

KI-generierte Abbildungen von Kindesmissbrauch

Technische Grundlagen und rechtliche Einordnung

Martin Steinebach¹, Julian Drechsler² und Lilly Schröder³

Abstract: Künstliche Intelligenz (KI), die kinderpornografisches Material (CSAM) erzeugt, stellt eine neue Herausforderung im Kampf gegen Kindesmissbrauch dar. Sie kann bestehende Inhalte sexuell aufladen oder gänzlich neue Inhalte erzeugen. Juristisch betrachtet wirft die Verwendung von KI-generiertem CSAM jedoch komplexe Fragen auf. Es gilt zu klären, inwiefern solches Material unter bestehende Gesetze fällt und wie es rechtlich zu bewerten ist. Eine klare Gesetzeslage und Richtlinien, die den Missbrauch dieser Technologien verhindern, sind notwendig, um auch aus Sicht der Technik weitere Schritte in der Bekämpfung des Phänomens zu planen.

Keywords: Künstliche Intelligenz, Jugendschutz, CSAM

1 Einleitung

Die Verbreitung von kinder- und jugendpornographischen Inhalten (englisch CSAM: Child Sexual Abuse Material) über das Internet ist ein bekanntes und viel diskutiertes Problem. Sie ist ein Grund dafür, dass das Darknet, beispielsweise in der Form von Tor-Foren, einen schlechten Ruf in der Öffentlichkeit bekommen hat. Sie ist Gegenstand von Diskussionen über den Umgang mit Inhalten im Internet und der Frage, ob solche Inhalte gelöscht werden können. Und sie ist einer der Gründe für die Bestrebungen der EU, eine Chatkontrolle mit Client-Side-Scanning einzuführen [Co24].

Bisher waren die Darstellungen entweder bereits bekannt und wurden aufgrund von strafbewehrtem Besitz und Verbreitung verfolgt, oder die Inhalte waren unbekannt und führten zusätzlich zur Aufnahme einer Suche nach den Opfern und Tätern. In der jüngeren Vergangenheit ist hier ein neues Phänomen zu beobachten, welches zu zahlreichen

¹ Fraunhofer SIT, Multimedia Sicherheit und IT-Forensik, Rheinstraße 75, 64296 Darmstadt,

martin.steinebach@sit.fraunhofer.de,  <https://orcid.org/0000-0002-0240-0388>

² Doktorand bei Prof. Dr. Indra Spiecker gen. Döhmman, LL.M., Lehrstuhl Recht der Digitalisierung, Universität zu Köln, Albertus-Magnus-Platz, 50923 Köln, Deutschland und Wiss. Mitarbeiter an der Goethe-Universität Frankfurt a.M., Theodor-W.-Adorno-Platz 4, 60629 Frankfurt am Main, Deutschland sowie Wiss. Mitarbeiter bei KASTEL am Karlsruher Institut für Technologie, Vincenz-Prießnitz-Str. 3, 76131 Karlsruhe, Deutschland, julian.drechsler@kit.edu.

³ Doktorandin bei Prof. Dr. Indra Spiecker gen. Döhmman, LL.M., Lehrstuhl Recht der Digitalisierung, Universität zu Köln, Albertus-Magnus-Platz, 50923 Köln, Deutschland und Wiss. Mitarbeiterin an der Goethe-Universität Frankfurt a.M., Theodor-W.-Adorno-Platz 4, 60629 Frankfurt am Main, Deutschland, l.schroeder@jur.uni-frankfurt.de.

Problemen führen kann: Die lebensechten und von realen Personen nicht oder nur schwer unterscheidbaren Darstellungen stammen vermehrt nicht von existierenden Opfern, sondern sind mittels KI-basierter generativer Systeme (GenAI) erstellt oder modifiziert worden. Untersuchungen von großen Trainingsdatensätzen solcher Systeme haben weiterhin gezeigt, dass dort auch reale CSAM eingeflossen sind [Vi23].

Dass CSAM heute auch synthetisch erstellt wird, stellt keine Überraschung dar. Bereits in der frühen Entwicklung von „Deepfakes“ (KI-basierten Manipulationen echter Darstellungen) wurden diese für erotische Inhalte verwendet, damals noch, um bekannte Schauspielerinnen in erotischen Filmen auftauchen zu lassen. Auch das Konzept der DeepNudes, also auf KI basierenden Verfahren, Fotos von bekleideten Personen künstlich in Nacktaufnahmen umzuwandeln, ist ein bereits lang bekanntes Phänomen.

Im Kontext von CSAM wurden ebenfalls schon früh synthetische Verfahren eingesetzt. Bekannt wurde hier Sweety von Terre des Hommes [St20], welche ein synthetisches weibliches Kind war, das als Lockvogel für illegale Videochats eingesetzt wurde. Die Risiken KI generierter Inhalte wurden im Kontext von CSAM bereits untersucht [Ra21] [F122] [Ke23].

Große Aufmerksamkeit erreichte das Thema, als die Internet Watch Foundation („IWF“) die Studie "How AI is being abused to create child sexual abuse imagery" veröffentlichte [In23]. Hier wurden die Ergebnisse einer einmonatigen Überwachung eines einzelnen Forums aufgeführt, die zeigten, dass tausende künstlich erstellter CSAM-Fälle zirkuliert wurden. In Großbritannien wurde nun erstmals eine Person wegen der Herstellung von CSAM durch KI zu einer Geldstrafe und einem KI-Verbot verurteilt [Da24].

Die vorliegende Arbeit stellt eine Erweiterung der rechtlichen Diskussion aus [St24] dar.

2 Methoden

Es existieren verschiedenen Verfahren, die zum Erzeugen künstlicher CSAMs eingesetzt werden können [St24]. Sie basieren auf verschiedenen Methoden des maschinellen Lernens. Ein grundlegender Unterschied ist, ob sie vorhandene Inhalte verändern oder Inhalte neu erstellen. Dabei gibt es allerdings auch Überlappungen, wenn Teile bestehender Inhalte durch künstlich erzeugte Inhalte ersetzt werden.

Die heute sicher bekanntesten Verfahren zum Modifizieren bestehender Inhalte, hier insbesondere Videos, sind Deepfakes, deren Potenzial für ethisch problematische Inhalte bereits frühzeitig diskutiert wurde [KR21] [We19]. Die Entwicklung von Deepfakes ist inzwischen weit vorangeschritten und ermöglicht in günstigen Konstellationen Ergebnisse, die nur noch schwer als Fälschung erkannt werden können. Im Kontext von CSAM bedeutet das, dass beispielsweise ein legales Erotikvideo mit einer jung aussehenden Darstellerin oder auch bereits vorhandenes CSAM so verändert wird, dass

das Gesicht einer Minderjährigen eingefügt wird. So entsteht ein Erotikvideo, in dem scheinbar das minderjährige Ziel gezeigt wird.

