

ZUKUNFTSWERKSTATT: Künstliche Intelligenz im Kinder- und Jugendmedienschutz

Rund 100 Expertinnen und Experten trafen sich am 17. September 2025 zur ZUKUNFTSWERKSTATT-Veranstaltung „Ich prompte mir die Welt, wie sie mir gefällt.‘ Kindheit und Jugend unter dem Einfluss künstlicher Intelligenz.“ in Berlin. Ziel des Fachaus-tauschs der Verantwortungsgemeinschaft für den Kinder- und Jugendmedienschutz aus Staat, Wirt-schaft und Zivilgesellschaft war der Diskurs über die Chancen und Risiken künstlicher Intelligenz (KI) für Kinder und Jugendliche.

KI als Instrument digitaler sexueller Gewalt

Das Vormittagsprogramm war geprägt von drei Fach-vorträgen. Kathi Heffe, Betroffenenberaterin bei der HateAid gGmbH sprach über Deepfakes und Deep-nudes, bei denen Künstliche Intelligenz gezielt für digitale sexuelle Gewalt eingesetzt werde, teils mit gravierenden Folgen für die Betroffenen. Die meisten Deepfakes seien pornografischer Natur. In rund 99 Prozent der Fälle seien Frauen von derartigen porno-grafischen Darstellungen, sogenannten Deepnudes, betroffen. Die Folgen reichten von Scham- und Schuldgefühlen bis hin zu schwerwiegenden psycho-logischen, sozialen sowie möglicherweise beruflichen Folgen. Erschwerend käme hinzu, dass betroffenen Personen oft nur inadäquate Hilfe angeboten würde und so unter Umständen eine Sekundär-Vikti-misierung erfolge.

Extremistische Ansprache von Kindern und Jugendlichen

Dr. Bernd Zywiets, Bereichsleiter Politischer Extre-mismus bei jugendschutz.net, dem gemeinsamen

Kompetenzzentrum von Bund und Ländern für den Schutz von Kindern und Jugendlichen im Internet, sprach über die extremistische Ansprache von Kin-dern und Jugendlichen online. Künstliche Intelligenz diene dabei als Attraktionsmittel und „Brandbe-schleuniger“. Er stellte fest, dass extremistische Ak-teurinnen und Akteure KI häufig als leicht zugängli-ches und günstiges Mittel zur Erstellung extremisti-schen Contents für Social Media nutzten. Dabei wür-de stark auf emotionalisierende Inhalte gesetzt oder alternativ in jugendaffiner Sprache und Ästhetik, oft unter dem Deckmantel des Humors, Hass und Hetze, verbreitet. Die durch die Nutzung von KI entstande-nen Zweifel an der Authentizität von Bildern und Vi-deos würden von extremistischen Akteurinnen und Akteuren instrumentalisiert, um eigentlich wirklich-keitsgetreue mediale Darstellungen im Rahmen von Pseudo-Faktenchecks in Zweifel zu ziehen und so politische Gegner zu diskreditieren.

Parasoziale Beziehungen mit KI-Chatbots

Fabian Sauer, Mitgründer und Projektleiter Bildung bei der mecodia GmbH, die unter anderem handy-sektor.de redaktionell betreut, sprach im Anschluss über parasoziale Beziehungen Kinder und Jugendli-cher zu Chatbots. Die dauerhaft verfügbaren Chat-bots böten vermeintliche Nähe und Intimität, In-teraktivität und Gegenseitigkeit. Für Kinder und Ju-gendliche auf der Suche nach Austausch bestünde die Gefahr der Isolation und (emotionalen) Abhän-gigkeit. Zudem sei die Sorge berechtigt, dass ihre real-weltlichen Kommunikationsfähigkeiten leiden könnten, da Chatbots in der Regel keine Konflikte böten und somit Gelegenheiten – beispielsweise zur Erlernung von Kompromissfähigkeit – verloren

gingen. Gleichzeitig könnten Chatbots jedoch auch als Übungsräume für Sozialverhalten dienen und als Lernhilfen eingesetzt werden. Erste Untersuchungen zeigten zudem, dass sich insbesondere einsame Kinder und Jugendliche Chatbots als „Freunde“ suchten, was die immer wieder geäußerte Befürchtung in Zweifel ziehen würde, dass KI-Chatbots junge Menschen erst einsam machen würden.

