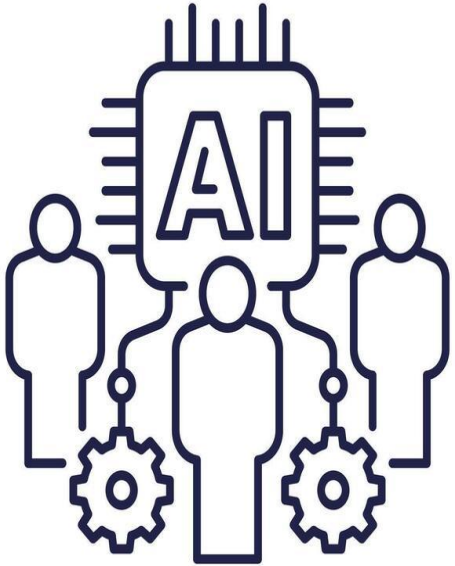
Meeting Online Threats: Supporting Critical Engagement and Well-being Online



Nino Gugushvili
Faculty of Psychology and Neuroscience

PSYCHOLOGY OF HUMAN-AI INTERACTION

Sub-line 1: Transparency,Explainability, and Controllability of AI

- Does User Control over Algorithmic Social Media Feeds Reduce Polarization?
- Does Explainability Help Users Calibrate Their Trust in AI Systems?
- Does digital literacy intervention on social media reduce misinformation susceptibility?

Sub-line 2: Affective use of AI and wellbeing

Disinformation & Democracy sector plan

Critical Engagement and Information Diversity Online

RESEARCH ARTICLE | PSYCHOLOGICAL AND COGNITIVE SCIENCES |

The echo chamber effect on social media

Matteo Cinelli, Gianmarco De Francisci Morales, Alessandro Galeazzi, +1, and Michele Starnini Authors Info & Affiliations

Edited by Arild Underdal, University of Oslo, Oslo, Norway, and approved January 14, 2021 (received for review November 15, 2020)

ELSEVIER

Body Image
Volume 41, June 2022, Pages 292-297

The dangers of the rabbit hole: Reflections on social media as a portal into a distorted world of edited bodies and eating disorder risk and the role of algorithms

Jennifer A. Harriger [a], Joshua A. Evans [a], J. Kevin Thompson [b], Tracy L. Tylka [c]

What Tiktok Personalization Reveals About Human-AI Interaction: Impact on Individual and Collective Dynamics

TIKTOK NEWS POLICY

TikTok's algorithm will be optional in Europe

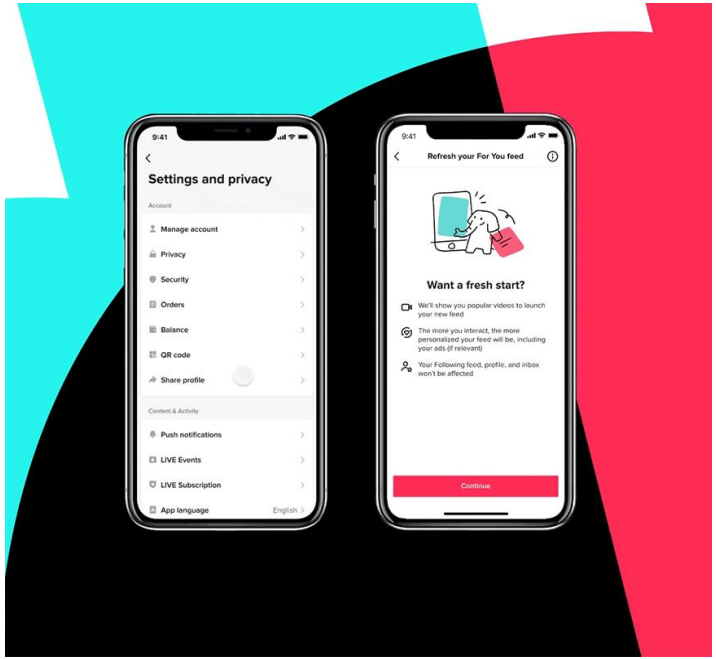
Illustration by Nick Barclay / The Verge

/ TikTok users in Europe will be able to see recommended 'For You' videos that don't rely on tracking their online activity.

by Jess Weatherbed
Aug 4, 2023, 2:45 PM GMT+2

7 Comments (All New)

- **RQ:** How does TikTok's curation algorithm affect political polarization?
- **Method**: EU Field Experiment, N=612, 2 weeks
    - Turning off personalization → No effects on political polarization
- **Other outcomes**: Mental health, addiction, perceived responsiveness.
    - Reduced problematic TikTok use; No effect on Mental health



Masciantonio, Alexandra and Gugushvili, Nino, What Tiktok Personalization Reveals About Human-Ai Interaction: Impact on Individual and Collective Dynamics.