Für einzelne Bilder kann auch DeepNude [YC20] [Gr21] zum Einsatz kommen. Im ursprünglichen Sinne war DeepNude eine Anwendung, die auf Generative Adversarial Networks („GANs“) basierte, einer spezifischen Form von neuronalen Netzwerken in der künstlichen Intelligenz. Im Falle von DeepNude wurde GAN verwendet, um Kleidung von Bildern von Frauen zu entfernen und nackt erscheinende Versionen zu erzeugen. Das GAN von DeepNude wurde auf einer großen Menge von Bildern trainiert, um die Fähigkeit zu entwickeln, realistisch aussehende nackte Versionen von Personen zu generieren. Zwar wurde das ursprüngliche DeepNude-Verfahren aufgrund ethischer Bedenken eingestellt, es existieren aber zahlreiche Nachahmungen, die entsprechende Bearbeitungen ermöglichen. Als Ausgangsmaterial eignen sich dabei Bilder, in denen die Personen bereits große Anteile von Haut zeigen, also beispielsweise Strandfotos im Bikini. Entsprechende Verfahren können auch bei Minderjährigen eingesetzt werden. Inwiefern dies auch bei Kindern realistische Ergebnisse erzielt, kommt dabei auf die Trainingsdaten an, auf denen das Netz basiert.

Zum Erstellen von neuen Inhalten haben sich Text-to-Image-Verfahren etabliert. Sie wandeln natürlichsprachlichen Textbeschreibungen zu visuellen Repräsentationen um. Der Prozess beginnt mit der Bereitstellung von Trainingsdaten, bestehend aus Paaren von textuellen Beschreibungen und zugehörigen Bildern. Das Text-to-Image-Modell wird dann auf diese Daten trainiert, wobei der Fokus darauf liegt, eine korrekte und semantisch kohärente Zuordnung zwischen den textuellen und visuellen Modalitäten herzustellen.

Ob sich Text-to-Image dazu eignet, CSAM zu erstellen [TS23], hängt von den Trainingsdaten ab. Etablierte Anbieter entsprechender Verfahren versuchen, entsprechende Daten aus ihren Trainingssets zu filtern. Diese können allerdings durch zusätzliche Module, die die trainierten Netze anhand weniger Beispiele anpassen, trotzdem in die Lage versetzt werden, CSAM zu erstellen. Das gilt insbesondere, wenn die Verfahren frei verfügbar und lokal installierbar sind. Dementsprechend wird auch von neu erzeugten Abbildungen bekannter CSAM-Opfer anhand von Modellen berichtet, die mittels vorhandener Abbildungen trainiert bzw. angepasst wurden [In23].

Die Text-to-Image-Verfahren müssen aber nicht dazu verwendet werden, Bilder von Grund auf neu zu erzeugen. Mit Inpainting ist eine Variante verfügbar, die gezielte Veränderungen von Bildern anhand von Textvorgaben ermöglicht. Inpainting ist eine fortgeschrittene Methode der Bildverarbeitung, die darauf abzielt, fehlende oder beschädigte Teile eines Bildes durch die Berechnung und Rekonstruktion von Pixelinformationen aus der umgebenden Bildumgebung zu restaurieren. Der Prozess der Inpainting-Algorithmen basiert auf der Annahme, dass die umliegenden Pixel relevante Informationen über die Struktur und Textur des Bildes liefern, die zur korrekten Wiederherstellung der fehlenden Bereiche herangezogen werden können. Die Verbindung zwischen diesen Konzepten liegt darin, dass Text-to-Image-Generierungstechniken

verwendet werden können, um fehlende oder beschädigte Teile eines Bildes zu rekonstruieren, was einem Inpainting-Ansatz ähnelt. Wenn ein Text-to-Image-Modell trainiert ist, Text in realistische Bilder umzuwandeln, kann es auch dazu verwendet werden, visuell konsistente Ergänzungen für fehlende Teile eines Bildes zu generieren.

Dementsprechend ergeben sich durch Inpainting für synthetisches CSAM verschiedene Möglichkeiten. Das harmlose Foto eines Kindes kann zu einem erotischen Foto umgewandelt werden, beispielsweise durch das Ersetzen oder Entfernen von Kleidung. Dieser Ansatz ist in seiner Wirkung ähnlich dem oben genannten DeepNude-Verfahren. Er ist aber weniger abhängig von geeignetem Ausgangsmaterial, kann also auch bei vollständig bekleideten Personen eingesetzt werden.

3 Herausforderungen

Die Risiken, die durch KI-generierte CSAM entstehen, können inzwischen anhand erster Fälle in der Praxis beobachtet werden. Ein Fall wurde in Spanien bekannt [He23]: Fotos von mindestens 20 Mädchen im Alter von 11 bis 17 Jahren wurden in Nacktaufnahmen umgewandelt. Die Verdächtigen waren zwischen 12 und 14 Jahre alt. In Florida erzeugten zwei Schüler Nacktfotos von zwei Dutzend Klassenkameraden. Nach Angaben der Ermittler hätten sie Bilder vom Social-Media-Konto der Schule verwendet und eine App eingesetzt, um die Gesichter mit KI-generierten Körpern zu verbinden [Tu23].

Das führt zu zahlreichen Herausforderungen und Problemen. Die Opfer solcher Handlungen können Gegenstand von Mobbing werden. Der Fall von Amanda Todd, die aufgrund verbreiteter Nacktaufnahmen und darauf basierendem Mobbing Suizid beging [LM13] hat gezeigt, welche dramatischen Folgen es haben kann, wenn Sexting-Aufnahmen von Kindern und Jugendlichen zirkuliert werden. Im Falle von künstlichen Bildern ist hier die Schwelle niedrig, entsprechende Aufnahmen zu erzeugen.

Denkbar ist auch, dass dieser Umstand ausgenutzt wird, um die Opfer zu erpressen. Ein Täter erstellt entsprechende Bilder und droht dann den Opfern, diese zu verbreiten. Dadurch können echte Aufnahmen oder sexuelle Handlungen erzwungen werden.

