

Gesetzentwurf

des Bundesrates

Entwurf eines Gesetzes zum strafrechtlichen Schutz von Persönlichkeitsrechten vor Deepfakes

A. Problem und Ziel

Realistisch wirkende Medieninhalte, die mit Hilfe Künstlicher Intelligenz (KI) erzeugt oder verändert worden sind, sogenannte Deepfakes, sind seit einigen Jahren auf dem Vormarsch und beginnen unser Verhältnis zur medial wahrgenommenen Realität zu verändern. Nie zuvor war es derart einfach, qualitativ hochwertige Fälschungen von Ton-, Bild- und Videoaufnahmen zu erstellen.

Die mit Deepfakes einhergehenden technischen Manipulationen derartiger Medieninhalte bergen erhebliche Gefahren sowohl für individuelle Persönlichkeitsrechte und Vermögenswerte als auch für den demokratischen Willensbildungsprozess. Durch die dynamische Entwicklung der Technik im Zeitalter (generativer) Künstlicher Intelligenz, die beständige Erweiterung der Anwendungsmöglichkeiten, die einfache und günstige Verfügbarkeit entsprechender Anwendungen und die inzwischen hohe Authentizität stellen Deepfakes eine für den Staat und seine Bürgerinnen und Bürger besonders gefährliche Form der Informationsmanipulation dar. Diese wird zunehmend auch von Straftäterinnen und Straftätern zum Erreichen ihrer missbräuchlichen Ziele eingesetzt.

Als solche Ziele kommen namentlich der Einsatz von Deepfakes zu Desinformationszwecken und zu Zwecken rechtswidriger Bereicherung in Betracht. Deepfakes schaffen aber vor allem auch Gefährdungen für den Schutz der Persönlichkeit. Diese Gefährdungen bilden Anlass und Gegenstand des vorliegenden Entwurfs. Beispielhaft hierfür stehen Fallgestaltungen, in denen Gesichter oder andere Körperteile in Videos ausgetauscht, Mimik und Gestik gezielt gesteuert oder Stimmen nachgeahmt werden und hierbei der Anschein einer authentischen Wiedergabe erweckt wird. Nach bisherigen Erkenntnissen handelt es sich häufig um Fälle, in denen Frauen und Mädchen durch technische Manipulation von Bild- oder Videoaufnahmen in einen zuvor nicht bestehenden und von den Betroffenen offensichtlich nicht gewollten sexuellen Kontext gesetzt werden (sogenannte Deepnudes). Den Tätern kommt es dabei oftmals darauf an, diese Personen zum Objekt eigener sexueller Interessen zu machen oder an ihnen ihre Rache- und Machtbedürfnisse auszuleben. Für die Betroffenen unterscheiden sich die schädlichen Auswirkungen kaum von denjenigen, in denen reale Nacktaufnahmen von ihnen unbefugt verbreitet werden. Beispielhaft zu nennen sind auch Fallgestaltungen, in denen die Täterinnen und Täter Deepfakes einsetzen, um damit andere Personen im politischen Meinungsstreit möglichst effektiv zu diskreditieren

und nachteilige Stimmungen gegen sie zu schüren, sowie „Schockanrufe“ mittels KI-generierter Stimmen enger Angehöriger. Beides ist ebenfalls persönlichkeitsrelevant und auch -verletzend.

Besondere strafrechtliche Regelungen gegen derartige missbräuchliche Deepfakes gibt es bislang nicht. Es existieren zwar zahlreiche strafrechtliche Regelungen, die im Zusammenhang mit persönlichkeitsrechtsverletzenden Deepfakes im Einzelfall einschlägig sein können. Diese Vorschriften erfassen das Phänomen aber jeweils – wenn überhaupt – nur in Teilaspekten und auch nicht in seinem Unrechtskern.

Das allgemeine Persönlichkeitsrecht schützt auch und gerade vor der Verbreitung eines technisch manipulierten oder künstlich hergestellten Medieninhalts, der den Anschein erweckt, eine authentische Video-, Bild- oder Tonaufnahme des äußeren Erscheinungsbilds, des Verhaltens oder der Stimme einer Person zu sein. Die betroffene Person hat ein berechtigtes Interesse, ohne ihre Zustimmung nicht in eine künstlich erzeugte, aber scheinbar authentische „Wirklichkeit“ hineingestellt zu werden mit Äußerungen, die sie selbst nicht getätigt hat, oder mit Handlungen, die sie selbst nicht vorgenommen hat. Durch eine nicht offengelegte technische Manipulation entsprechender Medieninhalte wird eine unzutreffende Tatsachenbehauptung über die Realität der betroffenen Person aufgestellt. Diese sieht sich in einer Situation, in der sie die Selbstbestimmung und die Kontrolle über das eigene Erscheinungsbild und Auftreten verliert und ihre identitätsprägenden Merkmale mit Außenwirkung ge- oder verfälscht oder in einen falschen Kontext gestellt werden. Bloße Kennzeichnungspflichten für Deepfakes werden dem persönlichkeitsrechtsverletzenden Charakter solcher Medieninhalte und den vielfach gravierenden Auswirkungen für die Betroffenen im Fall ihrer Verbreitung nicht gerecht. Auf den Einsatz des Strafrechts kann nicht verzichtet werden.

B. Lösung

Der Entwurf sieht daher eine spezifisch auf Deepfakes und vergleichbare technische Manipulationen zugeschnittene Vorschrift zum Persönlichkeitsschutz im Strafgesetzbuch vor. Wegen Verletzung von Persönlichkeitsrechten durch digitale Fälschung soll zukünftig bestraft werden, wer das Persönlichkeitsrecht einer anderen Person verletzt, indem er einen mit computertechnischen Mitteln hergestellten oder veränderten Medieninhalt, der den Anschein einer wirklichkeitsgetreuen Bild- oder Tonaufnahme des äußeren Erscheinungsbildes, des Verhaltens oder mündlicher Äußerungen dieser Person erweckt, einer dritten Person zugänglich macht. Darüber hinaus enthält der Entwurf Regelungen zum Schutz verstorbener Personen, zur Strafschärfung bei Vorliegen unrechtserhöhender Umstände und zur Strafflosigkeit sozialadäquater Handlungen sowie Folgeänderungen im Strafantragsrecht und in der Strafprozessordnung.

C. Alternativen

Beibehaltung des bisherigen, unbefriedigenden Zustands.

D. Haushaltsausgaben ohne Erfüllungsaufwand

Keine.

E. Erfüllungsaufwand

E.1 Erfüllungsaufwand für Bürgerinnen und Bürger

Keiner.

E.2 Erfüllungsaufwand für die Wirtschaft

Keiner.

E.3 Erfüllungsaufwand der Verwaltung

Keiner.

F. Weitere Kosten

Durch die Einführung der neuen strafrechtlichen Regelung können den Länderhaushalten Verfahrens- und Vollzugskosten in überschaubarem Umfang entstehen, deren Höhe sich nicht näher beziffern lässt.

BUNDESREPUBLIK DEUTSCHLAND
DER BUNDESKANZLER



Berlin, 21. August 2024

An die
Präsidentin des
Deutschen Bundestages
Frau Bärbel Bas
Platz der Republik 1
11011 Berlin

Sehr geehrte Frau Präsidentin,

hiermit übersende ich gemäß Artikel 76 Absatz 3 des Grundgesetzes den vom Bundesrat in seiner 1046. Sitzung am 5. Juli 2024 beschlossenen

Entwurf eines Gesetzes zum strafrechtlichen Schutz von Persönlichkeitsrechten
vor Deepfakes

mit Begründung und Vorblatt (Anlage 1).

Ich bitte, die Beschlussfassung des Deutschen Bundestages herbeizuführen.

Federführend ist das Bundesministerium der Justiz.

Die Auffassung der Bundesregierung zu dem Gesetzentwurf ist in der als Anlage 2 beigefügten
Stellungnahme dargelegt.

Mit freundlichen Grüßen

Olaf Scholz

Anlage 1

Entwurf eines Gesetzes zum strafrechtlichen Schutz von Persönlichkeitsrechten vor Deepfakes

Vom ...

Der Bundestag hat das folgende Gesetz beschlossen:

Artikel 1**Änderung des Strafgesetzbuches**

Das Strafgesetzbuch in der Fassung der Bekanntmachung vom 13. November 1998 (BGBl. I S. 3322), das zuletzt durch Artikel 1 des Gesetzes vom 24. Juni 2024 (BGBl. 2024 I Nr. 213) geändert worden ist, wird wie folgt geändert:

1. In der Inhaltsübersicht wird nach der Angabe zu § 201a folgende Angabe eingefügt:
„§ 201b Verletzung von Persönlichkeitsrechten durch digitale Fälschung“.
2. Nach § 201a wird folgender § 201b eingefügt:

„§ 201b

Verletzung von Persönlichkeitsrechten durch digitale Fälschung

(1) Wer das Persönlichkeitsrecht einer anderen Person verletzt, indem er einen mit computertechnischen Mitteln hergestellten oder veränderten Medieninhalt, der den Anschein einer wirklichkeitsgetreuen Bild- oder Tonaufnahme des äußeren Erscheinungsbildes, des Verhaltens oder mündlicher Äußerungen dieser Person erweckt, einer dritten Person zugänglich macht, wird mit Freiheitsstrafe bis zu zwei Jahren oder mit Geldstrafe bestraft. Gleiches gilt, wenn sich die Tat nach Satz 1 auf eine verstorbene Person bezieht und deren Persönlichkeitsrecht dadurch schwerwiegend verletzt wird.