Does digital literacy intervention on social media reduce misinformation susceptibility?

- In collaboration with **IMT Luca (Italy)**
- **RQ:** How does lateral reading training (digital literacy intervention) impact (1) truth discernment; (2) trust in media?

- **Design:**
  - Between-subjects: (C1) Lateral reading + booster (C2) Lateral reading only; (C3) Control
  - Within-subjects: T1 (baseline); T2 (in two weeks);
  - Sample: N=1,500 (Italian, Dutch)

- **Key outcomes:** detection of manipulative and non-manipulative content, and overall discernment; trust in news media;

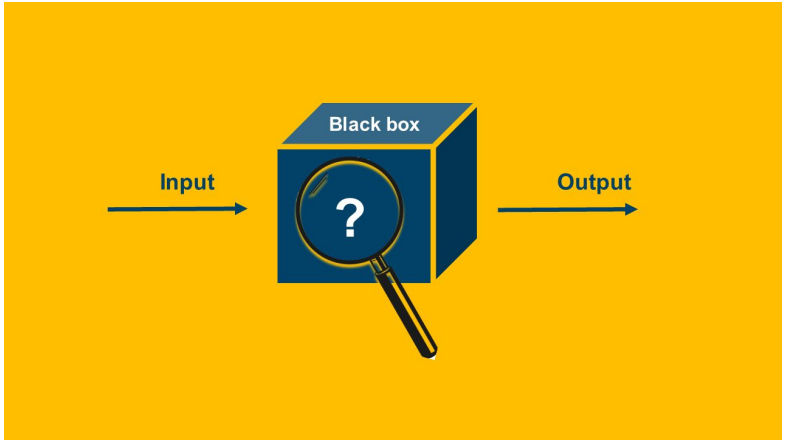
Universiteitsfonds Limburg | SWOL |
Supported by University Fund Limburg | SWOL

From Black Box to Transparency: Comparing the Impact of XAI on Trust and Adoption in Europe, North America, and Oceania

Explainable AI (XAI): An AI system that shows in a clear way how and why it made a decision.

Trust calibration: ensuring users do not over-trust or under-trust AI systems.

International collaboration among seven researchers across four regions (New Zealand, Netherlands, Indonesia, Canada).



WUN WORLDWIDE UNIVERSITIES NETWORK

Supported by The Worldwide Universities Network (WUN) Research Development Fund

- **Aim:** Examine whether Explainable AI (**XAI**) can help ordinary users develop more accurate, critical, and balanced trust in AI, especially in high-stakes situations

- **Study:** baseline → AI trust, understanding, and usage. Manipulation → Control (Standard GPT) versus Experimental group (XAI GPT). Post → Both groups complete identical high-stakes vignette tasks across health, education, and mental health

- **Tasks/vignettes:** Health → e.g., asking AI to explain a medical test result); Education → e.g., seeking guidance on choosing a study program; Mental health → asking for advice about emotion regulation

Maastricht University

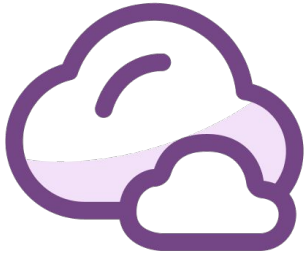
UNIVERSITY OF ALBERTA

UNIVERSITY OF AUCKLAND
Waipapa Taumata Rau
NEW ZEALAND

UNIVERSITAS PADJADJARAN

What is your profession, and what brings you here today?

What do you see as the most promising way to address these online threats?



How can universities foster students' development as critical, responsible, and engaged digital citizens?

THANK YOU!