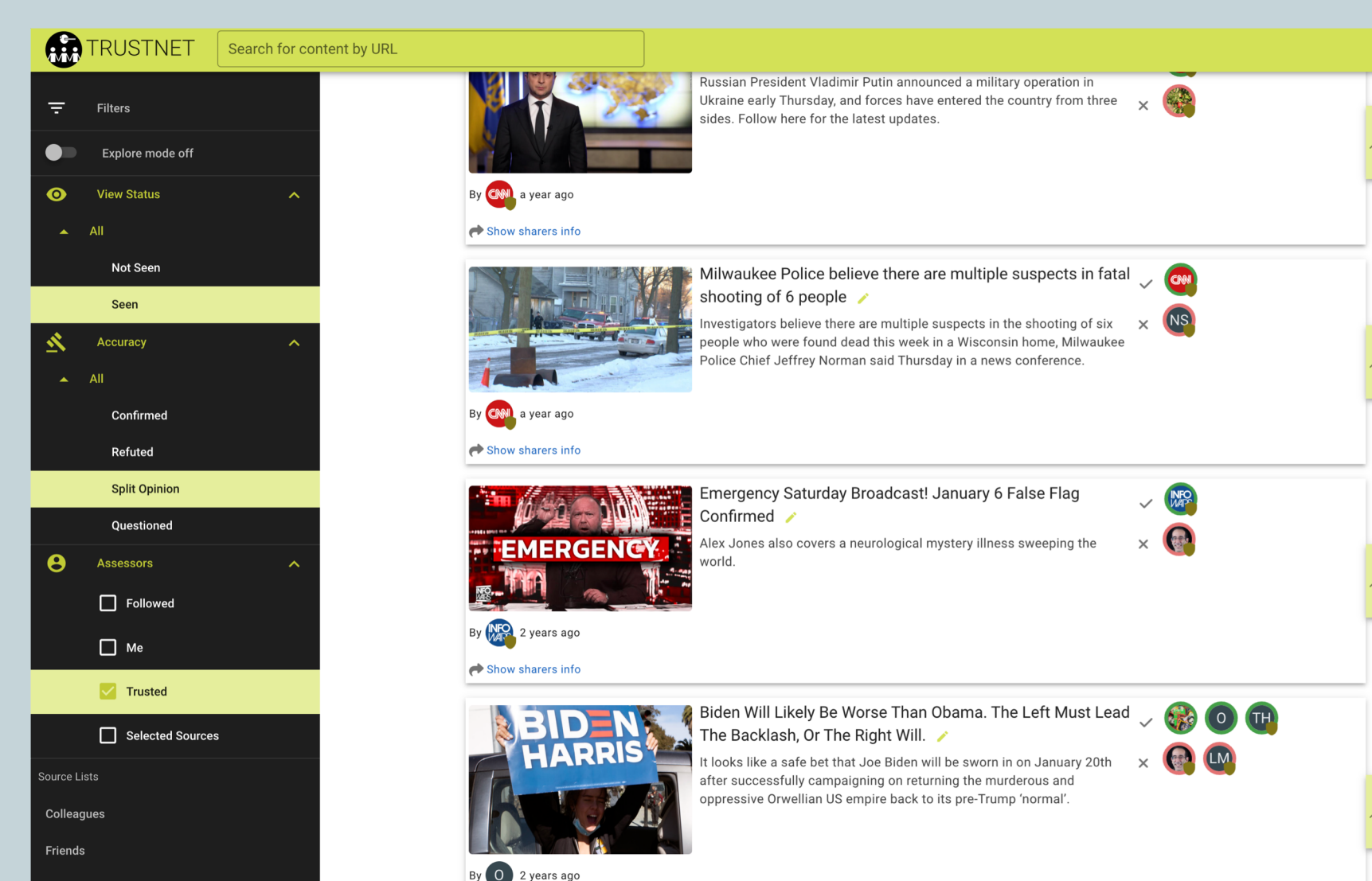


- 1 Assess content or ask about its accuracy. The system captures the assessments and questions as part of the data model.

- 2 Specify their trusted sources, separate from who they want to follow.



- 3 Filter the content on their feed based on the assessments of their trusted sources.