(2) Wer in den Fällen des Absatzes 1 Satz 1 den Medieninhalt der Öffentlichkeit zugänglich macht oder einen Medieninhalt zugänglich macht, der einen Vorgang des höchstpersönlichen Lebensbereichs zum Gegenstand hat, wird mit Freiheitsstrafe bis zu fünf Jahren oder mit Geldstrafe bestraft.

(3) Absatz 1 Satz 1, auch in Verbindung mit Absatz 2, gilt nicht für Handlungen, die in Wahrnehmung überwiegender berechtigter Interessen erfolgen, namentlich der Kunst oder der Wissenschaft, der Forschung oder der Lehre, der Berichterstattung über Vorgänge des Zeitgeschehens oder der Geschichte oder ähnlichen Zwecken dienen.

(4) Die Bild- oder Tonträger oder andere technische Mittel, die der Täter oder Teilnehmer verwendet hat, können eingezogen werden. § 74a ist anzuwenden.“

3. § 205 wird wie folgt geändert:
 - a) In Absatz 1 Satz 2 wird nach der Angabe „201a,“ die Angabe „201b,“ eingefügt.
 - b) In Absatz 2 Satz 4 wird nach der Angabe „Satz 2“ die Angabe „sowie § 201b Absatz 1 Satz 2“ eingefügt.

Artikel 2**Änderung der Strafprozessordnung**

Das Strafgesetzbuch in der Fassung der Bekanntmachung vom 13. November 1998 (BGBl. I S. 3322), das zuletzt durch Artikel 16 des Gesetzes vom 6. Mai 2024 (BGBl. 2024 I Nr. 149) geändert worden ist, wird wie folgt geändert:

1. In § 100k Absatz 2 Satz 1 Nummer 1 Buchstabe h wird nach der Angabe „201a,“ die Angabe „201b,“ eingefügt.
2. Nach § 374 Absatz 1 Nummer 2a wird die folgende Nummer 2b eingefügt:
„2b. eine Verletzung von Persönlichkeitsrechten durch digitale Fälschung (§ 201b des Strafgesetzbuches),“.

Artikel 3**Inkrafttreten**

Dieses Gesetz tritt am Tag nach der Verkündung in Kraft.

Begründung

A. Allgemeiner Teil

I. Zielsetzung und Notwendigkeit der Regelungen

Der Gesetzentwurf macht sich zur Aufgabe, den strafrechtlichen Schutz vor technisch manipulierten oder künstlich generierten Medieninhalten zu verbessern, die den Anschein erwecken, eine authentische Video-, Bild- oder Tonaufnahme des äußeren Erscheinungsbildes, des Verhaltens oder der Stimme einer Person zu sein. Er dient damit vor allem dem Schutz des allgemeinen Persönlichkeitsrechts der hiervon betroffenen Personen.

1. Realistisch wirkende Medieninhalte, die mit Hilfe Künstlicher Intelligenz (KI) erzeugt oder verändert worden sind, sog. Deepfakes, sind seit einigen Jahren auf dem Vormarsch und beginnen unser Verhältnis zur medial wahrgenommenen Realität zu verändern. Nie zuvor war es derart einfach, qualitativ hochwertige Fälschungen von Ton-, Bild- und Videoaufnahmen zu erstellen. So existiert bereits Software, die Sprachbefehle in erstaunlich realistisch wirkende Videos verwandeln kann. Eine Aufdeckung der Manipulation bzw. der künstlichen Herstellung dieser Inhalte ist ohne technische Hilfsmittel häufig kaum noch möglich. Die mit Deepfakes einhergehenden technischen Manipulationen von Audio-, Foto- oder Videodateien bergen erhebliche Gefahren sowohl für individuelle Persönlichkeitsrechte und Vermögenswerte als auch für den demokratischen Willensbildungsprozess. Durch die dynamische Entwicklung der Technik im Zeitalter (generativer) Künstlicher Intelligenz, die beständige Erweiterung der Anwendungsmöglichkeiten, die einfache und günstige Verfügbarkeit entsprechender Anwendungen und die inzwischen hohe Authentizität stellen Deepfakes eine für den Staat und seine Bürgerinnen und Bürger besonders gefährliche Form der Informationsmanipulation dar. Diese wird zunehmend auch von Straftäterinnen und Straftätern zum Erreichen ihrer missbräuchlichen Ziele eingesetzt. Es ist davon auszugehen, dass sich die benötigte Expertise und der notwendige Aufwand zur Erstellung von Fälschungen durch die Verbesserung und erhöhte Verfügbarkeit an öffentlich zugänglicher Software stetig verringern wird, so dass sich die Häufigkeit von Angriffen mittels dieser Technologie weiter signifikant erhöhen wird.
2. Auch wenn die missbräuchliche Verwendung von Deepfakes in der Praxis der Strafverfolgung bislang nur eine untergeordnete Rolle spielt, handelt es sich – gerade für den Bereich pornographischer Deepfakes – schon jetzt um ein praktisch relevantes Phänomen. Aus Medienberichten ergeben sich vielfältige Hinweise auf schädliche Einsatzformen, die im Lichte der vorbeschriebenen Entwicklung weiter an Bedeutung gewinnen dürften. Dabei lassen sich vor allem drei Zielrichtungen und Anwendungsbereiche identifizieren:
 - a) Der erste Bereich betrifft den Einsatz von Deepfake-Technologie zu Desinformationszwecken, namentlich zur Manipulation des demokratischen Willensbildungsprozesses. So ist es im Wege von Deepfakes möglich, glaubwürdige Desinformationskampagnen durchzuführen, indem manipulierte Medieninhalte von Personen des öffentlichen Lebens erzeugt und verbreitet werden. Hierzu aus dem Ausland berichtete Vorkommnisse betreffen häufig den Wahlkampf. Diese zeigen zugleich, dass Deepfakes neben dem Zweck der Fehlinformation und der Einflussnahme auf die politische Willensbildung auch dazu eingesetzt werden, um mit vergleichsweise geringem Aufwand politische Gegner möglichst effektiv zu diskreditieren oder lächerlich zu machen und nachteilige Stimmungen gegen sie zu schüren. Dies geschieht etwa dadurch, dass den Betroffenen Äußerungen in den Mund gelegt werden, die sie nie getätigt haben, bei denen aber ihre Stimme täuschend echt nachgeahmt wird.
 - b) Ein weiterer Bereich umfasst den Einsatz künstlich erstellten Video-, Bild- und vor allem Tonmaterials zur Verfolgung eigennütziger wirtschaftlicher Interessen, insbesondere zu Betrugs- oder Erpressungszwecken. Bekannt geworden sind insoweit zum einen Fallgestaltungen nach dem Muster der „Schockanrufe“ oder des „Enkeltricks“. So spiegeln Straftäter bei Dritten Anrufe von Angehörigen vor und arbeiten hierzu mit einer sog. Voice Cloning Software, die die Stimmen der Angehörigen täuschend echt imitiert. Die hierdurch bei den überrumpelten Angerufenen erzeugte Verwechslung, Verwirrung

oder Angst nutzen die Täterinnen und Täter, um sie zu Geldüberweisungen zu veranlassen, etwa für eine angeblich benötigte Kautions- oder ein Lösegeld. Eine andere Fallgestaltung betrifft Betrugstaten nach der Methode des „CEO/CFO-Fraud“. Bei dieser werden Unternehmen Opfer eines Betrugs, indem Mitarbeiter durch Anruf einer vermeintlichen Leitungsperson des Unternehmens mittels Verwendung einer Software zur Stimmenimitation oder sogar durch eine vorgespiegelte Videokonferenz zu einer Geldüberweisung auf ein tätereigenes Konto gebracht werden.

- c) Der dritte Bereich umfasst die Gruppe der Persönlichkeitsrechtsverletzungen durch Deepfakes. Beispielhaft hierfür stehen Fälle, in denen – auf der Basis entsprechender Bild-, Video- und/oder Audio-Originale – Gesichter oder andere Körperteile in Videos ausgetauscht, Mimik und Gestik gezielt gesteuert oder Stimmen nachgeahmt werden und hierbei der Anschein einer authentischen Wiedergabe erweckt wird. Die insoweit aus Deepfakes resultierenden Gefährdungen für die Persönlichkeit und die ungestörte Entwicklung von Kindern und Jugendlichen sind bislang nicht ausreichend in den Blick genommen worden. Sie bilden – auch aufgrund des unzureichenden strafrechtlichen Schutzes – den Anlass für den Gesetzentwurf.