Beim Cyber-Grooming, der gezielten Kontaktaufnahme mit Minderjährigen mit sexuellen Absichten, können die veränderten Aufnahmen eingesetzt werden. Täter können Kinder durch Zeigen von künstlich erstellten erotischen Aufnahmen aus deren Freundeskreis davon überzeugen, dass es 'in Ordnung' oder 'ganz normal' sei, wenn sie entsprechende Aufnahmen von sich machen lassen.

Auch das Umfeld der dargestellten Kinder und Jugendlichen kann in Mitleidenschaft gezogen werden. Erotische Aufnahmen von Kindern, die offensichtlich aus dem engeren Familienkreis stammen müssen, beispielsweise im Haus auf der Couch des Wohnzimmers oder im Schlafzimmer des Kindes können zu Untersuchungen durch die Polizei und zur Rufschädigung führen. Ein Nachweis, dass die Bilder im Original harmlos sind und über

soziale Medien verbreitet wurden, wo sie der eigentliche Täter gefunden und verändert hat, ist nicht einfach zu führen. Ebenso können Familienmitglieder sich gegenseitig verdächtigen, Täter zu sein.

Für Ermittler sind künstliche Aufnahmen, die virtuelle Opfer zeigen, potenziell eine zusätzliche Belastung. Können sie nicht erkennen, dass die Opfer nicht existieren, müssen sie davon ausgehen, dass diese Hilfe benötigen und verwenden Ressourcen darauf, sie zu finden und zu befreien. Abhängig von den erzeugten Inhalten kann weiterhin allein das Erzeugen und Verbreiten strafbewehrt sein, was eine Verfolgung der Täter erfordert.

4 Rechtliche Einordnung

Auch für das deutsche Recht ist die Generierung von CSAM mithilfe künstlicher Intelligenz ein neues Phänomen. Bislang ließ sich auch ohne den Einsatz von KI schon täuschend echtes (fiktives) CSAM herstellen. Voraussetzung dafür waren aber zunächst die dafür notwendigen Fähigkeiten (z.B. im Umgang mit Foto- und Videobearbeitungsprogrammen) sowie ein hoher Arbeitsaufwand. Bild- und videogenerierende KI-Anwendungen sind zum einen wesentlich unkomplizierter zu nutzen, zum anderen erleichtern sie den menschlichen Arbeitsaufwand bei der Herstellung von Inhalten um ein Weites. Es ist daher nicht unwahrscheinlich, dass die Anzahl neu hergestellter (täuschend echter) CSAM durch die Verbreitung von Bild- und videogenerierenden KI-Anwendungen deutlich steigern wird. Auf EU-Ebene wird derzeit ein Richtlinienentwurf diskutiert, der eine Strafbarkeit vorsieht für die nicht-einvernehmliche Herstellung oder Manipulation von Inhalten, die den Anschein erwecken, dass die abgebildete Person sexuelle Handlungen vornimmt (unabhängig davon, ob es sich um CSAM handelt) [Eu22]. Ebenfalls auf EU-Ebene wird Art. 50 Abs. 4 des AI Acts [Eu24] künftig eine Kennzeichnungspflicht für Deepfakes vorschreiben, . Im Übrigen finden aber die bereits bestehenden Gesetze Anwendung.

4.1 Strafrecht

§§ 184b und 184c des Strafgesetzbuches („StGB“) stellen bestimmte Handlungen in Bezug auf kinder- und jugendpornographische Inhalte unter Strafe. Die Tathandlung betreffend, laufen die Strafnormen im Wesentlichen parallel (die zugrundeliegende EU-Richtlinie 2011/93/EU differenziert begrifflich gar nicht erst zwischen Kindern und Jugendlichen). Das Verbot jugendpornographischer Inhalte (§ 184c StGB) hat jedoch einen geringeren Strafrahmen und sieht eine Ausnahme für den persönlichen Gebrauch einvernehmlich hergestellter Inhalte vor (§ 184c Abs. 4 StGB).

Ein Kind i.S.d. StGB ist jede Person unter 14 Jahren (§ 184b Abs. 1 Satz 1 Nr. 1 StGB), ein Jugendlicher eine Person im Alter von 15 bis 17 Jahren. Ist das Alter einer

erkennbaren, abgebildeten Person bekannt, kommt es für die rechtliche Einordnung allein auf das tatsächliche Alter an, selbst wenn die Person älter aussieht [Bu23][Bu01]. §§ 184b und 184c StGB betreffen allerdings nicht nur tatsächliche Darstellungen, sondern auch fiktive (z.B. auch Comics oder Romane [Bu13]). In der Regel wird man das Alter einer rein fiktiven Person, deren Abbildung mithilfe von KI generiert wurde, nicht anhand des Personalausweises bestimmen können. Darüber hinaus darf die Entscheidung über Strafbar- und Strafflosigkeit auch nicht allein in den Händen des Herstellers eines Inhalts liegen, indem dieser das Alter frei wählt. Es ist daher im Zweifel anhand von objektiven Kriterien zu beurteilen, ob eine dargestellte Person kindlich, jugendlich oder erwachsen aussieht. Die Rechtsprechung stellt auf den Eindruck ab, den ein objektiver, gewissenhaft urteilender Betrachter aufgrund einer Gesamtwürdigung erlangt [Bu23]. Beurteilungskriterien sind etwa die körperliche Entwicklung, Gestik und Mimik, ggf. Äußerungen und Verhalten der abgebildeten Person, aber auch weitere Faktoren wie die Räumlichkeit, in der sich das abgebildete Szenario abspielt, Bekleidungsstücke (z.B. Kinderbekleidung), weitere Gegenstände (z.B. Kinderspielzeug), sowie textliche oder sprachliche Altersangaben in dem Inhalt oder dessen Bezeichnung (Dateiname) [Bu23]. Letzteres dürfte auch auf Eingabeinformationen, auf deren Grundlage eine KI einen bestimmten Inhalt generiert (sog. „Prompts“), zutreffen.

Die inhaltlichen Anforderungen an kinder- und jugendpornographische Inhalte finden sich in § 184b Abs. 1 Satz 1 bzw. § 184c Abs. 1 StGB. Demnach erfassen kinder- und jugendpornographische Inhalte Darstellungen sexueller Handlungen von, an oder vor dem Opfer, Darstellungen eines ganz oder teilweise unbedeckten Opfers in aufreizend geschlechtsbetonter Körperhaltung und Darstellungen der unbedeckten Genitalien oder des unbedeckten Gesäßes des Opfers in sexuell aufreizender Weise.

