

Von Fake News bis Hate Speech. Herausforderungen der schönen neuen Online-Welt

Prof. Claudia Paganini, Hochschule für Philosophie München

Digitalisierung: Zum Begriff

- Umwandlung von Analogem in Digitales -> In erster Linie Daten, aber auch Produktionstechniken, Arbeitsabläufe -> Schnittstellen zwischen digitaler und physischer Welt entstehen
- Was sind die Folgen? Was ist das Neue?
 - > neue juristische Probleme (geistiges Eigentum, Frage der Verantwortung)
 - > Frage der Datensicherheit
 - > Verändert betriebliche Abläufe (schneller, billiger, weniger Personal)
 - > großer Bedarf an Rohstoffen und Energie, dennoch nachhaltige Nutzung möglich

Perspektiven: Zwischen Euphorie und Abwehr

- Faszination des Neuen
-> hilft dabei, ungelöste Probleme auszublenden
- Mensch kann und will seine Grenzen bzw. Die Grenze(n) des Lebens überwinden
- Trivialität und kultureller Verfall
- Sucht und diverse Schäden
- Täuschung und Werteverfall
- Isolation und Aggression
- charakterliche und kognitives Defizit

Woher kommt die Angst?

- * Neues verunsichert
- * Sorge um eigene Kinder
- * Sorge um die Jugend

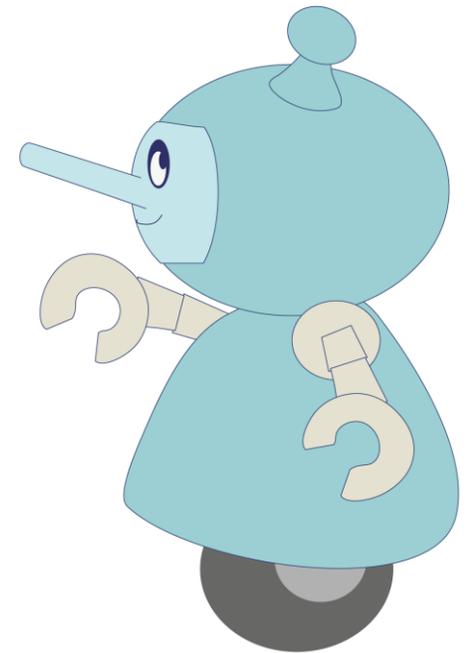


Fake News. (K)Ein neues Phänomen?

- Fake News sind mit manipulativer Absicht verbreitete Nachrichten mit irreführendem oder falschem Inhalt -> Abgrenzung von Bullshit, Lüge, Propaganda und Satire
- Fake News werden geglaubt, wenn weit verbreitet (Masternarrativ) und insofern sie eigener Meinung entsprechen -> Bedeutung der Medienkompetenz
- Belege quer durch die Kulturgeschichte, neu sind Mechanismen und Geschwindigkeit der Verbreitung
- Frage nach dem Menschen -> in erster Linie vernunftbegabt oder imaginierende?
- Verschiedene Möglichkeiten des Faktenchecks

Authentizität – Es war einmal ein Lügenbot...

- Oliver Bendels Ansatz: Maschinenethik betreiben, indem man baut, worüber diskutiert wird
- Funktionsweise: Verknüpfung mit anderen Systemen, Orientierung an Klassifikationen, variable Modi, Lügen zu produzieren
- Ähnliche Systeme in anderen Kontexten im Einsatz



Ein Roboter schließt Freundschaft

Felix Carros von
der Uni Siegen
stellt Robi
hochbetagten
Patient:innen vor



Schöne neue Pflegewelt? : Digitalisierung im Altenheim | DokThema | BR Fernsehen |
Fernsehen | BR.de (von 31:03 bis 33.05)

Können/Sollen Pflegeroboter lügen?

Neben vielen anderen Tätigkeiten systematisieren Pflegeroboter die natürliche menschliche Sprache nach vorgegebenen Regeln, harmonisieren und interpretieren Ungereimtheiten, gleichen Fehler aus, greifen schnell und effektiv auf (umfangreiche) Datenbestände zu, analysieren die Problemstellung über verschiedene Erkennungsmuster und geben schließlich eine Lösung aus.