Persönlichkeitsrelevant und auch -verletzend sind zum einen die bereits im Kontext des Betruges und der Erpressung genannten Fälle des „Voice Clonings“, also Fälle, in denen eine fremde Stimme zum Zweck der Täuschung oder Drohung und eigener rechtswidriger Bereicherung unbefugt nachgeahmt wird. Gleiches gilt für Deepfakes, die politische Gegner diskreditieren oder diffamieren sollen. Auch jenseits dessen finden sich manipulierte Medieninhalte, die den sozialen Geltungswert betroffener Personen beeinträchtigen, indem etwa ihre Gesichter auf fremde Körper in peinlichen Situationen gesetzt oder für sie verfängliche Äußerungen in den Mund gelegt werden.

Eine wichtige Gruppe von Persönlichkeitsverletzungen mit besonderem Schadenspotenzial betrifft die Fälle, in denen es den Tätern darauf ankommt, eine andere Person zum Objekt eigener sexueller Interessen zu machen oder an ihr Rache- und Machtbedürfnisse auszuleben. Diese Fallgruppe der Verletzung von Persönlichkeitsrechten ist bislang im Zusammenhang mit missbräuchlichen Deepfakes am häufigsten festzustellen. Über 90 Prozent der im Internet vorfindbaren Deepfakes sollen dabei den Bereich der Pornografie bzw. Nacktheit betreffen („Deepnudes“). Betroffen sind bislang fast ausschließlich Frauen. So wird das auf Original-Bildaufnahmen festgehaltene äußere Erscheinungsbild (Gesicht, Körper) von Ex-Partnerinnen, Schauspielerinnen, Popstars oder anderen Frauen per Computer in pornografische Filme oder Nacktaufnahmen eingefügt oder entsprechendes Material mittels KI generiert. Die betroffenen Frauen werden hierdurch in einen zuvor nicht bestehenden und von den Betroffenen offensichtlich nicht gewollten sexuellen Kontext gesetzt. Da die Täter in ihrem Handeln häufig von dem Bedürfnis nach Rache und Macht geleitet sind, werden die Aufnahmen vielfach im Internet veröffentlicht, z. B. auf Pornographie-Plattformen, und erzeugen so eine breitenwirksame Schädigung. Wenn die betroffenen Frauen Kenntnis von den Aufnahmen erlangen, ist der Schaden häufig genug schon längst eingetreten und Versuche, die Aufnahmen wieder aus dem Internet zu entfernen, ein schwieriges und in ihrem Erfolg höchst ungewisses Unterfangen. Für die Betroffenen unterscheiden sich die schädlichen Auswirkungen kaum von der Situation, in der reale Nacktaufnahmen von ihnen unbefugt verbreitet werden. Sie müssen damit rechnen, noch nach Jahren mit den manipulierten oder künstlich erstellten Medieninhalten in Verbindung gebracht und hiervon ausgehend als Person bewertet zu werden. Aufgrund der Digitalisierung und der damit einhergehenden Reproduzierbarkeit sind diese Inhalte der Verfügungsgewalt der Betroffenen in besonderer Weise entzogen. Sie können sich nie sicher sein, dass alle verfügbaren Inhalte gelöscht worden sind, also insbesondere im Internet nicht mehr verfügbar sind.

Eine besondere Schadensdimension haben hierbei Fälle, in denen die sexualbezogenen Deepfakes Kinder und Jugendliche zum Gegenstand haben. Dies zeigt sich an einem Fall aus Spanien, der im Herbst letzten Jahres medial bekannt geworden ist (<https://de.euronews.com/next/2023/09/20/gefalschte-kindernacktbilder-von-spanischen-jugendlichen-können-sie-dagegen-vorgehen>). Dort hatten mehr als 20 Mädchen Nacktfotos von sich auf ihren Mobiltelefonen erhalten. Die Bilder waren von ihren Instagram-Konten gestohlen, mit Hilfe generativer KI verändert und dann in WhatsApp-Gruppen geteilt worden. Die Kinder und Jugendlichen waren auf den echten Fotos vollständig bekleidet, aber die App ließ die Nacktheit vollkommen echt aussehen. Für die betroffenen Mädchen führte der Anblick der Bilder und die Ungewissheit über Art und Ausmaß der Verbreitung zu einer nachhaltigen psychischen Beeinträchtigung.

Der Fall zeigt beispielhaft, dass Deepfakes das Potenzial haben, erhebliche Schäden anzurichten, und dass sich mit wenigem und unverfänglichem Ausgangsmaterial unter Einsatz frei verfügbarer Software sehr schnell Material mit jedem beliebigen Inhalt herstellen und verbreiten lässt. Dies machen sich mittlerweile auch Hersteller und Konsumenten von kinderpornographischen Inhalten zunutze. Mittels Deepfake-Technologie können die Täter z. B. die von ihnen favorisierten Kinder in einen dem eigenen Fetisch entsprechenden Präferenzkontext bringen. Es genügen unverfängliche Bildaufnahmen und entsprechende Software zur Verarbeitung. Es ist daher auch nicht verwunderlich, dass sich Meldungen häufen, in denen von mittels generativer KI erzeugter Kinderpornographie berichtet wird.

3. Strafrechtliche Sonderregelungen im Kampf gegen missbräuchliche Deepfakes gibt es bislang nicht. Vielmehr gelten die allgemeinen Strafvorschriften. Der strafrechtliche Schutz der Persönlichkeitsinteressen vor entsprechenden Inhalten erweist sich dabei als unzureichend.
 - a) Dies zeigt sich bereits an den Fällen des „Voice Clonings“. Eine Strafbarkeit wegen Verletzung der Vertraulichkeit des Wortes gem. § 201 StGB kommt grundsätzlich nicht in Betracht (vgl. Fischer, StGB, 71. Aufl. 2024, § 201 Rn. 3). Schutzgut dieser Regelung ist das (nichtöffentlich) gesprochene Wort einer anderen Person. Durch „Voice Cloning“ generierte Äußerungen sind nicht Ausdruck einer tatsächlichen Äußerung der betroffenen Person. Vielmehr handelt es sich um völlig neue, computertechnisch hergestellte und die jeweilige Originalstimme des Sprechers imitierende Audiodateien. Da eine künstliche erzeugte („synthetische“) Stimme keine tatsächliche Äußerung der imitierten Person abbildet, entfällt insoweit richtigerweise auch eine Strafbarkeit nach dem Bundesdatenschutzgesetz. Die für die Erstellung der Fälschung verwendete tatsächliche Tonaufnahme hat zwar ein personenbezogenes Datum im Sinne des Datenschutzrechts zum Gegenstand, strafbar ist dessen Ver- und Bearbeitung aber nur dann, wenn es nicht allgemein zugänglich war und der Täter zugleich gegen Entgelt oder in Bereicherungs- oder Schädigungsabsicht gehandelt hat (vgl. § 42 Absatz 2 BDSG). Die unbefugte (digitale) Nachahmung der Stimme lässt sich – für sich gesehen – auch nicht als Beleidigungsdelikt erfassen.
 - b) Bei täuschend echt wirkenden Bild- und Videoaufnahmen bietet das geltende Recht zwar weitergehende Sanktionsmöglichkeiten. Aber auch diese erweisen sich bei näherer Betrachtung als nicht ausreichend und nicht zielgenau.
 - aa) Für die Verbreitung, den Erwerb und den Besitz kinder- und jugendpornographischer Deepfakes bietet das Strafrecht bereits weitreichende Sanktionsmöglichkeiten. So stellt das Gesetz in §§ 184b, 184c StGB auch den Umgang (Verbreitung, Erwerb, Besitz) mit solchen Inhalten unter Strafe, die ein wirklichkeitsnahes Geschehen wiedergeben, und erfasst damit auch die Fälle kinder- und jugendpornographischer Deepfakes. Allerdings trennt das Gesetz hierbei nicht klar zwischen dem traditionellen Schutz vor der Wahrnehmung pornographischer Inhalte und dem Schutz des Persönlichkeitsrechts der hiervon betroffenen Kinder und Jugendlichen. Die Regelungen in §§ 184b, 184c StGB dienen letztlich beiden Belangen. Anders ist dies jedoch mit Blick auf erwachsene Opfer. Deren Schutz bleibt defizitär. Denn der besondere persönlichkeitsverletzende Charakter pornographischer Deepfakes kommt in den Strafvorschriften der §§ 184, 184a StGB nicht hinreichend zum Ausdruck. So dient die Strafbarkeit der Verbreitung pornographischer Inhalte (§ 184 StGB) vor allem dem Jugendschutz sowie dem Schutz vor ungewollter Konfrontation mit Pornographie, nicht aber den Belangen der (tatsächlich oder scheinbar) abgebildeten Personen. Vergleichbares gilt für den Bereich der Gewalt- und Tierpornographie (§ 184a StGB).
 - bb) Schutzdefizite zeigen sich auch und gerade für den besonders bedeutsamen Bereich der mittels Deepfake-Technologie erzeugten (nicht pornographischen) Nacktaufnahmen und für die sonstigen Fälle, in denen ein reales Personenabbild mittels computertechnischer Manipulation in einen Kontext der Sexualität eingefügt wird.