Sexuelle Handlungen sind Verhaltensweisen, die allein gemessen an ihrem äußeren Erscheinungsbild einen Sexualbezug aufweisen [Bu15][Bu07]. Nacktheit ist für eine sexuelle Handlung keine zwingende Voraussetzung [Bu14][Bu13a]. Die vollständige oder teilweise Nacktheit ist aber für die folgenden Varianten relevant. Darstellungen eines ganz oder teilweise unbedeckten Opfers sowie Darstellungen der unbedeckten Genitalien oder des unbedeckten Gesäßes eines Opfers erfassen Inhalte, auf denen keine „Handlung“ wiedergegeben ist. Ersteres stellt auf die Körperhaltung ab, letzteres auf das abgebildete Körperteil. Auch bei diesen Darstellungen muss allerdings ein Sexualbezug erkennbar sein (aufreizend geschlechtsbetont bzw. sexuell aufreizend), sodass nichtsexuelle Badewannen- oder Strandfotos, sowie Abbildungen in medizinischem Lehrmaterial nicht erfasst werden [Bu20].

§§ 184b und 184c StGB richten sich primär gegen die Verbreitung oder Veröffentlichung solcher Inhalte. Für eine Verbreitung oder Veröffentlichung ist es notwendig, dass der Inhalt einem größeren, nach Zahl und Individualität unbestimmten Personenkreis zugänglich gemacht wird [Bu20a], z.B. durch das Teilen in Online-Foren. Wer einen kinder- oder jugendpornographischen Inhalt verbreitet oder der Öffentlichkeit zugänglich

macht, macht sich nach §§ 184b oder 184c StGB strafbar, unabhängig davon, ob dieser Inhalt ein tatsächliches, wirklichkeitsnahes oder rein fiktives Geschehen wiedergibt. Gleichmaßen ist die Herstellung von kinder- und jugendpornographischen Inhalten allumfassend dann verboten, wenn bei der Herstellung eine Absicht der späteren Verbreitung vorliegt (vgl. § 184b Abs. 1 Satz 1 Nr. 4 StGB).

Die Weitergabe an einzelne bestimmte Personen ist keine Verbreitung [Bu20a]. Sie ist nur strafbar, wenn es sich um einen Inhalt handelt, der ein tatsächliches oder wirklichkeitsnahes Geschehen wiedergibt (§ 184b Abs. 1 Satz 1 Nr. 2; § 184c Abs. 1 Satz 1 Nr. 2). Gleiches gilt für die Besitzverschaffung oder den Besitz, mit der Einschränkung, dass eine Strafbarkeit bei jugendpornographischen Inhalten nur vorliegt, wenn es sich um ein tatsächliches Geschehen handelt und soweit keine Einwilligung der dargestellten Person(en) nach § 184c Abs. 4 gegeben ist.

Herstellung an sich, d.h. ohne Verbreitungsabsicht, ist nur dann nach §§ 184b oder 184c StGB strafbar, wenn der Inhalt ein tatsächliches Geschehen wiedergibt (vgl. § 184b Abs. 3) und, wenn es sich um einen jugendpornographischen Inhalt handelt, wenn der Inhalt nicht ausschließlich zum persönlichen Gebrauch mit Einwilligung der dargestellten Person(en) hergestellt wurde. Die Herstellung rein fiktiver kinder- oder jugendpornographischer Inhalte zum persönlichen Gebrauch ist hingegen nicht nach §§ 184b oder 184c StGB strafbar [Bu18].

Da das StGB bei Inhalten danach unterscheidet, ob sie ein tatsächliches, ein wirklichkeitsnahes oder ein fiktives Geschehen wiedergeben, und dies nicht nur entscheidend für das Strafmaß, sondern auch für die Strafbarkeit einer bestimmten Handlung an sich ist, ist eine klare Abgrenzung von hoher Bedeutung. Es stellen sich komplexe Abgrenzungsfragen, wenn mithilfe von KI reale Darstellungen bearbeitet oder mit fiktiven Darstellungen vermischt werden. Etwa kann das Gesicht eines Kindes oder Jugendlichen mithilfe von Deepfake auf einen kindlich wirkenden Körper eines volljährigen Pornodarstellers gesetzt werden. Oder aber dem Gesicht eines Kindes kann mithilfe von Inpainting ein künstlicher Körper angefügt werden. Ebenso kann etwa das Gesicht eines real missbrauchten und gefilmten Kindes mithilfe von Inpainting durch ein fiktives Gesicht ersetzt werden.

Fiktive Darstellungen sind solche, die erkennbar künstlich sind [Bu19b]. Dass KI fiktive Darstellungen generieren kann, z.B. Text-to-Image, ist unstrittig.

Komplizierter wird es bei wirklichkeitsnahen oder tatsächlichen Abbildungen. Wirklichkeitsnahe Darstellungen sind fiktive Darstellungen, die objektiv wie tatsächliche Darstellungen aussehen, bei denen aber nicht ausgeschlossen werden kann, dass es sich um fiktive Darstellungen handelt. Dies dient primär der Beseitigung von Beweisschwierigkeiten [Bu13][Bu97]. Mit zunehmender Verbesserung von bild- und videogenerierender KI werden fiktive Inhalte immer häufiger als wirklichkeitsnah einzustufen sein. Relevant ist die Abgrenzung, weil bei wirklichkeitsnahen Darstellungen schon die Weitergabe an einzelne Personen unter Strafe steht, sowie bei kinderpornographischen Inhalten ihr Abruf und Besitz (§ 184 Abs. 1 Satz 1 Nr. 2, Abs. 3;

§ 184c Abs. 1 Nr. 2 StGB). Aus dem Gesetz selbst ergibt sich nicht, auf wessen Wahrnehmung es bei der „täuschenden Echtheit“ ankommt. Stellt man dieselben Erwägungen an, wie etwa bei der Beurteilung des Alters fiktiver Personen, so wäre auf den objektiven, gewissenhaft urteilenden Betrachter abzustellen (so wohl auch der BGH [Bu97a]). Bild- und videogenerierende KI sind jedoch mittlerweile so weit entwickelt, dass deren Output mit bloßem Auge, wie ihn ein objektiver Betrachter wohl wahrnehmen würde, kaum von tatsächlichen Inhalten zu unterscheiden ist, mit der Folge, dass KI-generiertes CSAM im Regelfall unter „wirklichkeitsnah“ fallen würde. Die Beurteilung eines fiktiven Inhalts als „wirklichkeitsnah“ sollte aber eine Ausnahme bleiben, weil damit fiktive Inhalte in Tatbestandsalternativen und Strafrahmen tatsächlichen Inhalten nahezu gleichgesetzt werden, ohne dass notwendigerweise ein Opfer vorliegt, dessen sexuelle Selbstbestimmung oder allgemeines Persönlichkeitsrecht durch Herstellung oder Verbreitung verletzt wurde. Die Gesetzesbegründung geht bei „wirklichkeitsnahen“ Darstellungen auch nicht von einem höheren Unrechtsgehalt aus als bei sonstigen fiktiven Darstellungen. Stattdessen sollen hierdurch primär Beweisschwierigkeiten behoben werden, die sich durch die technische Kreation von „Scheinwirklichkeiten“ ergeben [Bu97]. Es liegt daher näher, auf die Wahrnehmung und insbesondere die technischen Analysemöglichkeiten der Ermittlungsbehörden abzustellen. In diesem Fall wäre auch zu berücksichtigen, dass sich rein KI-generierte Inhalte auf Regelmäßigkeiten überprüfen lassen, um Rückschlüsse auf einen nicht-tatsächlichen Ursprung herzustellen [Ep23]. Hier ist eine gesetzgeberische Klarstellung wünschenswert.