Offering these user affordances on all the platforms on the web

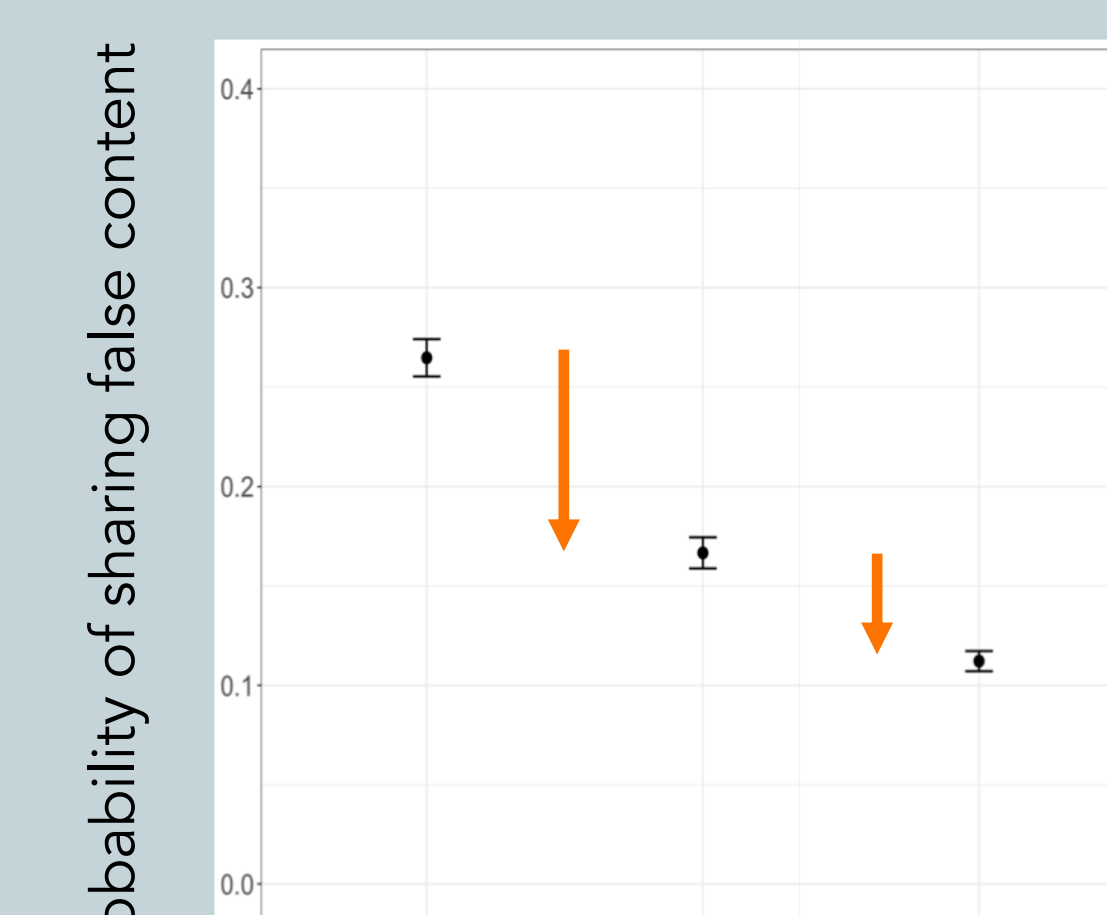


The Trustnet browser extension enables users to assess the accuracy of any content on the web and shows the user assessments from others that they trust in-situ on the page.