Zwar geht die überwiegende Auffassung zutreffend davon aus, dass das unbefugte Verbreiten oder öffentliche Zurschaustellen eines solches Bildnisses unter erkennbarer Wiedergabe des äußeren Erscheinungsbildes einer anderen Person einen strafbaren Verstoß gegen das Kunsturhebergesetz darstellt (vgl. Lantwin, MMR 2020, 78, 79; Heckmann/Paschke, in: Stern/Sodan/Möstl, Das Staatsrecht der Bundesrepublik Deutschland im europäischen Staatenverbund, 2. Aufl. 2022, § 121

Rn. 74; a. A. etwa Hartmann, K&R 2020, 350, 353). Den besonderen Unrechtsgehalt derartiger Deepfakes – die Informationsmanipulation und Dekontextualisierung von Bildern – vermag dieser Tatbestand aber nicht zu erfassen.

Die Anwendbarkeit des Datenschutzstrafrechts (§ 42 Absatz 2 BDSG) auf derartige Bildmanipulationen ist fraglich und der Schutz jedenfalls unvollkommen. Ungeklärt ist schon die grundsätzliche Frage, ob Täter nur für die Datenverarbeitung Verantwortliche oder Auftragsverarbeiter im Sinne des Bundesdatenschutzgesetzes sein können oder ob es sich um ein sog. Jedermannsdelikt handelt. Eine als Grundlage verwendete authentische Bildaufnahme einer realen Person kann jedenfalls nur dann Gegenstand strafbarer Verarbeitung sein, wenn sie nicht öffentlich zugänglich war. Dies wird häufig nicht der Fall sein (vgl. den o.g. Fall aus Spanien) oder sich jedenfalls nicht nachweisen lassen. Erforderlich ist zudem stets, dass der Täter gegen Entgelt oder in Bereicherungs- oder Schädigungsabsicht gehandelt hat.

In Betracht kommt dann vor allem eine Strafbarkeit nach § 201a Absatz 2 Satz 1 StGB. Danach macht sich derjenige strafbar, der unbefugt von einer anderen Person eine Bildaufnahme, die geeignet ist, dem Ansehen der abgebildeten Person erheblich zu schaden, einer dritten Person zugänglich macht. Hier stellen sich zunächst zwei Fragen: Zum einen ist ungeklärt, ob eine Bildaufnahme in diesem Sinne auch dann vorliegt, wenn sie sich – wenn auch unter Verwendung realer Bilder als „Lernmaterial“ – als computergenierte Nachahmung der Wirklichkeit, also als „synthetische“ bzw. manipulierte Bildaufnahme, darstellt. Insoweit spricht – insbesondere mit Blick auf den Gesetzeswortlaut („Bildaufnahme“) – manches dafür, dass unter einer Bildaufnahme nur eine echte, fotografisch hergestellte Aufnahme, also die Aufnahme wahrer Ereignisse, zu verstehen ist (vgl. Greif, Strafbarkeit von bildbasierten, sexualisierten Belästigungen, 2023, S. 231 ff.; Dorerbeck, Cybermobbing, 2019, S. 182 ff.). Zum anderen ist fraglich, ob und unter welchen Umständen die Zurschaustellung von Nacktheit in einer täuschend echt wirkenden Bild- oder Videoaufnahme als erheblich ansehenschädigend zu bewerten ist (sehr zurückhaltend etwa Greif a. a. O. S. 235 ff. m.w.N.). Die Bezugnahme auf den Begriff der Ansehenschädigung wirft im Übrigen die ebenfalls ungeklärte Frage auf, ob daneben oder stattdessen eine Strafbarkeit wegen Verleumdung (§ 187 StGB) in Betracht kommt. Durch die Beschränkung der Strafbarkeit auf Bildaufnahmen können Fälle des „Voice Clonings“ durch § 201a StGB von vornherein nicht erfasst werden.

- c) Insgesamt ist die Rechtslage daher unklar und unübersichtlich. Es gibt zwar eine Vielzahl strafrechtlicher Regelungen, die im Zusammenhang mit persönlichkeitsrechtsverletzenden Deepfakes im Einzelfall einschlägig sein können (vgl. etwa Lantwin, MMR 2019, 574, 576; 2020, 78 ff.). Die Vorschriften erfassen das Phänomen aber jeweils – wenn überhaupt – nur in Teilaspekten und auch nicht in seinem Unrechtskern (dazu unten).

Die aktuelle Gesetzgebung der Europäischen Union wird daran absehbar wohl nichts ändern. So ist in der kommenden Richtlinie zur Bekämpfung von Gewalt gegen Frauen und häuslicher Gewalt zwar vorgesehen, dass die Mitgliedstaaten eine Strafvorschrift zur nicht-einvernehmlichen Weitergabe von intimem oder manipuliertem Material vorsehen müssen. Unter Strafe gestellt werden sollen u.a. die Herstellung, Manipulation oder Veränderung von Bildern, Videos oder ähnlichem Material, die den Anschein erwecken, dass eine andere Person eindeutige sexuelle Handlungen vornimmt, und deren anschließende Zugänglichmachung gegenüber der Öffentlichkeit mittels Informations- und Kommunikationstechnologien ohne Einwilligung der betreffenden Person. Diese Regelung zielt gerade auf die Verbreitung missbräuchlicher Deepfakes ab. Allerdings ist sie in ihrem Anwendungsbereich eng gefasst, da sie nur den Bereich sexualbezogener Bildaufnahmen betrifft und zusätzlich voraussetzt, dass die Handlung geeignet ist, der Person ernsthaften Schaden zuzufügen. Sie dürfte kaum über den Regelungsgehalt des § 201a Absatz 2 StGB hinausgehen, sofern man Deepfakes von dieser Regelung grundsätzlich als erfasst ansieht.

4. Vor dem Hintergrund der aufgezeigten Bedrohung der Persönlichkeitsinteressen durch Deepfakes und die Schwierigkeit deren strafrechtlicher Erfassung schlägt der Gesetzentwurf eine spezifisch auf Deepfakes und vergleichbare technische Manipulationen zugeschnittene Regelung zum Persönlichkeitsschutz im Strafgesetzbuch vor. Eine solche Regelung soll dazu dienen, das damit verbundene Unrecht zielgenau zu erfassen und prägnant zum Ausdruck zu bringen.

Ausgangspunkt der Überlegungen ist dabei, dass das allgemeine Persönlichkeitsrecht auch und gerade vor der Verbreitung eines technisch manipulierten oder künstlich hergestellten Medieninhalts schützt, der den Anschein erweckt, eine authentische Video-, Bild- oder Tonaufnahme des äußeren Erscheinungsbilds, des Verhaltens oder der Stimme einer Person zu sein (vgl. BVerfG, NJW 2005, 3271, 3272). Die betroffene Person hat ein berechtigtes Interesse, ohne ihre Zustimmung nicht in eine künstlich erzeugte, aber authentisch wirkende „Wirklichkeit“ hineingestellt zu werden mit Äußerungen, die sie selbst nicht getätigt hat, oder mit Handlungen, die sie selbst nicht vorgenommen hat. Das Bundesverfassungsgericht hat dies bereits im Jahr 2005 – bezogen auf die Verwendung einer Fotomontage in einer Zeitschrift – wie folgt ausgesprochen (NJW 2005, 3271, 3273):

„Das fotografische Abbild übermittelt ohne Verwendung von Worten Informationen über die abgelichtete Person. Fotos suggerieren Authentizität und die Betrachter gehen davon aus, dass die abgebildete Person in Wirklichkeit so aussieht. Diese Annahme aber trifft bei einer das Aussehen verändernden Bildmanipulation, wie sie heute relativ einfach mit technischen Mitteln herbeigeführt werden kann, nicht zu. Der Träger des Persönlichkeitsrechts hat zwar kein Recht darauf, von Dritten nur so wahrgenommen zu werden, wie er sich selbst gerne sehen möchte (vgl. BVerfGE 97, 125 [148f.] = NJW 1998, 1381; BVerfGE 97, 391 [403] = NJW 1998, 2889; st. Rspr.), wohl aber ein Recht, dass ein fotografisch erstelltes Abbild nicht manipulativ entstellt ist, wenn es Dritten ohne Einwilligung des Abgebildeten zugänglich gemacht wird. Die Bildaussage wird jedenfalls dann unzutreffend, wenn das Foto über rein reproduktionstechnisch bedingte und für den Aussagegehalt unbedeutende Veränderungen hinaus verändert wird. Solche Manipulationen berühren das Persönlichkeitsrecht, einerlei ob sie in guter oder in verletzender Absicht vorgenommen werden oder ob Betrachter die Veränderung als vorteilhaft oder nachteilig für den Dargestellten bewerten. Stets wird die in der bildhaften Darstellung in der Regel mitschwingende Tatsachenbehauptung über die Realität des Abgebildeten unzutreffend.“

Es geht damit um den Schutz vor unbefugter, täuschend echter „digitaler Repräsentation“ (Lennartz, NJW 2023, 3543), also um die Verfälschung des Persönlichkeitsbildes. Durch eine nicht offen gelegte technische Manipulation wird eine unzutreffende Tatsachenbehauptung über die Realität des Betroffenen, sein äußeres Erscheinungsbild, sein Verhalten oder seine mündlichen Äußerungen, aufgestellt. Die in dem Medieninhalt liegende Information über die Person wird von der wirklichen Person des Betroffenen gleichsam entkoppelt und ein konkreter Ausschnitt von dessen Lebenswirklichkeit falsch wiedergegeben. Der Betroffene sieht sich in einer Situation, in der er die Selbstbestimmung und die Kontrolle über das eigene Erscheinungsbild und Auftreten verliert und identitätsprägende Merkmale mit Außenwirkung in – dauerhaft verkörperten und damit mannigfaltig reproduzierbaren – Medieninhalten ge- oder verfälscht oder in einen falschen Kontext gestellt werden.