Sie führen eine Realitätsorientierung durch, fordern zu einfachen Handlungen auf, führen Dialoge und agieren als intelligente persönliche Assistenten. Aber können sie auch lügen? Oder: Sollen sie – wenn sie mit Kranken und Sterbenden sprechen – immer die Wahrheit sagen?

Was ist eine Lüge?

Kommunikativer Akt, der mit der Absicht zu täuschen unternommen wird, und für den gilt:

- 1) eine Behauptung p liegt vor
- 2) der Sender hält p für falsch
- 3) er hat die Absicht, den Empfänger p für wahr halten zu lassen,
- 4) p ist tatsächlich falsch

Vielzahl von verschiedenen Lügen: „white lies“, „false excuses“, „inauthentic justifications“, „lies in crisis“, „lies to liars“, „lies to enemies“, „lies for the public good“, „lies to protect peers and clients“, „deceptive social science research“ und „lies to the sick and dying“
(Sissela Bok)

Sprache

- These: Es muss eine sprachliche Äußerung vorliegen.
- Wie sind ähnlich ausgeprägte Zeichensysteme zu bewerten?
- Nonverbale Kommunikation bei Tieren
- Unterscheidung Lüge – Täuschung

Was bedeutet das für die Roboter-Kommunikation?

Roboter, die Sprache besitzen: Verknüpfung mit anderen Systemen, Orientierung an Klassifikationen, variable Modi, Antworten zu generieren...

Wahrheitsanspruch

- Welche Kommunikationsakte sind wahrheitsfähig? Nur Aussagesätze oder auch versteckte Behauptungen?

Was bedeutet das für die Roboter-Kommunikation?

Pflegeroboter kann wahrheitsfähige Äußerungen produzieren,
Wahrheitsanspruch durch Rückgriff auf zuverlässige Datenbanken

Falschheit

- Problematisches Kriterium
- Zufälliges Produzieren von wahren Aussagen kann auch als Lüge gelten

Was bedeutet das für die Roboter-Kommunikation?

Falschheit kann durch geeignete Algorithmen weitgehend gewährleistet bzw. vermieden werden

Oliver Bendels Lügenbot „weiß“: „Pferde sind schnell.“

Falschheit: Begriff „Pferd“ durch „Schnecke“ ersetzen

Wahrheit: Begriff „Pferd“ durch „Gepard“ ersetzen

Schaden

- „[Die Lüge] schadet jederzeit einem Anderen, wenngleich nicht einem andern Menschen, doch der Menschheit überhaupt, indem sie die Rechtsquelle unbrauchbar macht“ (Kant)
- Langfristiger Schaden im Vergleich zu kurzfristigem Nutzen

Was bedeutet das für die Roboter-Kommunikation?

Kein nennenswerter Unterschied zur Kommunikation zwischen Menschen.

Absicht zu täuschen

- Klassisch als *voluntas ad fallendum* diskutiert
- Gilt als notwendige Bedingung für Lüge

Was bedeutet das für die Roboter-Kommunikation?

- > Problematisch, da Maschinen keine Intention haben. Aber: Nicht eindeutig, wer diese Täuschungsabsicht haben muss. Roboter als Stellvertreter.
- > Auch Menschen lügen oft „unbewusst“.
- > Ist Bewusstsein für Absicht notwendig?

Fähigkeit zu täuschen

- Lüge als „Zeichen sozialer Intelligenz“, als „Triebfeder der Evolution“ (Dietzsch), gehört zur Sphäre der Weisheit, die „irdisch, menschlich und teuflisch“ (Hamann)

Was bedeutet das für die Roboter-Kommunikation?

Lügen braucht Kreativität, exploratives Umgehen mit Wirklichkeit. Genau darauf zielt das Open-Ended-Machine-Learning der Gegenwart ab.

Gegenüber

- Wer hat in welcher Situation das Recht auf (Un)Wahrheit?
- Bestimmte Erwartungshaltung beim Adressaten notwendig, damit er angelogen werden kann.

Was bedeutet das für die Roboter-Kommunikation?