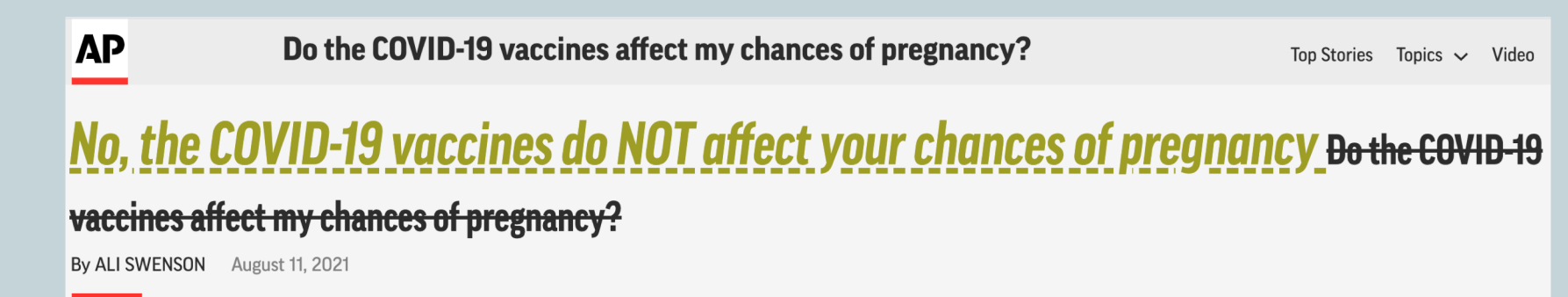
It also places any available trusted accuracy assessments next to outgoing links found on the page and fades inaccurate links.



Can asking users to assess content accuracy reduce the likelihood that they share misinformation?



Enabling users to modify misinforming content



The Reheadline browser extension enables users to suggest alternative headlines for news articles whose headlines they find misleading, clickbait, or otherwise problematic.

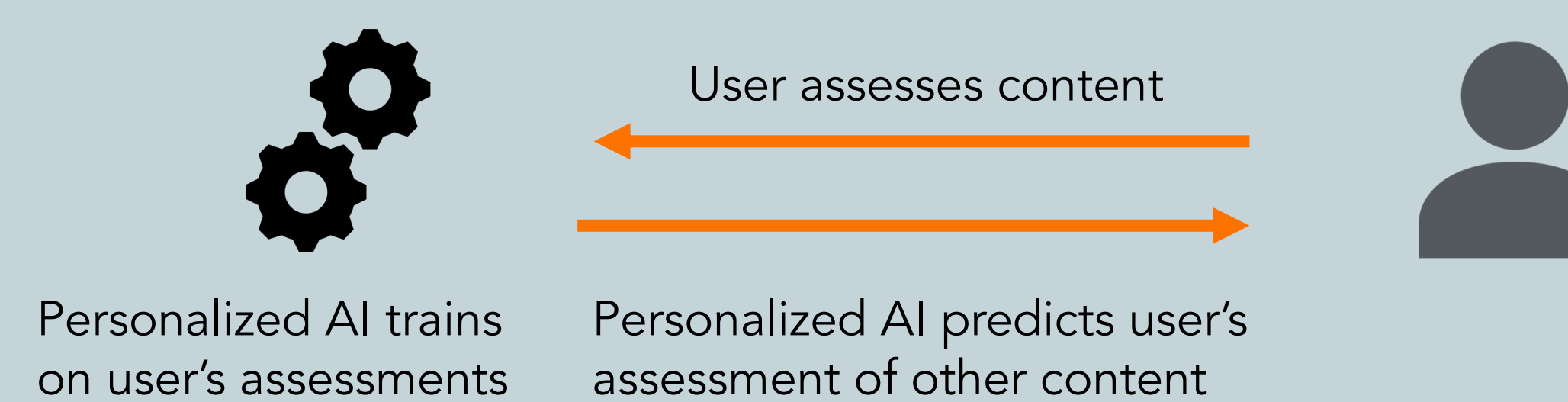
Other users who follow the headline suggester will see the alternative headline next to the original wherever the original headline appears on the web.



Scaling a user's assessments

With so many different sources publishing similar claims and stories, not showing a user's assessment of a piece of content on other similar content would be a missed opportunity.

Solution:



Leveraging Structured Trusted-Peer Assessments to Combat Misinformation. Jahanbakhsh et al. CSCW'22
 Our Browser Extension Lets Readers Change the Headlines on News Articles, and You Won't Believe What They Did!. Jahanbakhsh et al. CSCW'21
 Exploring lightweight interventions at posting time to reduce the sharing of misinformation on social media. Jahanbakhsh et al. CSCW'21
 Exploring the User of Personalized AI for Identifying Misinformation on Social Media. Jahanbakhsh et al. to appear in CHI'23

Sincere thanks are due to my advisor David Karger, and my collaborators Amy X. Zhang, Adam Berinsky, David G. Rand, Gordon Pennycook, Karrie Karahalios, Yannis Katsis, Dakuo Wang, Lucian Popa, and Michael Muller