Handelt es sich letztlich um eine tatsächliche Darstellung, so ist auch die Herstellung (ohne Verbreitungsabsicht) von kinder- und jugendpornographischen Inhalten strafbar, sowie auch bei jugendpornographischen Inhalten der Abruf und Besitz (wenn nicht die Ausnahme aus § 184c Abs. 4 StGB greift). Hier stellt sich insbesondere die Frage, wann ein Inhalt noch ein tatsächliches Geschehen wiedergibt, wenn es mithilfe von KI manipuliert wurde.

Pornographische Inhalte mit volljährigen Darstellern, die ohne Beteiligung eines Kindes hergestellt und mithilfe von KI „verjüngt“ wurden, stellen keine tatsächlichen Geschehen dar. Dem strengen und allumfassenden Verbot tatsächlicher kinderpornographischer Darstellungen liegt die Annahme zugrunde, dass zur Herstellung eines solchen Inhalts oftmals ein sexueller Missbrauch eines Kindes stattgefunden hat [Bu70]. Je nach „Glaubwürdigkeit“ der Bearbeitung kann ein solcher Inhalt allerdings als wirklichkeitsnah gesehen werden.

Wird ein tatsächlicher kinderpornographischer Inhalt als Ausgangsmaterial verwendet und mithilfe von KI bearbeitet, kommt es bei der Frage, ob hier eine Herstellung eines tatsächlichen CSAM vorliegt, darauf an, ob das im Ausgangsmaterial wiedergegebene tatsächliche Geschehen nach der Bearbeitung noch als solches erkennbar ist. Denn würde ein tatsächlicher kinderpornographischer Inhalt etwa soweit editiert, dass überhaupt kein pornographischer Inhalt mehr zu sehen wäre, kann das Endergebnis kein CSAM mehr darstellen, auch wenn bei der Herstellung des Ursprungsmaterials tatsächlich ein sexueller

Missbrauch stattgefunden hat. Eine Schutzlücke entsteht hierdurch aber nicht, da für die Bearbeitung in der Regel zuvor eine Besitzverschaffung oder ein Abruf des Ausgangsmaterials stattgefunden hat, was seinerseits strafbar ist. Die Bearbeitung eines Inhalts für sich kann noch nicht ausschlaggebend für die Annahme eines fiktiven Inhalts sein, da sonst die Umgehung des Strafmaßes oder der Strafbarkeit, etwa durch Spiegelung des Bildes, ohne Schwierigkeiten möglich wäre. Der Wortlaut des § 184b StGB stellt auch nicht auf eine originalgetreue Wiedergabe eines Inhalts ab, sondern auf die Wiedergabe eines tatsächlichen Geschehens. Solange erkennbar ist, dass das wiedergegebene Geschehen tatsächlich stattgefunden hat, handelt es sich um einen tatsächlichen Inhalt.

4.2 Ausnahmen zu Präventionszwecken

§ 184b Abs. 5 und 6 StGB (ggf. i.V.m. § 184c Abs. 6 StGB) sehen Ausnahmen für den Umgang mit kinder- und jugendpornographischen Inhalten zur Erfüllung staatlicher Aufgaben oder dienstlicher/beruflicher Pflichten sowie zur strafrechtlichen Ermittlung vor. § 184b Abs. 5 StGB erlaubt den Besitz und die Weitergabe tatsächlicher und wirklichkeitsnaher Inhalte an einzelne Personen (bei fiktiven Inhalten sind diese Handlungen ohnehin nicht strafbar), sofern dies im Rahmen der Erfüllung staatlicher Aufgaben oder dienstlicher oder beruflicher Pflichten geschieht. Dies betrifft insbesondere Angehörige von Polizei- und Justizbehörden ebenso wie Anwälte und Sachverständige in Bezug auf die Auswertung solcher Inhalte im Strafverfahren, aber auch Forschende im Rahmen eines konkreten wissenschaftlichen Forschungsauftrags sowie den behandelnden Arzt eines zur Herstellung eines Inhalts missbrauchten Kindes oder Jugendlichen [Bu12]. § 184b Abs. 6 StGB enthält eine Ausnahme für insb. die Herstellung in Verbreitungsabsicht und Verbreitung von kinder- bzw. jugendpornographischen Inhalten zu Ermittlungszwecken, wenn die Aufklärung des Sachverhalts auf andere Weise jedenfalls wesentlich erschwert wäre. Erfasst sind fiktive und wirklichkeitsnahe Inhalte, sofern bei der Herstellung dessen keine (tatsächliche) Bildaufnahme eines Kindes oder Jugendlichen verwendet wurde. Die Gesetzesbegründung erwähnt hier auch ausdrücklich die Herstellung von Inhalten mithilfe von KI sowie deren Verwendung zu Ermittlungszwecken [Bu19a]. Das Verwenden von kinder- oder jugendpornographischen Inhalten zum Training einer KI wird durch diese Ausnahme jedoch nicht gedeckt [Bu19a]. Die Herstellung und Verbreitung ist für Ermittlungsbeamte deswegen erforderlich, weil sie häufig eine sog. "Keuschheitsprobe" ablegen müssen, d.h. strafrechtlich relevante Inhalte verbreiten müssen, um Zugang zu Foren zu erlangen oder das Vertrauen potenzieller Täter zu gewinnen [Bu19].