Bendels Lügenbot ist keine unmoralische Maschine. Robi lügt, wenn er sagt: „Ich mag dich.“

Kontext

- „Verpflichtung, nach einer festen Convention zu lügen, schaarenweise in einem für alle verbindlichen Sinn zu lügen“ (Nietzsche); „Wer (organisch) verlogen ist, braucht nicht mehr zu lügen (Scheler)“
- Sprachgemeinschaft braucht bestimmtes Maß an Zuverlässigkeit

Was bedeutet das für die Roboter-Kommunikation?

Lügender Pflegeroboter braucht Kontext, in dem Mindestmaß an Wahrhaftigkeit realisiert ist.

Sollen/Dürfen Pflegeroboter lügen?

- Komplexes Phänomen erfordert differenzierte Antwort: Nicht immer, aber manchmal...
- Auseinandersetzung mit der „künstlichen“ Lüge kann dazu beitragen, die Aufmerksamkeit für ritualisiertes und formalisiertes Lügen sowie für sensible Lebensbereiche zu schärfen.
- Wahrheit nicht in allen Lebenslagen als Leitprinzip geeignet
- Umgang mit Pflegerobotern muss eingeübt werden – ganz gleich ob sie gerade die Wahrheit sagen oder lügen.

Hatespeech – Worin äußert sich der Hass?

- Social Media zielt auf emotionale Zustimmung bzw. Ablehnung, aus einfachen „Daumen hoch“ - oder „Wut“. Emojis entwickeln sich extreme Emotionen: begeisterte Zustimmung oder radikale Ablehnung
- vielfältige Formen von negativen Emotionsäußerungen: extremistische Propaganda bis Mobbing von Einzelpersonen
- Hass-Postings ursprünglich auf rechtsextremistische Homepages beschränkt, heute primär Social Media, aber auch Online-Games
- Täter hauptsächlich junge Männer aus rechter Ecke, dennoch Phänomen, das ganze Gesellschaft betrifft
- Zielscheibe von Hass sind häufig Randgruppen oder besonders exponierte Personen

Woher kommt der Hass?

- meist spontan, beim Surfen staut sich Wut auf, Gruppenpolarisierung verstärken positive wie negative Emotionen
- Trolle mit sadistischen Eigenschaften, teilweise bezahlt
- Internet bietet Möglichkeit, Angriffe schnell und effektiv zu koordinieren
- entkörperlichtes Handeln fördert distanzloses Kommunizieren von Emotionen, senkt Hemmschwelle, fehlende digitale Empathie
- Emotionalisierung von Diskursen in Gesellschaft und Politik ist altes Problem

Wo liegt das Problem?

- Emotionalisierung erschwert rationalen Diskurs bzw. überhaupt Argumente wahrzunehmen, vermehrte Entstehung von Filter Bubbles
- Bei Hate Speech leistet Dehumanisierung des Gegners realer Gewalt Vorschub
- Hate Speech zielt auf kognitive und emotionale Verletzbarkeit ab -> „Körperkraft der Sprache“
- Minderheit diktiert den Diskurs, indem Angst verbreitet wird bzw. weil lautes Auftreten (z.B. von Fans) Menschen mit anderer Meinung verstummen lässt
- Angst verursacht Selbstzensur, Meinungen werden nicht mehr gehört
- Sensibilisierung einer Gesellschaft gegenüber Aggression und Gewalt als zunehmende Humanisierung

Was kann man dagegen tun?

- Medienkompetenz fördern, auf rationale Diskurse setzen, Filter Bubbles aufbrechen -> Möglichkeiten der Kunst
- technische Lösungen bei Hate Speech: Popup-Fenster halten Täter Spiegel vor Augen
- juristische Lösung: Task Force und Strafverfolgung, damit bei Opfern nicht Gefühl der Hilflosigkeit bleibt, außerdem präventive Wirkung
- Schuldzuweisungen an Opfer klar zurückweisen
- Vermittlung von Wissen darüber, dass Menschen durch Sprache verletzlich sind
- Kampagnen lancieren: „Wer hetzt, macht sich strafbar“, Netzkodex schaffen

Leben in der Online-Welt...

- **Für mich selbst:**
 - > sich schützen und Hilfe annehmen
 - > Ursachen für eigene Reaktion suchen
 - > differenzierte Position entwickeln
 - > auf „Framing“ achten
 - > Selbstwirksamkeit erleben

