ICT20-016 - Young People Against Online Hate: Computer-assisted Strategies for Facilitating Citizen-generated Counter Speech

Zusammenfassung

Hintergrund

Der Umgang mit Online-Hass, Cybermobbing und digitaler Gewalt ist ein massives gesellschaftliches Problem. Jugendliche sind durch ihre intensive Internetnutzung besonders stark betroffen. Sie begegnen solchen Übergriffen meist mit Ignorieren oder Rückzug, hätten aber hohes Interventionspotenzial, vor allem mittels gezielter Förderung von Counter Speech (aktiver Gegenrede). Diese Form der digitalen Zivilcourage scheitert jedoch häufig an fehlenden Kompetenzen und wird durch plattformspezifische Besonderheiten erschwert. In diesem Projekt entwickelten daher Soziolog*innen und Informatiker*innen gemeinsam mit Jugendlichen computergestützte Strategien und Tools, um jugendliche Counter Speech zu unterstützen und die Umsetzbarkeit von digitaler Zivilcourage zu fördern.

Methoden

Folgende Methoden wurden verwendet: (i) automatisierte, computergestützte Datensammlung im Rahmen von Fallstudien, um typische Erscheinungsformen von Online-Hass und Gegenrede auf Plattformen wie TikTok zu erfassen; (ii) quantitative und qualitative Inhaltsanalysen sowie maschinelle Lernverfahren, um spezifische Muster zu identifizieren und diese zur computergestützten Erkennung jugendlicher Gegenrede sowie zur Entwicklung praktikabler, jugendgerechter Formulierungen zu nutzen; (iii) partizipative Forschungsmethoden, um gemeinsam mit Jugendlichen Gegenrede mit Hilfe digitaler Tools so zu gestalten, dass sie auch angenommen und praktiziert wird .

Wichtigste Ergebnisse

Die Studie zeigt deutlich die Herausforderungen im Umgang mit digitaler Gewalt sowie plattformspezifische Besonderheiten. Die Praktikabilität von Counter Speech wird massiv erschwert, da sich Online-Hass sowohl auf Text - als auch auf Bildebene subtil äußert und sich ständig verändert. Counter Speech ist dadurch sehr schwer identifizierbar. Gleichzeitig sind Jugendliche je nach Plattformalgorithmus, Nutzungsverhalten und Community -Interaktion in sehr unterschiedlichem Ausmaß mit Hass konfrontiert. Besonders jene, die Zivilcourage zeigen wollen, erleben in ihrer eigenen „Bubble“ oft nur wenig Hass. Wird gezielt nach negativen Inhalten gesucht, um zivilcouragiert darauf reagieren zu können, werden sie algorithmusbedingt unerwünschterweise vermehrt mit negativen Inhalten konfrontiert. Fehlende Online-Zivilcourage unter Jugendlichen kann daher nicht unbedingt mit mangelnder Empathie begründet werden, empathische Jugendliche haben oft kaum Gelegenheit einzugreifen oder sind, wenn sie es tun, noch stärker mit Hass konfrontiert. Gleichzeitig ist mangelnde Online-Zivilcourage durch Einschränkungen digitaler Räume erklärbar: Jugendliche scheitern mit ihrer Gegenrede oft an mangelnder Sichtbarkeit, Reichweite und Wirkung , was zu einem Gefühl der Machtlosigkeit führt. In der Studie wurde ein unterstützendes und spielerisches Counter-Speech -Tool entwickelt, das es ermöglicht, Hass zu markieren, Reaktionen anzuregen und Gegenrede vorzuschlagen, die von Jugendlichen als attraktiv und motivierend wahrgenommen wird.

Schlussfolgerungen

Die Ergebnisse zeigen: Die Hauptursache für mangelnde Online-Zivilcourage und die geringe Wirksamkeit von Counter Speech sind plattformspezifische Algorithmen, die polarisierende Inhalte bevorzugen, die Sichtbarkeit von Counter Speech einschränken und konstruktives Engagement behindern. Zentral wäre, dass Plattform-Betreiber*innen digitale

Tools zur Förderung von Online-Zivilcourage auch implementieren.

Weitere Infos zu den Ergebnissen in diesem Bericht der FH St. Pölten.

Wissenschaftliche Disziplinen:
Sociology of youth (50%) | Machine learning (50%)

Schlagworte:
Hate speech, counter speech, online civil courage, information retrieval, multimodal analysis

---

Principal Investigator:	Ulrike Zartler
Institution:	University of Vienna
Co-Principal Investigator(s):	Matthias Zeppelzauer (University of Applied Sciences St. Pölten) Christiane Atzmüller (University of Vienna)



v.l.n.r. Ulrike Zartler ©citronenrot, Matthias Zeppelzauer ©FHSanktPölten, Christiane Atzmüller ©Gilbert Handler

---

Status: Abgeschlossen (15.09.2021 – 14.09.2025)
GrantID: 10.47379/ICT20016

Weiterführende Links zu den beteiligten Personen und zum Projekt finden Sie unter
https://wwtf.at/funding/programmes/ict/ICT20-016/