Diskussion mit Familienministerin Karin Prien

Auf die Vorträge des Vormittags folgte – dem Gedanken der ZUKUNFTSWERKSTATT als Diskursraum treu bleibend – der interaktive Austausch der Teilnehmenden in verschiedenen Formaten. Den Beginn machte ein Jugendpanel mit fünf jungen Menschen zwischen 12 und 19 Jahren. Die Jugendlichen berichteten von ihren Erfahrungen mit KI, aber auch mit Social Media und Smartphones, insbesondere im schulischen Kontext. Im Anschluss fand ein World Café statt, in dem alle Teilnehmenden in Kleingruppen diskutierten und gemeinsam Erkenntnisse sammelten. Die Veranstaltung schloss mit einer großen Paneldiskussion unter Beteiligung von Karin Prien, Bundesministerin für Bildung, Familie, Senioren, Frauen und Jugend.

Erkenntnisse und Anregungen

Im Folgenden sind einige der Kernerkenntnisse und Anregungen der Teilnehmenden, die in der ZUKUNFTSWERKSTATT-Fachtagung entstanden sind, überblicksartig dargestellt. Diese beziehen sich naturgemäß nicht nur auf KI im Kinder- und Jugendmedienschutz, sondern beziehen weitere allgemeine Erkenntnisse und Ideen mit ein.

Erkenntnis 1: Social Media ist Kern sozialer Teilhabe für Kinder und Jugendliche. Insbesondere viele Jugendliche sind auf Social-Media-Plattformen aktiv und nutzen diese für Austausch und Interaktion, Information und Unterhaltung – sie halten also so Kontakt zu ihren Altersgenossen und erschließen sich die Welt.

Erkenntnis 2: Pauschale Verbote bedeuten der Annahme der Teilnehmenden der ZUKUNFTSWERKSTATT nach ein Ungleichgewicht zwischen Schutz, Befähigung und Teilhabe – zugunsten von Schutz, aber insbesondere auf Kosten der Teilhabe und zumindest in Teilen auf Kosten der Befähigung.

Erkenntnis 3: Mit Blick auf eine mögliche Prävention für KI-bedingte Risiken bedarf es eines

Zusammenspiels aus gesetzlichen Regelungen und KI-Kompetenzen.

Erkenntnis 4: Medienpädagogische Angebote, die unter anderem auch für die Vermittlung von KI-Kompetenzen notwendig sind, stehen vor erheblichen Budgetherausforderungen. Entsprechende Bildungsangebote sind ressourcenintensiv, das heißt es bräuchte mehr, dauerhafte und an den richtigen Stellen gezielt eingesetzte Finanzierungen.

Altersgerechte Zugänge

Mit Blick auf die Anbieter forderten die Teilnehmenden erneut, dass Plattformen im Sinne des Digital Services Act positiv und sicher gestaltet werden. Sie sollen sichere Räume für Kinder und Jugendliche schaffen und deren altersangemessene Teilhabe ermöglichen. Als notwendige Voraussetzungen nannten sie unter anderem wirksame Altersüberprüfungssysteme in Kombination mit Alterskennzeichnungen, sichere Voreinstellungen sowie funktionierende Melde- und Abhilfemechanismen. Als qualitative Merkmale funktionierender Melde- und Abhilfeverfahren nannten die Teilnehmenden, dass die Verfahren von vornherein offen und transparent gestaltet sein sollten. Nutzende sollten bereits vor dem Absetzen einer Meldung in altersangemessener Weise über den Ablauf der Meldung und das nachfolgende Verfahren in Kenntnis gesetzt werden. Damit Meldeverfahren nicht abschreckend wirkten, sei auf eine einladende, ermutigende und für Kinder und Jugendliche zugängliche beziehungsweise verständliche Gestaltung zu achten. Die auf eine Meldung folgenden Abhilfeverfahren sollten seitens der Anbieter schnell bearbeitet werden. Diejenigen Nutzenden, die eine Meldung machen, sollten eine Rückmeldung zur Bearbeitung der Meldung erhalten – damit bekannt werde, wie mit dem gemeldeten Inhalt weiter verfahren wurde. Zuletzt wurde angeregt, in den Meldeverfahren auf Beratungsangebote hinzuweisen, bei denen Betroffene anbieterunabhängige Hilfe erhalten können.