Der Aspekt einer verfälschenden, aber zugleich authentisch wirkenden Darstellung zentraler Persönlichkeitsmerkmale wird durch die bestehenden Strafvorschriften, namentlich durch die auf den Persönlichkeitsschutz gerichteten Vorschriften in §§ 201, 201a StGB, nicht erfasst. Die Verletzung der Vertraulichkeit des Wortes (§ 201 StGB) betrifft nur tatsächliche Äußerungen der betroffenen Person; bei der Verletzung des höchstpersönlichen Lebensbereichs und von Persönlichkeitsrechten durch Bildaufnahmen (§ 201a StGB) geht es um die authentische Abbildung der Person bei realen Vorkommnissen. Beide Vorschriften weisen aber auf die besondere Schutzwürdigkeit von Aufnahmen des gesprochenen Wortes und des äußeren Erscheinungsbildes hin: In jeder mündlichen Äußerung teilt sich die Persönlichkeit des Sprechers mit, jede bildhafte Wiedergabe einer Person führt beim Wahrnehmenden zur Aufnahme von Informationen, die er der Persönlichkeit des Abgebildeten zuschreibt. Äußeres Erscheinungsbild, Verhalten und Äußerungen sind daher zentrale Elemente, die die Persönlichkeit (Wesen und Charakter) einer Person formen und prägen und nach außen hin festlegen.

Aus dieser Relevanz für die Persönlichkeitsentfaltung ergibt sich auch die grundsätzliche Strafbarkeit von Manipulationen, die eine Person – scheinbar authentisch – so darstellen, wie sie aber gerade nicht ist. Die rasante Entwicklung der Künstlichen Intelligenz und die mit ihrer Hilfe erstellten Deepfakes haben hier zu einer Gefährdung der Persönlichkeit geführt, die gerade auch im Lichte des im Grundgesetz verankerten allgemeinen Persönlichkeitsrechts (Artikel 2 Absatz 1 i. V. m. Artikel 1 Absatz 1 GG) eines Schutzes durch das Strafrecht bedarf. Berechtigte gegenläufige Interessen an einer technisch manipulierten Darstellung einer Person in einer vermeintlich authentischen Bild-, Ton- oder Videoaufnahme ohne deren Zustimmung und

ohne Offenlegung der Manipulation werden demgegenüber in der Regel nicht vorliegen und zumeist auch kein die Interessen des Betroffenen überwiegendes Gewicht haben. Eine Strafbarkeit hier erst dann beginnen zu lassen, wenn die manipulierte Aufnahme geeignet ist, dem Ansehen der Person erheblich zu schaden (so wie dies § 201a Absatz 2 Satz 1 StGB und sachlich vergleichbar auch § 187 StGB vorsieht), wird daher dem Persönlichkeitsschutz des Betroffenen nicht gerecht. Das zeigt sich insbesondere an den Fällen des unbefugten „Voice Clonings“, wie etwa in den genannten Betrugsfällen, wo der Aspekt der Ansehensschädigung ersichtlich keine Rolle spielt, oder in der kontroversen Bewertung der Ansehensschädlichkeit von Nacktaufnahmen. Hinzu kommt, dass die Tat auch Allgemeininteressen berührt. So hat die Gesellschaft ein vitales Interesse an unverfälschten Informationen, da Wahrheit und Vertrauen in Informationen Grundlage für eine freie Meinungs- und Willensbildung ist. Die primär dem Persönlichkeitsschutz dienende Regelung hat insoweit eine ergänzende Schutzfunktion und wirkt auch dem drohenden Verlust des Vertrauens in Medieninhalte entgegen. Sofern im Einzelfall berechnete gegenläufige Interessen an einem Zugänglichmachen von Deepfakes bestehen, wird dem u.a. durch das Erfordernis täuschungsgleichen Täterverhaltens und durch eine Ausnahmeregelung für sozialadäquate Verwendungsformen Rechnung getragen.

5. Die vorgeschlagene Strafvorschrift zum Schutz von Persönlichkeitsinteressen vor einer Verletzung durch Deepfakes stellt nicht nur eine spezifische Regelung des Phänomens und des damit verbundenen Unrechts dar. Sie trägt auch zu einem wirksameren Schutz der von manipulierten Medieninhalten betroffenen Personen bei und schärft das Bewusstsein für den Unrechtscharakter derartiger Manipulationen. Zugleich gibt die vorgeschlagene Strafvorschrift für die Betreiber sozialer Netzwerke und sonstiger Internetplattformen einen klar(er)en Anhaltspunkt dafür, welche Inhalte als rechtswidrig im Sinne des Digital Service Act (DSA) anzusehen und nach den dort geltenden Maßgaben (z. B. notice-and-take-down) zu behandeln sind. Das ist vor allem deswegen bedeutsam, weil Deepfakes häufig über digitale Plattformen verbreitet werden und den Diensteanbietern daher eine erhebliche Bedeutung im Kampf gegen missbräuchliche Deepfakes zukommt.

Die in der (kommenden) KI-Verordnung vorgesehene Offenlegungs-/Kennzeichnungspflicht für Deepfakes ist für sich gesehen kein ausreichendes Instrument gegen technisch manipulierte oder künstlich generierte und persönlichkeitsrechtsverletzende Video- und Audioinhalte. Dies gilt zum einen, weil sich die Verordnung nicht an Nutzer richtet, die Künstliche Intelligenz im Rahmen einer persönlichen und nicht beruflichen Tätigkeit verwenden. Zum anderen werden sich insbesondere Personen, die diese Inhalte auf strafbare Weise einsetzen wollen, von einer ordnungsrechtlichen Transparenzpflicht nicht beeindruckt lassen. Auf den Einsatz des Strafrechts kann daher nicht verzichtet werden.

II. Gesetzgebungskompetenz; Vereinbarkeit mit EU-Recht

Die Gesetzgebungskompetenz des Bundes folgt aus Artikel 74 Absatz 1 Nummer 1 des Grundgesetzes (Strafrecht).

Der Entwurf ist mit dem Recht der Europäischen Union und völkerrechtlichen Verträgen, die die Bundesrepublik Deutschland abgeschlossen hat, vereinbar.

III. Auswirkungen

Auswirkungen auf den Bundeshaushalt sind durch den Entwurf nicht zu erwarten. Durch die Einführung der neuen strafrechtlichen Regelung können den Länderhaushalten Verfahrens- und Vollzugskosten in überschaubarem Umfang entstehen, deren Höhe sich nicht näher beziffern lässt. Für Bürgerinnen und Bürger entsteht kein Erfüllungsaufwand.

B. Besonderer Teil

Zu Artikel 1 (Änderung des Strafgesetzbuches)

Zentraler Gegenstand des Gesetzentwurfs ist die Einführung des Straftatbestandes „Verletzung von Persönlichkeitsrechten durch digitale Fälschung“ in § 201b StGB. Da die Regelung – ebenso wie die Taten nach §§ 201, 201a StGB – dem Persönlichkeitsschutz der von dem manipulierten oder künstlich hergestellten Medieninhalt betroffenen Person dient, wird sie unmittelbar nach diesen Vorschriften in den 15. Abschnitt des Besonderen Teils des Strafgesetzbuches („Verletzung des persönlichen Lebens- und Geheimbereichs“) eingeordnet.

Zu Nummer 1 (Inhaltsübersicht)

Es handelt sich um eine redaktionelle Folgeänderung zur Einführung von § 201b StGB.

Zu Nummer 2 (§ 201b)

In Nummer 2 wird der neue Straftatbestand der „Verletzung von Persönlichkeitsrechten durch digitale Fälschung“ (§ 201b StGB) in das Strafgesetzbuch eingefügt. Die Vorschrift dient vor allem dem Schutz des Persönlichkeitsrechts der betroffenen Personen vor der Verbreitung eines technisch manipulierten oder künstlich hergestellten Medieninhalts, der den Anschein erweckt, eine authentische Video-, Bild- oder Tonaufnahme des äußeren Erscheinungsbildes, des Verhaltens oder der Stimme einer Person zu sein. Hinsichtlich der Einzelheiten wird auf die Ausführungen im Allgemeinen Teil unter I.4. Bezug genommen.

Zu § 201b Absatz 1

Die Bestimmung in Satz 1 enthält die zentrale Regelung zur Strafbarkeit wirklichkeitsverfälschender Persönlichkeitsdarstellungen. Strafbar macht sich, wer das Persönlichkeitsrecht einer anderen Person verletzt, indem er eine mit computertechnischen Mitteln hergestellte oder veränderte Bild- oder Tonaufnahme, die den Anschein einer wirklichkeitsgetreuen Wiedergabe des äußeren Erscheinungsbildes, des Verhaltens oder mündlicher Äußerungen dieser Person erweckt, einer dritten Person zugänglich macht. Während der erste Satzteil („Wer ... verletzt“) das (primäre) Schutzgut und die Angriffsrichtung der Tat bezeichnet, folgt der eigentliche Unrechtstatbestand mit dem „indem“-Satz.

