Deutsche Stiftung für Recht und Informatik

CHATBOTS ALS KI-SYSTEME MIT BESONDEREN TRANSPARENZPFLICHTEN NACH ART. 52 KI-VERORDNUNG

RA Kilian Georg Wolf
Simmons & Simmons LLP

Herbstakademie 2022

Agenda

- Technische Grundlagen
- Typische Anwendungsfälle
- Kategorisierung von KI-Systemen nach der KI-Verordnung
- Kennzeichnungspflicht für Chatbots

Technologischer Fortschritt von ELIZA zu LaMDA

CHATBOT – WIE FUNKTIONIERT DAS?

ELIZA – Prototyp regelbasierter Chatbots

ELIZA (1996)

- Musterübereinstimmungsregeln (”pattern matching”): manuell kodierte Schlüsselwörter
- Begrenzte Möglichkeiten → Trick: Eingabe als Frage zurückspielen

A.L.I.C.E. (1995): AIML - „Artificial Intelligence Markup Language“

Weizenbaum, Joseph: ELIZA – A Computer Program for the Study of Natural Language Communication between Man and Machine, Computational Linguistics 1966, S. 36-45

→ Regelbasierte Chatbots als KI-Systeme?

Machine Learning-Chatbots

- Selbstlernende Systeme: Korpus entscheidend

  → für Hochrisiko-KI-Systeme hochwertige Trainingsdaten vorgeschrieben, Art. 10 KI-VO-E

- Information Retrieval Modelle vs. Generative Modelle

Information Retrieval Modelle:
Fundus an Antworten

Generative Modelle:
Neuronale Netze/Deep Learning (Vokabular + Syntax), Programm erstellt Antwort selbst (NLP „Natural Language Processing“)

Transformer: Gewichtung des Inputs nach Relevanz

→ Google 2021 „LaMDA“

Kundendienst, FAQ & Co.

USE CASES

Kostengünstiger „first point of contact“

Bereiche

- Kundendienst
  → häufig hybrides Setup/vorgeschalteter Chatbot
- Marketing
- Sales
- Recruiting
- Interne Prozesse

Kanäle

- Webseite
- SMS
- Voicebots/Kundenhotline
- Virtual personal assistants (Siri, Alexa…)
- „Facebook Messenger Bot“, „Slackbot“

Risikobasierter Ansatz der KI-Verordnung

CHATBOTS ALS „KI-SYSTEME MIT GERINGEM RISIKO“?

Verordnung zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz 2021/0106 (COD)

- Definition KI-System sehr weit, Art. 3 Nr. 1 i.V.m. Anhang I
- Regulierung KI-Systeme je nach Risiko

Höchstes Risiko	Hohes Risiko	Niedriges Risiko
Verbot, Titel II/Art. 5	Regulierung, Titel III, VII, VIII	Keine Regulierung; Rahmen für freiwillige Codes of Conduct, Titel IX

Transparenzpflichten (Titel IV/Art. 52) gelten für KI-Systeme mit niedrigem Risiko *und* für Hochrisiko-KI-Systeme (Art. 52 Abs. 4)

Kennzeichnungspflicht für Chatbots

…UNLESS THIS IS OBVIOUS…

Transparenzpflicht nach Art. 52 Abs. 1 KI-VO-E

„Die Anbieter stellen sicher, dass KI-Systeme, die für die Interaktion mit natürlichen Personen bestimmt sind, so konzipiert und entwickelt werden, dass natürlichen Personen mitgeteilt wird, dass sie es mit einem KI-System zu tun haben, es sei denn, dies ist aufgrund der Umstände und des Kontexts der Nutzung offensichtlich. (…)“

- Anbietereigenschaft: Niedrige Schwelle (Art. 3 Nr. 2 KI-VO-E)
- Kennzeichnungspflicht, wenn nicht KI-System offensichtlich
- Maßstab für Offensichtlichkeit?

Chatbots aus der Sicht des verständigen Verbrauchers

- KI-Verordnung kein Verbraucherschutzrecht, aber: Transparenzpflichten primär zugunsten Verbraucher

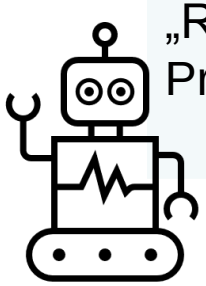
  → Orientierung an EU- / lauterkeitsrechtl. Verbraucherleitbild sachgerecht

> Offensichtlichkeit, wenn einem durchschnittlich informierten, verständigen und situationsangemessen aufmerksamen User unter den gegebenen Umständen ohne weiteres klar ist, dass er es mit einem KI-System zu tun hat

- Gesamtbetrachtung der Interaktionssituation erforderlich

Offensichtlichkeit aufgrund „der Umstände und des Kontexts der Nutzung“ (1)

Indizien für Offensichtlichkeit	Indizien gegen Offensichtlichkeit
Bedienung über Schaltflächen / Vorauswahl Themenfelder	Kommunikation in natürlicher Sprache / Freitexteingabe
„Formular-ähnliche“, technisch anmutende grafische Gestaltung	Verwendung von Sprechblasen, „Messenger-ähnliche“ grafische Gestaltung
Verzögerungsfreie Antwortausgabe, sofortige Bereitstellung von Links/Grafiken	KI „tippt“ augenscheinlich Buchstabe für Buchstabe; „Authentische“ Umgangssprache…
„Roboterkopf“ o.ä. als Profilbild/Avatar des Chatbots	„Mitarbeiterpiktogramme“ o.ä.



Offensichtlichkeit aufgrund „der Umstände und des Kontexts der Nutzung“ (2)

KI-System offensichtlich bei Chatfenster-Aufschrift:

- „Kundendienst-Chatbot“
- „elektronischer/automatischer Kundenbetreuer“
- „Robo-Expertin“
- „virtueller Assistent“

Bot vermenschlicht, aber wohl dennoch offensichtlich:

- „Chatten Sie hier mit unserem Bot“
- „Tippen Sie hier Ihre Frage ein, unser Bot wird Ihnen weiterhelfen“
- „Hey, ich bin [Name], dein Support-Chatbot“

Offensichtlichkeit aufgrund „der Umstände und des Kontexts der Nutzung“ (3)

KI-System nicht offensichtlich:

- „Hallo, ich bin [Name], wie kann ich dir helfen?“
- „[B.] Kundenbetreuung: Willkommen im [B.] Chat. Bitte formulieren Sie uns vorab Ihr Anliegen.“
- „[N.]-Kundendienst: Hallo, was können wir heute für Sie tun?“

Kennzeichnungspflicht

- Spätestens zu Beginn der Unterhaltung
- Eng an den Verordnungstext angelehnte Formulierungen:
  - „Bitte beachten Sie, dass Sie es mit einem KI-System zu tun haben“
  - „Sie interagieren mit einem KI-System“
- Varianten:
  - „Sie unterhalten sich mit einer AI“
  - „Dies ist ein intelligentes Chat-System“

Kennzeichnung

Offensichtlichkeit

Kilian Georg Wolf, Rechtsanwalt
Simmons & Simmons LLP, München

Lehel Carré
Thierschplatz 6
80538 München

089 20 80 77 63-36
kilian.wolf@simmons-simmons.com

VIELEN DANK