4.3 Zivilrecht

Neben strafrechtlichen Sanktionen aus den §§ 184b und 184c StGB stehen der dargestellten Person zivilrechtliche Ansprüche gegen den Täter zu. Denn die Verbreitung entsprechender Inhalte verletzt regelmäßig das allgemeine Persönlichkeitsrecht des

dargestellten Kindes oder Jugendlichen. Dieses aus Art. 2 I, Art. 1 I des Grundgesetzes („GG“) abgeleitete Recht schützt insbesondere die persönliche Ehre und soll die Grundrechtsträger vor herabsetzenden Darstellungen bewahren [Lo20]. Es steht nach der Rechtsprechung des Bundesverfassungsgerichts auch Kindern zu, die wegen ihrer erhöhten Schutzbedürftigkeit sogar noch umfassender geschützt werden als Erwachsene [Bu03]. Die Verbreitung manipulierter [Bu05] und/oder pornographischer [KR22] Inhalte sowie die Verbreitung von Nacktbildern [Bu15a] beeinträchtigt ebendieses Recht in erheblicher Weise. Die dargestellten Kinder und Jugendlichen haben dann einen Anspruch auf Unterlassen nach § 1004 Abs. 1 des Bürgerlichen Gesetzbuchs („BGB“) analog i.V.m. § 823 Abs. 1 BGB i.V.m. Art. 2 Abs. 1, Art. 1 Abs. 1 GG sowie auf Ersatz des entstandenen Schadens, den sie – vertreten durch ihre Eltern, §§ 1626, 1629 BGB – gegenüber dem Täter geltend machen können [Ke19]. In praktischer Hinsicht besteht aber oft die Schwierigkeit, dass die dargestellten Kinder und Jugendlichen einerseits nichts von dem Bild- und Videomaterial erfahren und andererseits die Täter für sie nur schwer zu identifizieren sind.

4.4 Urheberrecht und KUG

Die Generierung oder Bearbeitung von Werken (z.B. Bildern oder Videos) durch KI hat auch urheberrechtliche Relevanz. Werden etwa Bilder oder Videos einer Person bearbeitet, um daraus CSAM (z.B. Deep Nudes oder Deepfakes) herzustellen, kann der Rechteinhaber einen Anspruch auf Unterlassen und Beseitigung geltend machen sowie Schadensersatz verlangen (§ 97 des Urheberrechtsgesetzes („UrhG“)). Urheber ist nach § 7 UrhG der Schöpfer des Werks. Das kann die abgebildete Person selbst sein, sofern sie das Bild oder Video selbst aufgenommen hat, aber auch Dritte. Die Abbildung auf einem Bild oder Video begründet noch kein Urheberrecht. Hier kommt gegebenenfalls ein Leistungsschutzrecht des ausübenden Künstlers nach § 73 ff. UrhG in Betracht, welches ebenfalls zu einem Anspruch auf Unterlassen, Beseitigung und Schadensersatz nach § 97 UrhG berechtigt. Ein Verstoß gegen urheberrechtliche Vorschriften kann zudem strafrechtliche Folgen nach sich ziehen (§ 106 UrhG) [Ho20].

Des Weiteren wird bereits für das Training der KI Bildmaterial von Kindern und Jugendlichen verwendet. KI-Systeme wird mit großen Datenmengen trainiert, unter denen sich auch urheberrechtlich geschützte Werke befinden können. Möglich ist beispielsweise, dass Bildaufnahmen oder Videos, die Kinder oder Jugendliche selbst veröffentlicht haben, oder die Eltern von ihren Kindern ins Netz gestellt haben, als KI-Trainingsdaten verwendet werden. Die Urheber dieser Inhalte (§§ 7 ff. UrhG) - teilweise die Kinder und Jugendlichen selbst, teilweise ihre Eltern oder andere Personen - können dann mitunter urheberrechtliche Ansprüche geltend machen. Erforderlich für einen etwaigen Anspruch der Urheber ist eine Verletzung des Urheberrechts, vgl. § 97 UrhG. Hierzu wird vertreten, dass dies bereits im Training der KI mit urheberrechtlich geschützten Werken bestehen

kann, da hierin eine Vervielfältigung im Sinne des § 16 UrhG zu sehen sei [Ma23; Ba23]. In Bezug auf das hergestellte Ergebnis, also den „Output“ der KI, muss zwischen den verschiedenen technischen Möglichkeiten differenziert werden: Besteht das Endergebnis zumindest teilweise aus einem bereits vorhandenen urheberrechtlich geschützten Werk, wie es beispielsweise bei Deepfakes oder Inpainting regelmäßig der Fall sein kann, ist auch hierin eine Verletzung des Urheberrechts zu sehen, wenn das Ursprungswerk urheberrechtlich geschützt ist [Ba23].

Weitere mögliche zivilrechtliche Ansprüche der dargestellten Kinder und Jugendlichen, können sich aus dem Kunsturhebergesetz („KUG“) ergeben; ein etwaiger Verstoß kann aber auch strafrechtlich verfolgt werden, § 33 KUG. Denn da auch Fotomontagen unter den Bildnisbegriff fallen [GP21], ist zumindest nicht ausgeschlossen, dass die generierten Inhalte im Einzelfall auch gegen § 22 KUG verstoßen. Erforderlich ist aber, dass die dargestellten Kinder und Jugendlichen durch die Nennung ihres Namens oder aufgrund ihrer körperlichen Merkmale (Gesichtszüge, Körperbau) erkennbar sind [DS22; GP21]. Das KUG schützt dann ebenfalls vor entsprechenden Darstellungen. Insbesondere besteht ein Schutz vor der Verbreitung von Nacktbildern in Form von Deepfakes [La20].

4.5 Datenschutzrecht

Auch die Datenschutz-Grundverordnung („DSGVO“) sieht Regelungen vor, die in verschiedenen Stadien und Konstellationen der Herstellung KI-generierter kinderpornographischer Inhalte greifen können. Keine Anwendung findet die DSGVO jedoch auf Verarbeitungsvorgänge, die ausschließlich im Rahmen persönlicher Tätigkeiten erfolgen (Art. 2 Abs. 2 lit. c DSGVO). Sie findet daher keine Anwendung auf Personen, die KI-generierte CSAM, insbesondere Deepfakes oder Inpaintings, nur zu privaten Zwecken herstellen, speichern oder hochladen [Ho20; KR22]. Eine Verarbeitung zu privaten Zwecken besteht allerdings nicht mehr, sobald Daten öffentlich zugänglich gemacht werden, weil die Verarbeitung dann nicht mehr ausschließlich in der persönlichen Sphäre des Verantwortlichen vorgenommen wird [Eu14; Eu03].

