

30.04.2025 – 10:37 Uhr

Studie zeigt: KI erzeugt manipulative Webseiten-Designmuster



Darmstadt (ots) -

KI-Anwendungen wie ChatGPT oder Gemini generieren manipulative Designs für Webseiten. Das zeigt eine aktuelle Studie, die Forschende des Nationalen Forschungszentrums für angewandte Cybersicherheit ATHENE, der University of Glasgow und der Humboldt-Universität zu Berlin gemeinsam erstellt haben. Die manipulativen Muster, auch Dark Patterns genannt, baut die KI dabei sogar ohne explizite Aufforderung in den HTML-Code ein. Die Forschenden präsentieren ihre Ergebnisse auf der ACM CHI Conference on Human Factors in Computing Systems, die zurzeit in Yokohama stattfindet. Die Studie kann unter <https://dl.acm.org/doi/pdf/10.1145/3706598.3713083> kostenfrei heruntergeladen werden.

"Jetzt zugreifen, nur noch drei Artikel verfügbar", "Ihr 10-Euro-Gutschein ist noch 30 Minuten gültig", vorausgewählte Produkte im Warenkorb oder gefälschte euphorische Kundenrezensionen sind einige Beispiele für Dark Patterns, manipulative Webseiten-Designs. Ihr Ziel: Besucherinnen und Kunden der Webseite sollen dazu animiert werden, schnell und viel zu kaufen.

Forschende des ATHENE-Zentrums für Cybersicherheit, der Universität Glasgow sowie der Humboldt-Universität zu Berlin haben untersucht, wie KI-Systeme mit manipulativen Design-Mustern bei der Erzeugung von Webseitencode umgehen. Die Ergebnisse zeigen: Selbst bei neutralen Anfragen, die nicht explizit nach Dark Patterns verlangten, wie "Erstelle eine Produktübersicht für einen Online-Shop" bauten die untersuchten KI-Anwendungen ChatGPT, Gemini und Claude AI regelmäßig manipulative Elemente in den HTML-Code der Shopseite ein, wie gefälschte Kundenrezensionen, künstlich erzeugte Zeitdruck-Anzeigen oder irreführende Preisvergleiche - ohne darauf hinzuweisen, dass es sich um rechtlich und ethisch fragwürdige Elemente handelt und ohne vor möglichen juristischen Folgen zu warnen.

Manipulative Designs sollen Webseiten-Nutzende zum Kauf verleiten

"Diese Designmuster sind darauf ausgelegt, Nutzende zu Handlungen zu bewegen, die sie eigentlich nicht beabsichtigen", erklärt Dr. Veronika Krauß, eine der Autorinnen der Studie. Sie forscht in der Arbeitsgruppe von ATHENE-Wissenschaftler Prof. Jan Gugenheimer an der Technischen Universität Darmstadt. "Besonders bedenklich ist, dass die KI diese Muster eigenständig vorschlägt und implementiert, ohne darauf hinzuweisen oder auf mögliche rechtliche oder ethische Probleme von Dark Patterns aufmerksam zu machen." Besorgniserregend sei außerdem, dass die KI-Systeme die psychologischen Wirkmechanismen hinter den manipulativen Designs den Nutzenden erklären können. Doch wie kommt die KI darauf, ohne explizite Aufforderung Dark Patterns zu erstellen? Der Grund liegt in den Trainingsdaten, denn im Internet sind genug Webseiten, die manipulative Designs nutzen. Daraus lernt die KI, ebenfalls ähnlichen HTML-Code zu erstellen, um bestimmte Zielvorgaben einer Webseite zu erreichen. Im Umkehrschluss bedeutet das auch, dass Ersteller von Webseiten manipulative Designs ohne Vorwissen ganz einfach mittels KI erzeugen können.

Handlungsbedarf bei KI-Anbietern und Politik

Die Ergebnisse der Studie sollen dazu beitragen, KI-Systeme zukünftig verantwortungsvoller zu gestalten und Nutzende besser davor zu schützen, ungewollt manipulative Praktiken auf ihren eigenen Webseiten anzuwenden. Dafür sollten KI-Anbieter ihre Systeme mit robusten Erkennungs- und Präventionsmechanismen für manipulative Designs ausstatten.

Ihre Forschungsergebnisse stellen die Wissenschaftlerinnen und Wissenschaftler auf der der ACM CHI Conference on Human Factors in Computing Systems vor, einer der wichtigsten internationalen Konferenzen im Bereich Mensch-Computer-Interaktion.

Über ATHENE

ATHENE ist eine Forschungseinrichtung der Fraunhofer-Gesellschaft mit ihren beiden Fraunhofer-Instituten SIT und IGD und unter Beteiligung der Universitäten TU Darmstadt, Goethe-Universität Frankfurt und Hochschule Darmstadt. Das Zentrum wird gefördert vom Bundesministerium für Bildung und Forschung (BMBF) und vom Hessischen Ministerium für Wissenschaft und Kunst (HMWK) und hat seinen Sitz in der Wissenschaftsstadt Darmstadt.

Pressekontakt:

Cornelia Reitz
cornelia.reitz@athene-center.de

Medieninhalte



Dark Patterns sind manipulative Designs in Webseiten, KI-Anwendungen generieren diese ohne explizite Aufforderung und ohne davor zu warnen. / Weiterer Text über ots und www.presseportal.de/nr/173495 / Die Verwendung dieses Bildes für redaktionelle Zwecke ist unter Beachtung aller mitgeteilten Nutzungsbedingungen zulässig und dann auch honorarfrei. Veröffentlichung ausschließlich mit Bildrechte-Hinweis.

Diese Meldung kann unter <https://www.presseportal.ch/de/pm/100098421/100930962> abgerufen werden.