Tatobjekt ist ein mit computertechnischen Mitteln hergestellter oder veränderter Medieninhalt. Der Begriff des Medieninhalts ist neu. Er meint den Informationsgehalt, der in einem Medium verkörpert ist, sich an Adressaten (z. B. Internetnutzer) richtet und dem ein Bedeutungsgehalt zukommt. Durch den nachfolgenden Gesetzestext (Anschein einer wirklichkeitsgetreuen Bild- oder Tonaufnahme) wird der Medieninhalt noch weiter konkretisiert. Daraus ergibt sich, dass dieser (nur) einen visuell oder auditorisch erfassbaren Inhalt meint, wobei auch Bewegtbilder und die Kombination von Bild- und Toninhalten erfasst werden. Der Medieninhalt muss mit computertechnischen Mitteln hergestellt oder verändert sein. Erfasst werden damit nur solche Inhalte, die unter Verwendung entsprechender Computerprogramme (Software), insbesondere unter Nutzung künstlicher neuronaler Netzwerke, erzeugt oder geändert werden und damit in digitaler Form als Bild- und/oder Tondateien vorliegen. Die Regelung schließt daher insbesondere Manipulationen in Gestalt von manuell erstellten Bildcollagen oder persönlich erzeugten Stimmenimitationen aus. Die Eingrenzung trägt dem Umstand Rechnung, dass Gefahren für den Persönlichkeitsschutz vor allem von computertechnisch manipulierten oder künstlichen generierten Medieninhalten ausgehen. Diese sind mittlerweile in der Lage, ein sehr hohes Maß an Authentizität vorzuspiegeln und können auf einfachem Weg einer großen Öffentlichkeit zugänglich gemacht werden. Unerheblich ist, ob der Medieninhalt mit entsprechender Software – auf der Basis entsprechenden „Lernmaterials“ – vollständig künstlich erzeugt worden ist oder ob eine bestehende Inhalt, also eine existierende (Original-)Bild- oder Tondatei, mit computertechnischen Mitteln nachträglich geändert worden ist. Der Tatbestand erfasst beide Varianten.

Der derart hergestellte oder veränderte Medieninhalt muss den Anschein einer wirklichkeitsgetreuen (lebensechten) Bild- (inkl. Video-) oder Tonaufnahme des äußeren Erscheinungsbildes, des Verhaltens oder – als besonders hervorgehobener Teilaspekt des Verhaltens – mündlicher Äußerungen einer anderen Person erwecken. In dieser näheren Bestimmung des Charakters des Inhalts kommt der durch die Tat bewirkte Übergriff in die fremde Persönlichkeitssphäre zum Ausdruck: Es werden Tatsachenbehauptungen zur Realität der betroffenen Person vorge spiegelt, die auf einer (nicht offen gelegten) Manipulation bzw. künstlichen Herstellung des Informationsträgers

beruhen. Den Anschein einer wirklichkeitsgetreuen Bild- oder Tonaufnahme erwecken nur solche Medieninhalte, die für den Betrachter den Eindruck vermitteln, sie stellen – so wie sie vorliegen – eine authentische Aufnahme des äußeren Erscheinungsbildes, des Verhaltens oder mündlicher Äußerungen der betroffenen Person dar. Dies muss auch vom Vorsatz des Täters erfasst sein. Ein solcher Anschein kann durch eine klar erkennbare Manipulation oder eine klar erkennbare künstliche Herstellung entfallen. Von der Strafbarkeit ausgenommen sind daher Fallgestaltungen, in denen jeder verständige Mensch die fehlende Authentizität der Aufnahme erkennen würde. Der Anschein kann aber auch durch eine deutliche Kennzeichnung der Aufnahme als künstlich erzeugter oder veränderter Medieninhalt aufgehoben sein. Strafrechtlicher Schutz wird dann durch anderweitige Regelungen, wie namentlich § 33 des Kunsturheberrechtsgesetzes, bewirkt. Die Manipulation muss sich nicht notwendig auf die erwähnten Persönlichkeitsmerkmale, d.h. äußeres Erscheinungsbild, Verhalten oder mündliche Äußerungen, selbst beziehen, sondern kann sich auch durch deren Einbettung in einen falschen Zusammenhang ergeben, da der Aussagegehalt des Medieninhalts erst dadurch bestimmt werden kann. Den Anschein einer wirklichkeitsgetreuen Bild- oder Tonaufnahme erwecken daher auch solche Medieninhalte, die zwar für sich gesehen ein authentisches Abbild der genannten Persönlichkeitsmerkmale darstellen, dieses Abbild aber in einen von der Realität abweichenden, künstlich geschaffenen Kontext stellen. Zu denken ist etwa an Medieninhalte, in denen eine authentisch wiedergegebene Person durch manipulative Bearbeitung – den tatsächlichen Begebenheiten zuwider – in einer kompromittierenden Situation dargestellt wird.

Aus dem genannten Schutzzweck der Strafvorschrift folgt im Übrigen, dass die vorgenommene Manipulation Persönlichkeitsrelevanz haben muss und ihr im Zusammenhang mit der Tathandlung Eingriffscharakter zukommt. Für den Aussagegehalt unbedeutende Veränderungen (vgl. BVerfG, NJW 2005, 3271, 3273, s.o.) erfüllen den Tatbestand nicht. Hierdurch wird keine Verletzung des Persönlichkeitsrechts bewirkt. Dies soll durch den einleitenden Satzteil „Wer das Persönlichkeitsrecht einer anderen Person verletzt“ auch im Tatbestand klargestellt werden. Was als eine unbedeutende Veränderung anzusehen ist, lässt sich nur auf der Grundlage einer Gesamtwürdigung des Medieninhalts und des Persönlichkeitsbezugs des künstlich hergestellten oder veränderten Inhalts bestimmen. In Betracht kommen namentlich untergeordnete Änderungen des äußeren Erscheinungsbildes und für die Information über die Lebenswirklichkeit des Betroffenen unmaßgebliche Veränderungen der Bildkomposition. Für über diese (Geringfügigkeits-)Schwelle hinausgehende Veränderungen bieten die Regelungen in Absatz 3 (Sozialadäquanz), zum Strafantrag (§ 205 StGB) und zur Privatklage (§ 374 StPO) sowie die allgemeinen strafrechtlichen Vorschriften ausreichend Möglichkeiten zu einer flexiblen, tat- und schuldangemessenen Handhabung.

Die Tathandlung besteht darin, dass der Täter den vorstehend beschriebenen Medieninhalt einer dritten Person zugänglich macht. Dies ist der Fall, wenn ihr die Möglichkeit eröffnet wird, sich durch sinnliche Wahrnehmung vom Inhalt Kenntnis zu verschaffen. Es gelten hierzu die auch für andere Straftatbestände mit gleicher Tatmodalität anerkannten Regelungen.

Die vorgesehene Strafdrohung – Freiheitsstrafe bis zu zwei Jahren oder Geldstrafe – ermöglicht eine dem Unrechtscharakter der Tat angemessene, differenzierende Ahndung. Sie entspricht der in § 201a Absatz 1, 2 StGB vorgesehenen Strafe.

Satz 2 erweitert den durch Satz 1 gewährten Strafrechtsschutz auf zum Tatzeitpunkt bereits verstorbene Personen. Die Regelung dient dem postmortalen Persönlichkeitsschutz und fußt insoweit auf dem Schutzauftrag des Artikel 1 Absatz 1 GG zur Wahrung der Menschenwürde. Auch die verstorbene Person soll dagegen geschützt werden, dass ihr äußeres Erscheinungsbild, ihr Verhalten oder ihre mündlichen Äußerungen zum Gegenstand manipulierter Medieninhalte gemacht wird. Sie ist insoweit schutzbedürftig, weil sie sich hiergegen nicht mehr zur Wehr setzen kann und die scheinbar authentische Aufnahme auch eine verfälschende Darstellung ihrer (früheren) Lebenswirklichkeit und Persönlichkeit beinhaltet. Da der postmortale Persönlichkeitsschutz nach dem Grundgesetz nur Schutz gegen grobe Entstellungen bietet (vgl. BVerfG, Beschl. v. 24. Oktober 2022 – 1 BvR 19/22, ZUM 2023, 189, 191 f.), soll auch der strafrechtliche Schutz – über die Anforderungen des Satzes 1 hinaus – nur auf schwerwiegende Verletzungen des (postmortalen) Persönlichkeitsrechts erstreckt werden. Tatbestandsmäßig ist daher das Zugänglichmachen nur solcher Aufnahmen, die durch ihre Manipulation den Achtungsanspruch der verstorbenen Person besonders nachhaltig verletzen, sie also beispielsweise grob herabwürdigen oder in ihrem Ansehen schädigen. Angesichts der vergleichsweise hohen Schwelle der Strafbarkeit besteht daher auch kein Bedarf für eine Strafschärfungs- und eine Ausnahmeregelung, wie sie in Absatz 2 und 3 für Taten nach Absatz 1 Satz 1 vorgesehen sind.

