ICT20-015 - Transparent Automated Content Moderation (TACo)

Zusammenfassung

Politische Diskussionen im Internet sind zunehmend durch Negativität und Konflikt gekennzeichnet – mit negativen Konsequenzen für BürgerInnen, Politik und Gesellschaft. Diese “toxischen” Inhalte werden darum auf vielen Plattformen durch (automatisierte) Content Moderationsverfahren eingedämmt. Diese Verfahren vernachlässigen jedoch fast immer die Perspektive der NutzerInnen selbst, und fällen so scheinbar arbiträre Entscheidung darüber, welche Beiträge “toxisch” sind. Auch gibt es viele Moderationsverfahren, die auf simplen und verzerrten algorithmischen Modellen basieren. Dies kann sich rufschädigend für Organisationen auswirken, die solche intransparente Moderationsverfahren anwenden. Viel wichtiger ist jedoch, dass diese den BürgerInnen die Handlungsmacht entziehen, und das obwohl sich viele in digitalen Informationsumgebungen bereits machtlos fühlen. Das TACo-Projekt wendet darum einen nutzerzentrierten Ansatz an, durch den politische Beiträge im Netz effektiv und transparent moderiert werden sollen. In einer qualitativen, explorativen Studie wird zunächst erörtert, was NutzerInnen unter toxischer Sprache verstehen, und welche Moderationsverfahren sie vorziehen. Ausgehend von diesem Wissen werden neue automatisierte Modelle zur Bestimmung toxischer Inhalte entwickelt. Schließlich wird mit experimentellen Befragungen untersucht, ob das, was NutzerInnen „wollen“, auch wirklich positive Effekte hat, z.B. auf politisches Vertrauen, politische Beteiligung und Wohlbefinden.

Wissenschaftliche Disziplinen:
Political communication (50%) | Data science (50%)

Keywords:
content moderators, user comments, incivility, toxic talk, citizen engagement

---

Principal Investigator:	Sophie Lecheler
Institution:	University of Vienna
Co-Principal Investigator(s):	Allan Hanbury (TU Wien)

v.l.n.r. Sophie Lecheler, Allan Hanbury ©Foto Schoerg

Status: Abgeschlossen (01.09.2021 - 30.04.2026)
GrantID: 10.47379/ICT20015

Weiterführende Links zu den beteiligten Personen und zum Projekt finden Sie unter
https://www.wwtf.at/funding/programmes/ict/ICT20-015/