Darüber hinaus findet die DSGVO jedenfalls Anwendung auf die Anbieter von Diensten (z.B. Anbieter von KI-Diensten), die zur Herstellung verwendet werden oder Betreiber von Plattformen, auf die CSAM hochgeladen wird, sofern sie Verantwortliche i.S.d. Art. 4 Nr. 7 DSGVO sind.

Bereits das Erheben und Speichern von Daten zum Zwecke des KI-Trainings kann einen Verstoß gegen die DSGVO darstellen, wenn sich in dem Datensatz personenbezogene Daten befinden und keine entsprechende Rechtsgrundlage vorliegt [Di24]. Gleiches gilt für das KI-Training selbst [Fr23].

Ob der KI-Output, d.h. mit Hilfe von KI hergestelltes CSAM, ein personenbezogenes Datum ist und damit der DSGVO unterfällt, hängt davon ab, ob darauf eine reale, natürliche Person ohne unverhältnismäßigen Aufwand an Zeit, Kosten und Arbeitskraft identifiziert werden kann (Art. 4 Nr. 1, ErwGr. 26 DSGVO) [Eu16]. Identifikator ist hier insbesondere das Gesicht (z.B. bei Deepfakes [KR22], wenn ein Gesicht originalgetreu auf eine andere Person gesetzt wird, oder DeepNudes/Inpainting, wenn nur der Rest des Körpers durch KI verändert wurde). Es können aber auch abstraktere Merkmale wie spezielle körperliche Eigenschaften oder gewisse Kleidung in Betracht kommen, wenn sie so signifikant einer natürlichen Person zuzuordnen sind, dass diese ohne unverhältnismäßigen Aufwand eindeutig identifiziert werden kann. Hierbei wird in Zukunft auch die Verfügbarkeit von KI-gestützten Online-Bildsuchediensten (z.B. "AI Lens") eine immer größere Rolle spielen.

Gegen eine unrechtmäßige Verarbeitung personenbezogener Daten besteht insbesondere ein Recht auf Löschung (Art. 17 DSGVO) und ein Anspruch auf Schadensersatz (Art. 82 DSGVO) gegenüber dem oder den Verantwortlichen. Darüber hinaus kann eine unrechtmäßige Verarbeitung von der zuständigen Datenschutzbehörde mit einem Bußgeld geahndet werden (Art. 83 DSGVO).

5 Fazit

Unsere Arbeit zeigt, dass die aktuelle technische Entwicklung bezogen auf CSAM zwar rechtlich keine gänzlich neuen Sachverhalte schafft, sich aber dennoch Klärungsbedarf auf rechtlicher und politischer Ebene ergibt. Durch die rasanten Fortschritte in der KI-Technologie verschwimmen die Grenzen zwischen realen und künstlich generierten Inhalten, was dazu führen kann, dass sich die als Ausnahme gedachte Strafbarkeit des Besitzes oder der vereinzelt Weitergabe „wirklichkeitsnaher“ Darstellung zum neuen Regelfall für nicht-tatsächliche Inhalte entwickelt, wenngleich sie nur eine ausnahmsweise Gleichstellung mit tatsächlichen Inhalten darstellen soll. Dies macht eine präzise Definition des Begriffs "wirklichkeitsnah" erforderlich, damit die Strafbarkeit und das Strafmaß verhältnismäßig zum Unrechtsgehalt der jeweiligen Handlung stehen.

Zudem bleibt abzuwarten, wie der Gesetzgeber mit der massenhaften Erzeugung fiktiven CSAMs ohne die Absicht zur Verbreitung umgehen soll, was bislang nicht strafbar ist. Hierbei stehen sowohl ethische als auch rechtliche Überlegungen im Vordergrund. Die Technologie bietet die Möglichkeit, hochrealistische Inhalte mit geringem Aufwand zu erstellen, birgt jedoch auch das Potenzial, missbraucht zu werden.

Insgesamt steht die Gesellschaft vor der Herausforderung, die Auswirkungen dieser Technologie zu erkennen und angemessene Regelungen zu entwickeln, die sowohl den Schutz von Kindern als auch die Nutzung innovativer Technologien in Einklang bringen.

Hier ist enge Zusammenarbeit zwischen dem Gesetzgeber, Technologieexperten, Kinderrechtsorganisationen und Therapeuten unerlässlich.

Danksagungen

Diese Forschungsarbeit wurde unterstützt vom Nationalen Forschungszentrum für angewandte Cybersicherheit ATHENE. ATHENE wird gemeinsam vom Bundesministerium für Bildung und Forschung (BMBF) und vom Hessischen Ministerium für Wissenschaft und Kunst (HMWK) gefördert. Dieser Beitrag wurde vom Topic “Engineering Secure Systems” der Helmholtz-Gemeinschaft mit Unterstützung der KASTEL Security Research Labs Karlsruhe gefördert.

Literaturverzeichnis

- [Ba23] Baumann, M.: Generative KI und Urheberrecht – Urheber und Anwender im Spannungsfeld. *Neue Juristische Wochenschrift*, S. 3673, 2023.
- [Bu01] BGH, Urt. v. 27.6.2001 – 1 StR 66/01.
- [Bu03] BVerfG, Beschl. v. 29.7.2003 – 1 BvR 1964/00.
- [Bu05] BVerfG, Beschl. V. 14.2.2005 – 1 BvR 240/04.
- [Bu07] BGH, Urt. v. 20.12.2007 – 4 StR 459/07.
- [Bu12] BT-Drs. 12/4883, S. 8 f.
- [Bu13] BGH, Beschl. v. 19.3.2013 – 1 StR 8/13.
- [Bu13a] BGH, Beschl. v. 22.11.2013 – 2 StR 459/13.
- [Bu14] BGH, Beschl. v. 3.12.2014 – 4 StR 342/14.
- [Bu15] BGH, Beschl. v. 19.8.2015 – 5 StR 275/15.
- [Bu15a] Busch, R.: Strafrechtlicher Schutz gegen Kinderpornographie und Missbrauch, *Neue Juristische Wochenschrift*, S. 977, 2015.
- [Bu18] Bundestagsfraktionen CDU/CSU und SPD, BT-Drs 18/2601, S. 30.
- [Bu19] Bundesregierung, BT-Drs. 19/13836, S. 15.
- [Bu19a] Bundestagsausschuss für Recht und Verbraucherschutz, BT-Drs. 19/16543, S. 11.
- [Bu19b] Bundestagsfraktionen CDU/CSU und SPD, BT-Drs. 19/23707, S. 41.
- [Bu20] BGH, Beschl. v. 1.9.2020 – 3 StR 275/20.
- [Bu20a] BGH, Urt. v. 14.10.2020 – 1 StR 234/20.
- [Bu23] BGH, Urt. v. 14.12.2023 – 3 StR 183/23.