Zu § 201b Absatz 2

Absatz 2 enthält für die Fälle des Absatzes 1 Satz 1 eine Qualifikation, wenn der Täter oder die Täterin den Medieninhalt der Öffentlichkeit zugänglich macht oder einen Medieninhalt zugänglich macht, der einen Vorgang des höchstpersönlichen Lebensbereichs zum Gegenstand hat. Dies trägt dem Umstand Rechnung, dass die Tat in diesen Fällen zu einer schwerwiegenden Beeinträchtigung des Persönlichkeitsrechts der Betroffenen führen kann – einerseits vor allem wegen der Breitenwirkung der Tat, andererseits wegen des besonders sensiblen Inhalts der vermeintlich authentischen Aufnahme.

In der ersten Variante stellt die Regelung darauf ab, dass der Täter oder die Täterin den in Absatz 1 Satz 1 bezeichneten Medieninhalt der Öffentlichkeit zugänglich macht, also für einen grundsätzlich unbeschränkten, im einzelnen nicht überschaubaren Personenkreis wahrnehmbar macht. Das wird häufig dadurch geschehen, dass der Inhalt auf Instant-Messengern mit einer großen Reichweite oder in öffentlichen Bereichen des Internets gepostet wird. Angesichts der weltweiten Zugriffsmöglichkeit auf das Internet kann ein solcher Inhalt eine außerordentliche Reichweite erlangen und, für die Betroffenen kaum kontrollier- oder überschaubar, innerhalb kürzester Zeit einen großen Adressatenkreis erreichen und von diesem weiter verbreitet werden. Zugleich stehen den Betroffenen häufig nur sehr eingeschränkte Möglichkeiten für eine wirksame und endgültige Entfernung der Aufnahme zur Verfügung. Oftmals bleibt daher die Ungewissheit, dass die Aufnahme weiterhin im Internet abrufbar ist.

In der zweiten Variante knüpft die Strafschärfung daran an, dass der Täter oder die Täterin einen in Absatz 1 Satz 1 bezeichneten Medieninhalt zugänglich macht, der einen Vorgang des höchstpersönlichen Lebensbereichs zum Gegenstand hat. Ob die Verbreitung solcher Inhalte gegenüber einer einzelnen Person oder gegenüber der Öffentlichkeit erfolgt, ist für die Erfüllung des Tatbestands unerheblich. Die Regelung knüpft mit dem Begriff des „höchstpersönlichen Lebensbereichs“ an die Terminologie des § 201a StGB an, so dass die dort anerkannten Grundsätze zum Ausgangspunkt für die Bestimmung seines Inhalts herangezogen werden können. Letztlich geht es um die Darstellung von Vorgängen und Äußerungen, die den Kernbereich der Persönlichkeit und die Intimsphäre betreffen, vor allem die Bereiche Sexualität, Krankheit, Tod und (bezogen auf mündliche Äußerungen) die innere Gedanken- und Gefühlswelt – also Umstände, die unbeteiligten Dritten nicht oder nicht ohne Weiteres zugänglich sind und die gesteigerten Schutz vor dem Einblick Außenstehender verdienen. Die Qualifikation erfasst damit insbesondere Sachverhalte, wie den bereits erwähnten Fall der manipulierten bzw. künstlich hergestellten Nacktaufnahmen von Schülerinnen in Spanien oder die vielfältigen Fallgestaltungen, in denen reale Bildaufnahmen von Frauen, insbesondere deren Gesichter, in einen pornographischen Kontext gestellt werden. Derartige Medieninhalte führen, auch wenn sie keine reale Grundlage haben, regelmäßig zu einer besonders nachhaltigen Verletzung der Persönlichkeitsinteressen der betroffenen Personen (dazu bereits im Allgemeinen Teil unter I. 2. c).

Als Sanktion sieht Absatz 2 Freiheitsstrafe bis zu fünf Jahren oder Geldstrafe vor. Dieser weite Strafrahmen soll der Vielgestaltigkeit der in Betracht kommenden Taten und deren Auswirkungen für die betroffenen Personen Rechnung tragen. Er gewährleistet, dass gerade in Fällen, in denen der Täter oder die Täterin Medieninhalte, die den höchstpersönlichen Lebensbereich betreffen, der Öffentlichkeit zugänglich macht, also kumulativ beide Strafschärfungsgründe erfüllt, spürbare Sanktionen verhängt werden können.

Absatz 2 findet keine Anwendung auf die in Absatz 1 Satz 2 geregelten Fälle der Verletzung des postmortalen Persönlichkeitsrechts, da diese von vornherein auf schwerwiegende Verletzungen des Persönlichkeitsrechts begrenzt sind.

Zu § 201b Absatz 3

Absatz 3 ordnet an, dass die in Absatz 1 Satz 1, auch in Verbindung mit Absatz 2, angedrohte Strafbarkeit nicht für Handlungen gilt, die in Wahrnehmung überwiegender berechtigter Interessen erfolgen. Als in Betracht kommende Interessen nennt die Regelung beispielhaft solche der Kunst, der Wissenschaft, der Forschung, der Lehre, der Berichterstattung über Vorgänge des Zeitgeschehens oder der Geschichte oder ähnliche Zwecke. Absatz 3 entspricht damit der Regelung in § 201a Absatz 4 StGB, so dass grundsätzlich an die dort anerkannten Auslegungsgrundsätze angeknüpft werden kann. Für die Regelung in Absatz 3 ergibt sich aber die wichtige Besonderheit, dass der Tat nach Absatz 1 Satz 1 ein scheinbar authentischer, tatsächlich aber manipulierter oder künstlich generierter Medieninhalt zugrunde liegt und der Täter oder die Täterin hiervon Kenntnis hat und dies bei Verbreitung nicht oder nicht hinreichend deutlich offenlegt. Es besteht regelmäßig kein anerkanntes, überwiegendes Interesse, eine solche Darstellung einer Person ohne deren Zustimmung und ohne Offenlegung der Manipu-

lation Dritten zugänglich zu machen. Denn die in der Darstellung mitschwingende Tatsachenbehauptung über die Realität der betroffenen Person ist unzutreffend und an der Verbreitung falscher Tatsachenbehauptung besteht grundsätzlich kein schützenswertes Interesse. Auch wird den Interessen der verbreitenden Person an der Zugänglichmachung solcher Inhalte regelmäßig dadurch ausreichend Rechnung getragen, dass sie die Manipulation bei deren Verbreitung (klar) offenlegen und damit den Anschein einer wirklichkeitsgetreuen Aufnahme verhindern kann. Derartige Fälle unterfallen dann bereits nicht dem Tatbestand nach Absatz 1. Als mögliche Anwendungsfälle des Absatzes 3 ist insbesondere an solche aus dem Bereich von Kunst und Satire zu denken. Erforderlich ist stets eine auf den Einzelfall bezogene Abwägung zwischen den Persönlichkeitsinteressen der betroffenen Person und den gegenläufigen Interessen des den Medieninhalt Verbreitenden. Dabei wird sich der Rechtsanwender auch an den für § 23 des Kunsturheberrechtsgesetzes entwickelten Abwägungsgrundsätzen orientieren können. In den Fällen, in denen die Tat den höchstpersönlichen Lebensbereich zum Gegenstand hat, wird eine Straffreiheit nach Absatz 3 allerdings kaum in Betracht kommen.

Einer darüber hinausgehenden Ausnahmeregelung für die Fälle, in denen die Strafverfolgungsbehörden rein computergenerierte kinderpornographische Deepfakes zur Verfolgung von Kinderpornographie einsetzen, bedarf es nicht. Die einen solchen Einsatz ermöglichende Regelung in § 184b Absatz 6 StGB erfasst nur dienstliche Handlungen, die sich auf einen kinderpornographischen Inhalt beziehen, der kein tatsächliches Geschehen wiedergibt und auch nicht unter Verwendung einer Bildaufnahme eines Kindes oder Jugendlichen hergestellt worden ist. Die Handlung kann daher nicht zu einer Verletzung von Persönlichkeitsrechten führen und wird von § 201b StGB nicht erfasst.

Absatz 3 findet keine Anwendung auf die in Absatz 1 Satz 2 geregelten Fälle der Verletzung des postmortalen Persönlichkeitsrechts. Eine Abwägung mit gegenläufigen Belangen kann hier bereits im Rahmen der Prüfung einer „schwerwiegenden“ Verletzung des Persönlichkeitsrechts erfolgen.

Zu § 201b Absatz 4

Die Vorschrift eröffnet die Möglichkeit, Bild- oder Tonträger und andere technische Mittel, die der Täter oder Teilnehmer verwendet hat, einzuziehen. Auch auf die in § 74a StGB normierte erweiterte Einziehungsmöglichkeit wird verwiesen, um tatgegenständliche Datenträger möglichst umfassend einzuziehen zu können. Die Regelung in Absatz 4 ist angelehnt an § 201a Absatz 5 StGB, verzichtet aber auf die ausdrückliche Erwähnung von Bildaufnahmegeräten, da diese im Rahmen der Tat nach Absatz 1 keine Rolle spielen.

