

Liebe Mitglieder des Netzwerkes der ehemaligen SSAB, liebe Interessierte

Deep Fakes – mit KI gefälschte Audio-, Video- und Bilddateien – haben beachtliches Manipulationspotenzial. Betroffen sind damit die Grundlagen demokratischer Meinungsbildung, insbesondere bei Wahlen. Kritisches Denken und Hinterfragen von Informationen werden immer wichtiger. Die Bildungsinstitutionen aller Stufen sind gefordert. Was, wenn Deep Fakes nicht als solche erkannt werden? Was, wenn mit der Verunsicherung darüber, was fake, was echt ist, generell das Vertrauen in politische Prozesse und letztlich auch in den Staat erodiert? Hierzu ein paar Beispiele.

Debatten mit KI-Modellen – überzeugender als Diskussionen mit Menschen?

Eine am [Data Science Lab](#) der EPFL erstellte Studie zur Überzeugungskraft von Menschen im Vergleich zu KI-Modellen kam zu einem überraschenden Ergebnis, [hier](#). Von den insgesamt 850 Teilnehmenden änderten jene, die mit GPT-4 debattierten, ihre Meinung viel eher als jene, die mit Menschen diskutierten; dies umso deutlicher, mit rund 82 Prozent höherer Wahrscheinlichkeit, wenn die KI auf die persönlichen Daten der Teilnehmenden zugreifen und massgeschneidert argumentieren konnte. Gemäss Professor Robert West, Leiter des Labors, ist zu erwarten, dass solche Sprachmodelle in Wahlkämpfen eingesetzt werden.

Als KI-generiertes Video gekennzeichnet – trotzdem grosse Wirkung

Ein gut gemachtes Fake Video ([hier](#)) liess die dänische Premierministerin Mette Frederiksen an einer Pressekonferenz erklären, alle christlichen Feiertage würden abgeschafft; künftig gebe es nur noch einen Feiertag, das muslimische Fest des Fastenbrechens. Das Video enthielt in kleinem Logo den Hinweis «KI-generiert» und der Autor erschien am Schluss im Bild. Es war sein Wunschtraum – in Form einer Satire. Trotzdem scheint es, dass zahlreiche Menschen auf das Fake Video hereinfielen, was eine hitzige Debatte auslöste.

KI in den Wahlen in Indien vom 19. April bis 1. Juni 2024

Gemäss verschiedenen Medienberichten (z.B. [Reuters](#), [swissinfo](#)) zeige sich bereits jetzt, dass in der weltweit grössten Demokratie (rund 970 Millionen Inderinnen und Inder, mehr als 800 Millionen Internetbenutzende) die Bekämpfung der Verbreitung von Deep Fakes sehr schwierig sei. Regierung und Opposition werfen sich gegenseitig Manipulationen in den Social Media vor. Zudem mischen eigentliche Fälscherbanden mit. Das Internet wird intensiv überwacht, Inhalte werden gelöscht, Verhaftungen vorgenommen, aber was hier vorgeht, lässt sich nur schwer überblicken. Deep-Fakes haben meist bereits Wirkung erzielt, auch wenn sie rasch aufgedeckt und von den Plattformen entfernen werden. Interessant dürften Auswertungen im Nachgang sein.

Es lässt sich etwas gegen Fake News machen – so im Swing-State Arizona

Ein kleines Radiostudio in Phoenix, Arizona, widerlegt laufend Fake News, die für den bevorstehenden US-Wahlkampf von Bedeutung sind. Das Radio Campesina wendet sich in spanischer Sprache an die Latinos, die in Arizona rund einen Drittel der Bevölkerung ausmachen. Die Latinos beziehen ihre Informationen vor allem aus den mit Deep Fakes überfluteten Social Media. Die amerikanische Nachrichtenagentur AP (Associated Press) mit Sitz in New York berichtete, [hier](#).

Abkommen gegen Wahlmanipulation durch KI

In der Pflicht wären aber eigentlich die Tech-Giganten. Bereits im Februar 2024, anlässlich der Münchner Sicherheitskonferenz, unterzeichneten mehrere Tech-Giganten, darunter Google, Meta, Microsoft, OpenAI und TikTok, ein Abkommen gegen manipulative Desinformation bei Wahlen, den [AI Elections Accord](#). Im Fokus waren vor allem Deep Fakes. Absichtserklärungen der Tech-Giganten zu KI-Sicherheitsfragen gibt es mehrere. Eben erst, am AI-Summit vom 21./22.5.2024 in Seoul (Südkorea) haben 14 KI-Unternehmen wiederum ein entsprechendes Dokument unterzeichnet, [hier](#).

Aber wo bleiben die Wirkungen dieser Abkommen? Wer hat welche Verantwortung zu übernehmen – die Tech-Giganten, die Staatengemeinschaft, die Zivilgesellschaft, die Bildungsinstitutionen?

Mit freundlichen Grüßen

Für das Netzwerk der ehemaligen SSAB: Hanna Muralt Müller

Neues Datenschutzrecht: Falls Sie diese E-Mail nicht mehr erhalten möchten, melden Sie sich bitte bei mir!