

Anna-Clara Pentz: KI für den digitalen Jugendschutz

Beitrag aus Heft »2022/02 Sprache in den Medien – Deutungshoheit und Sprachschlachten«

In ihrem Alltag mit Social-Media-Plattformen, Chat-Portalen und Messenger-Apps sind Kinder und Jugendliche neben den vielen Möglichkeiten zur Kommunikation, Selbstdarstellung und Identitätsbildung vermehrt auch mit Gefahren und Problemen konfrontiert. Gut, wenn hier Technologien beispielsweise vor Sexting und Cybergrooming (digitale Kontaktaufnahme Erwachsener mit Minderjährigen, mit dem Ziel, diese zu missbrauchen) schützen können. Das Fraunhofer-Institut für Sichere Informationstechnologie (SIT) hat technische Möglichkeiten geprüft, die Minderjährige vor dem unachtsamen Verschicken von Nacktfotos (sogenanntes ‚Sexting‘) warnen könnten. Außerdem wurden Verfahren untersucht, mit denen etwa erwachsene Personen erkannt werden können, die sich in Online-Foren oder Chat-Portalen als Minderjährige ausgeben.

Die Studie ‚Technik für den digitalen Jugendschutz: Automatische Erkennung von Sexting und Cybergrooming‘ stellt unterschiedliche technische Ansätze vor, mit denen sich der Jugendschutz in der digitalen Welt verbessern lässt. So werden etwa Verfahren zur ‚Hauterkennung‘ in den Blick genommen, durch die mithilfe automatischer Bilderkennung verhindert werden könnte, dass Apps auf intime Bilder zugreifen und Minderjährige Nacktbilder versenden. Technische Verfahren aus der Textforensik, wie etwa Verfahren zur Chat-Analyse oder zur Identifizierung von Personen anhand ihres Schreibstils, können dabei helfen, potenzielle Täter*innen rechtzeitig zu erkennen. Die Studie zeigt auf, dass eine deutliche Verbesserung maschineller Lernverfahren zu beobachten ist, Künstliche Intelligenz aber hier auch an ihre Grenzen stößt.

www.sit.fraunhofer.de