Als weitere Schlüsseltechnologie wurden Altersüberprüfungssysteme gefordert, wie sie beispielsweise in den Guidelines der EU-Kommission gemäß Artikel 27 Absatz 4 des Digital Services Act beschrieben werden. Diese sollen sicherstellen, dass Kinder und Jugendliche nur entsprechend ihres Alters einen adäquaten Zugang erhalten. Zudem können sie als Grundlage zur altersabgestuften Steuerung sicherer Voreinstellungen dienen. Je nach ermitteltem Alter sollten diese mehr oder weniger strikt sein. Dennoch

sollten sie im Nachhinein veränderlich sein, gegebenenfalls nach Aufklärung über mit einer Änderung einhergehende Risiken und gegebenenfalls unter Einbezug der personensorgeberechtigten Personen. Die Plattformen und ihre Voreinstellungen sollten damit dem Prinzip Safety by Design/Safety by Default (inklusive Opt-out) entsprechen. Unter Einbezug der genannten Gestaltungsprinzipien könnten dem Alter entsprechend angemessene Angebote geschaffen und so sichere und altersgerechte Zugänge ermöglicht werden.

Präventive Angebote

Mit Blick auf präventive Angebote im Kinder- und Jugendmedienschutz wurde angeregt, Prävention grundsätzlich ganzheitlich und vom Kind aus zu denken. Dabei sollte die gesamte Lebensrealität von Kindern und Jugendlichen betrachtet werden – nicht nur ihre medialen Erfahrungen.

KI-Inhalte sollten stets als solche gekennzeichnet werden, um insbesondere mit Blick in die Zukunft die Erkennung zu erleichtern und Verwirrungseffekten entgegenwirken zu können. Es sollten Akteurinnen und Akteure gestärkt werden, die positive Botschaften kommunizieren, um Kindern und Jugendlichen einen Zugang zur KI-Technologie zu vermitteln sowie ihnen die altersgerechte Nutzung von Chancen, ihre Teilhabe und Befähigung zu ermöglichen. Sowohl Eltern, Bezugspersonen als auch Lehr- und Fachkräfte müssten mit Blick auf KI mehr Bildungsangebote erhalten und damit Fachwissen erwerben. Bestehende Hilfsangebote sollten sichtbarer gemacht werden, damit Kinder und Jugendliche schnell Hilfe finden können. Die Fortbildung von Eltern, Fach- und Lehrkräften sowie kompetente Hilfsangebote können auch bei der Vermeidung von Sekundär-Viktimisierung helfen.

An die gesamte Kinder- und Jugendmedienschutzszene richteten die Teilnehmenden der ZUKUNFTSWERKSTATT die Erwartungshaltung, sich angesichts der raschen technischen Entwicklung weiterhin mit den Risiken und Chancen von KI zu beschäftigen und das entsprechende Wissen an Kinder und Jugendliche zu vermitteln. Allein schon das Wissen um und das Verstehen von Chancen und Risiken könne ein erster Ansatz sein, um Sicherheit zu schaffen und Kindern und Jugendlichen die Nutzung von Chancen zu ermöglichen. Gerade bei diesen sollte Medienkompetenz als zentrale Kulturtechnik unserer heutigen Zeit bestärkt und somit eine positive Haltung gefördert werden. Dazu gehöre, im

Lernprozess KI gemeinsam mit Eltern oder Lehrkräften auszuprobieren. So ließen sich Chancen, aber auch Grenzen und Risiken der Technologie praxisnah erfahrbar machen und das gemeinsame Ausprobieren sei ein geschützter Raum für Reflektion und kritische Auseinandersetzung, beispielsweise mit den Outputs eines Chatbots.

Um Medienkompetenz zielgruppengerecht und unterhaltsam-subtil zu vermitteln, wurde angeregt, mit bereits etablierten, großen Influencerinnen und Influencern zusammenzuarbeiten. Zudem stand die Forderung im Raum, Medienbildung, insbesondere in schulischen Kontexten, verstärkt voranzutreiben und stärker zu etablieren.