- [Bu70] Bundesregierung, BT-Drs. VI/1552, S. 35 f.
- [Bu97] Bundestagsausschuss für Bildung, Wissenschaft, Forschung, Technologie und Technikfolgenabschätzung, BT-Drs. 13/7934.
- [Bu9a] BGH, Urt. v. 17.12.1997 – 3 StR 567-97.
- [Co24] Council of the European Union, 2022/0155(COD), 8579/24, <https://www.patrick-breyer.de/wp-content/uploads/2024/04/2024-04-11-conseil-csam-compromis-9-avril.pdf>, 2024.
- [Da24] Das, S.: Sex offender banned from using AI tools in landmark UK case, *The Guardian*, 21.4.2024.
- [Di24] Diecker, A.: Datenschutzrechtliche Zulässigkeit der Trainingsdatensammlung. Wie Scraping und Crawling zur KI-Entwicklung eingesetzt werden können. *Zeitschrift für Datenschutz*, S. 132, 2024.
- [DS22] Dreier, T. et al.: Urheberrechtsgesetz, Urheberrechts-Diensteanbieter-Gesetz, Verwertungsgesellschaftengesetz, Nebenurheberrecht, Kunsturhebergesetz Kommentar. C.H. Beck, München, 2022.
- [Ep23] Epstein, D. et al.: Online Detection of AI-Generated Images. *IEEE/CVF International Conference on Computer Vision Workshops (ICCVW)*, 2023.
- [Eu03] EuGH, Urt. v. 6.11.2003 – C101/01.
- [Eu14] EuGH, Urt. v. 11.12.2014 – C-212/13.
- [Eu16] EuGH, Urt. v. 19.10.2016 – C-582/14.
- [Eu22] European Commission, Art. 7 lit. b des Vorschlags für eine Richtlinie [...] zur Bekämpfung von Gewalt gegen Frauen und häuslicher Gewalt, COM(2022) 105 final, 2022.
- [Eu24] European Parliament, P9_TA(2024)0138, <https://data.consilium.europa.eu/doc/document/ST-7536-2024-INIT/DE/pdf>, 2024.
- [Fl22] Flynn, A. et al.: Deepfakes and digitally altered imagery abuse: A cross-country exploration of an emerging form of image-based sexual abuse. *The British Journal of Criminology*, 62(6):1341–1358, 2022.
- [Fr23] Franke, L.: Datenschutzrechtskonformes Training von KI-Systemen mit öffentlich verfügbaren personenbezogenen Daten. *Recht Digital*, S. 565, 2023.
- [GP21] Gersdorf, H.; Paal, B. Hrsg.): *BeckOK Informations- und Medienrecht*, 2021.
- [Gr21] Graaf, J. W.: *Augmented reality: too good to be true?*, 2021.
- [He23] Hedgecoe, G.: Ai-generated naked child images shock spanish town of almendralejo. *BBC News*, 2023.
- [Ho20] Hoche, A.: Hat Obama Präsident Trump wirklich als Vollidiot bezeichnet? *IP-Rechtsberater*, S. 240–244, 2020.
- [In23] IWF. *How ai is being abused to create child sexual abuse imagery*, 2023.

- [Ke19] Keye, J.: Persönlichkeitsrechtsgefährdung durch Deep Fakes. IT-Rechtsberater, S. 290–291, 2019.
- [Ke23] Kempen, A.: The ugly side of creating life-like child abuse images. Servamus Community-based Safety and Security Magazine, 116(10):24–25, 2023.
- [KN23] Papathanasiou, in: Kindhäuser/Neumann/Paeffgen/Saliger, § 184b StGB, Rn. 4, 2023.
- [KR21] Kerner, C. and Risse, M.: Beyond porn and discreditation: Epistemic promises and perils of deepfake technology in digital lifeworlds. *Moral Philosophy and Politics*, 8(1):81–108, 2021.
- [KR22] Kumkar, L. K.; Rapp, J. P.: Deepfakes. Eine Herausforderung für die Rechtsordnung. *Zeitschrift für Digitalisierung und Recht*, S. 199, 2022.
- [La20] Lantwin, T.: Strafrechtliche Bekämpfung missbräuchlicher Deep Fakes, *MultiMedia und Recht*, S. 78, 2020.
- [LM13] Lester, D.; McSwain, S.; Gunn III, J. F.: Suicide and the internet: the case of amanda todd. *International journal of emergency mental health*, 15(3):179–180, 2013.
- [Lo20] Lothar, M., *Grundrechte*, 7. Auflage, 2020.
- [Ma23] Maamar, N.: Urheberrechtliche Fragen beim Einsatz von generativen KI-Systemen. *Zeitschrift für Urheber- und Medienrecht*, S. 481, 2023.
- [Ra21] Ratner, C.: When “sweetie” is not so sweet: artificial intelligence and its implications for child pornography. *Family Court Review*, 59(2):386–401, 2021.
- [St20] Stănilă, L.: Memories of the future-sweetie and the impact of the new technologies on the criminal justice system. *EU and comparative law issues and challenges series (ECLIC)*, 4:557–575, 2020.
- [St24] Steinebach, M.: KI-generierte Abbildungen von Kindesmissbrauch: Entwicklungen, Gegenmaßnahmen und Probleme. *Datenschutz und Datensicherheit-DuD*, 48(5), 304–309, 2024.
- [TS23] Thiel, D.; Stroebel, M.; Portnoff, R.: *Generative ml and csam: Implications and mitigations*, 2023.
- [Tu23] Tutten, J.: Florida students accused of creating nude photos of classmates using ai. *Yahoo News*, Dec 2023.
- [Vi23] Vigliarolo, B.: Csam found in large ai image generator-training dataset. *The Register*, Dec 2023.
- [We19] Westerlund, M.: The emergence of deepfake technology: A review. *Technology innovation management review*, 9(11), 2019.
- [WS19] Wittmer, S.; Steinebach, M.: Computergenerierte Kinderpornografie zu Ermittlungszwecken im Darknet. *Rechtliche Rahmenbedingungen und technische Umsetzbarkeit*, *MMR*, S. 650–653, 2019.
- [YC20] Yeh, C. et al.: Disrupting image-translation-based deepfake algorithms with adversarial attacks. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV) Workshops*, March 2020.