Zu Nummer 3 (Änderung § 205)

Zu Buchstabe a (Absatz 1)

Die Einführung des neuen Straftatbestands in § 201b StGB macht es erforderlich, auch die Regelung zum Strafantrag in § 205 StGB zu ändern. Durch eine Aufnahme des § 201b StGB in die Regelung des Absatz 1 Satz 2 wird sichergestellt, dass die Strafverfolgung grundsätzlich nur auf Antrag stattfindet. Dies entspricht dem Charakter der Tat als einer primär dem Persönlichkeitsschutz des Einzelnen dienende Regelung. Berechtigt zur Stellung des Strafantrags ist zunächst der Verletzte (§ 77 Absatz 1 StGB), also die Person, die durch die Tat nach § 201b StGB in ihren Rechten verletzt ist. Da die Tat jedoch auch über den Rechtskreis der verletzten Person hinausgehende Wirkung entfalten kann, ist vorgesehen, dass sie – unabhängig von einem Strafantrag – auch dann verfolgt werden kann, wenn die Strafverfolgungsbehörde wegen des besonderen öffentlichen Interesses an der Strafverfolgung ein Einschreiten von Amts wegen für geboten hält. In Betracht kommt dies beispielsweise, wenn die Tat das Bild des Betroffenen in der Öffentlichkeit nachhaltig beeinträchtigt oder die Tat geeignet ist, auch den öffentlichen Meinungsbildungsprozess zu verfälschen.

Zu Buchstabe b (Absatz 2)

§ 201b Absatz 1 StGB enthält in Satz 2 eine Regelung zum Schutz verstorbener Personen. Für diese Fälle sieht der Gesetzentwurf durch entsprechende Ergänzung der Regelung in § 205 Absatz 2 Satz 4 StGB die Schaffung eines Antragsrechts der Angehörigen der verstorbenen Person vor. Dadurch können die Angehörigen die postmortalen Persönlichkeitsinteressen des Verstorbenen (auch) für den Bereich strafrechtlicher Verfolgung wahren. Andernfalls wäre eine Verfolgung der Tat nach der oben (Buchstabe a) vorgesehenen Ergänzung des § 205 Absatz 1 StGB nur möglich, wenn die Strafverfolgungsbehörde wegen des besonderen öffentlichen Interesses an der Strafverfolgung ein Einschreiten von Amts wegen für geboten hält. Der in § 205 Absatz 2 Satz 1 StGB vorgese-

hene Übergang des Antragsrechts für Fälle, in denen der Verletzte stirbt, greift hier nicht, da der Verletzte zum Zeitpunkt der Tat nach § 201b Absatz 1 Satz 2 StGB bereits verstorben war.

Zu Artikel 2 (Änderung der Strafprozessordnung)

Artikel 2 enthält Folgeänderungen in der Strafprozessordnung.

Zu Nummer 1 (Änderung § 100k)

Es handelt sich um eine Folgeänderung, die der Erweiterung des 15. Abschnitts des Besonderen Teils des Strafgesetzbuches Rechnung trägt. Ebenso wie in den bereits in § 100k Absatz 2 Satz 1 Nummer 1 Buchstabe h StPO genannten Fällen soll auch für Taten nach § 201b StGB die Möglichkeit der Erhebung von Nutzungsdaten eröffnet werden.

Zu Nummer 2 (Änderung § 374)

Aufgrund seiner Nähe zu Taten nach § 201a StGB und zu anderen in § 374 Absatz 1 StPO genannten Delikten, die dem Schutz höchstpersönlicher Rechtsgüter dienen, soll auch § 201b StGB in den Kreis der Privatklagedelikte aufgenommen werden. Wie die in § 374 Absatz 1 StPO genannten Vergehen kann eine Verletzung von Persönlichkeitsrechten durch digitale Fälschung die Allgemeinheit mitunter so wenig berühren, dass kein öffentliches Interesse an der Strafverfolgung besteht. In Durchbrechung des Offizialprinzips soll in einem solchen Fall ausnahmsweise der Verletzte bzw. dessen gesetzlicher Vertreter oder ein sonstiger Strafantragsberechtigter die Strafverfolgung als Privatkläger selbst betreiben dürfen.

Zu Artikel 3 (Inkrafttreten)

Die Vorschrift regelt das Inkrafttreten.

A n l a g e

E n t s c h l i e ß u n g

zum

Entwurf eines Gesetzes zum strafrechtlichen Schutz von Persönlichkeitsrechten vor Deepfakes

a) Technologische Lösungen:

Der Bundesrat bittet die Bundesregierung, Programme zur Förderung der Entwicklung und Implementierung von Technologien zur Erkennung und Kennzeichnung von Deepfakes zu initiieren und zu unterstützen. Diese Maßnahmen sollen insbesondere Unternehmen und staatlichen Stellen zugutekommen, um effektiv und effizient gegen missbräuchliche Deepfakes vorgehen zu können.

b) Evaluierung und Anpassung:

Zu wünschen wäre eine regelmäßige Evaluierung der Umsetzung der Gesetze, die den Umgang mit KI inklusive Deepfakes regeln. Die hierfür zuständige Stelle soll Empfehlungen aussprechen, um den Schutz von Persönlichkeitsrechten in Einklang mit technologischen Entwicklungen und dem Bürokratieabbau zu optimieren.

c) Unterstützung der Wirtschaft:

Der Bundesrat bittet, eine zentrale Anlaufstelle einzurichten, die KMU in Fragen der Deepfake-Erkennung und -Kennzeichnung sowie der rechtlichen Anforderungen berät, um wirtschaftliche Belange zu berücksichtigen und bürokratische Hürden abzubauen.

Anlage 2

Stellungnahme der Bundesregierung

Die Bundesregierung nimmt zu dem Gesetzentwurf des Bundesrates wie folgt Stellung:

Die Bundesregierung nimmt die im Gesetzentwurf geschilderten Phänomene ernst und prüft fortlaufend, ob aufgrund aktueller, insbesondere auch technologischer Entwicklungen gesetzgeberischer Handlungsbedarf besteht und das geltende Strafrecht alle strafwürdigen Konstellationen erfasst. Das geltende Recht ermöglicht eine Strafverfolgung der Verbreitung missbräuchlicher Deepfakes bereits in vielfältigen Konstellationen, insbesondere mit den bestehenden Tatbeständen der Verleumdung gemäß § 187 des Strafgesetzbuches (StGB) und der Verletzung des höchstpersönlichen Lebensbereichs und von Persönlichkeitsrechten durch Bildaufnahmen gemäß § 201a Absatz 2 StGB. Letzterer erfasst auch Verfremdungen und Montagen sowie unter Verwendung realer Bilder hergestellte computergenerierte Aufnahmen. Wenn es sich bei den Deepfakes um Darstellungen mit pornographischem Inhalt handelt, kommt eine Strafbarkeit nach § 184 StGB (Verbreitung pornographischer Inhalte), § 184b StGB (Verbreitung, Erwerb und Besitz kinderpornographischer Inhalte) oder § 184c StGB (Verbreitung, Erwerb und Besitz jugendpornographischer Inhalte) in Betracht.

Daneben können abhängig von den Umständen des Einzelfalles weitere Straftatbestände des StGB einschlägig sein. Wenn etwa manipulierte Stimmaufzeichnungen zur Täuschung eingesetzt werden, kommt eine Strafbarkeit nach § 263 StGB (Betrug) in Betracht.

Wer ein Bildnis ohne eine Einwilligung des Betroffenen im Sinne der §§ 22 f. des Gesetzes betreffend das Urheberrecht an Werken der bildenden Künste und der Photographie (KunstUrhG) verbreitet oder öffentlich zugänglich macht, kann sich außerdem gemäß § 33 KunstUrhG strafbar machen. Denn nach dem von der Rechtsprechung verwendeten sogenannten subjektiven Bildnisbegriff fällt jede erkennbare Wiedergabe des äußeren Erscheinungsbildes einer Person in den Anwendungsbereich von § 22 KunstUrhG, ohne dass es auf deren Herstellungsform ankäme.

Werden urheberrechtlich geschützte Inhalte als Ausgangsmaterial genutzt, kann die Verbreitung eines Deepfakes eine Urheberrechtsverletzung darstellen und nach den §§ 106, 108 des Urheberrechtsgesetzes strafbar sein. Darüber hinaus kann je nach Fallgestaltung § 42 Absatz 2 des Bundesdatenschutzgesetzes erfüllt sein.

Hinsichtlich des vorgeschlagenen § 201b StGB (Verletzung von Persönlichkeitsrechten durch digitale Fälschung) ist anzumerken, dass dieser teilweise enger gefasst ist als die §§ 187, 201a Absatz 2 StGB, da er einen Verletzungserfolg voraussetzt und nicht eine bloße Eignung zur Ansehenschädigung genügen lässt. Gleichzeitig erscheint die Formulierung dieses Verletzungserfolgs im vorgeschlagenen § 201b Absatz 1 StGB („Wer das Persönlichkeitsrecht einer anderen Person verletzt...“) angesichts der Weite und Offenheit des Rechtsgutes des allgemeinen Persönlichkeitsrechts mit Blick auf das Bestimmtheitsgebot (Artikel 103 Absatz 2 des Grundgesetzes) nicht unbedenklich.

