06. Februar 2026

Stellungnahme zur Expertenanhörung zu „Technischer Jugendmedienschutz, Plattformgestaltung und Verantwortung digitaler Dienste“

Die gemeinnützige Organisation HateAid wurde 2018 gegründet und hat ihren Hauptsitz in Berlin. Sie setzt sich für einen sicheren digitalen Raum ein und engagiert sich auf gesellschaftlicher wie politischer Ebene gegen digitale Gewalt und ihre Folgen. HateAid unterstützt deutschlandweit Betroffene von digitaler Gewalt konkret durch psychosoziale Beratung und Sicherheitsberatung.

Die Organisation ist vom Programm “Demokratie Leben!” beim BMBFSFJ gefördert und Teil des Netzwerks “toneshift”. HateAid setzt sich mit der Unterstützung von “Demokratie Leben!” für die Aufklärung über digitale Gewalt und die Rolle von Online-Plattformen bei ihrer Verbreitung ein. Auch Kinder- und Jugendliche stehen hierbei vermehrt im Fokus. Mit der Universität Klagenfurt wurde eine Studie zur Betroffenheit von jungen Erwachsenen durchgeführt.[1] Im Februar 2026 eröffnet HateAid gemeinsam mit neuland&gestalten eine Ausstellung mit dem Titel “Alles im Angebot”. Ziel ist es die Funktionsweise von Online-Plattformen und Mechanismen digitaler Gewalt bildhaft zu erläutern. Mit ihr wurde ein niedrigschwelliges und interaktives Bildungs- und Präventionsangebot geschaffen, welches deutschlandweit auch von Schulklassen und Bildungseinrichtungen besucht werden kann. Darüber hinaus ist die Einrichtung eines Supervisionsangebots für Eltern in der Betroffenenberatung von HateAid geplant, um Eltern zu befähigen, informierte Entscheidungen zur Nutzung sozialer Medien durch ihre Kinder zu treffen. Die Notwendigkeit dafür zeichnete sich in der Betroffenenberatung in den stark gestiegenen Anfragen durch Eltern und Lehrpersonal ab. Aus dieser Erfahrung und dem langjährigen Einsatz für sichere Plattformen gründet sich HateAids Expertise zum Schutz von Kindern und Jugendlichen im digitalen Raum – sowohl zu praxisnahen Einschätzungen der derzeitigen Lage auf Online-Plattformen für Kinder und Jugendliche als auch für Vorschläge wie den Gefahren regulatorisch begegnet werden kann.

I. Risiken, Schutzmaßnahmen und Gestaltung digitaler Dienste

Die digitale Welt bietet Kindern und Jugendlichen enorme Chancen – doch sie birgt auch nie dagewesene Risiken. Der aktuelle Grok-Skandal auf der Plattform X zeigt dies auf erschreckende Weise: Innerhalb von nur 11 Tagen generierte das KI-Modell, neben missbräuchlichen Darstellungen von Erwachsenen, über 23.000

[1] HateAid gGmbH. (2024). *In meinem Netz soll es keine Gewalt geben! Wie junge Erwachsene digitale Gewalt erleben und wie sie damit umgehen* (Studie). Berlin. https://hateaid.org/wp-content/uploads/2024/07/hateaid-studie-junge-erwachsene-2024.pdf

sexualisierte Deepfakes von Kindern, die zum Teil bis heute online zu finden sind. Dieser Vorfall ist kein Einzelfall, sondern ein Symptom für ein systemisches Versagen beim Schutz von Minderjährigen im digitalen Raum. Um solche Risiken wirksam zu bekämpfen, müssen technische, regulatorische und pädagogische Maßnahmen eng ineinandergreifen.

Designentscheidungen der Plattformen sind maßgeblich, um schädliche Auswirkungen, wie z.B. Suchtgefahren, Cybergrooming oder auch der Exposition mit extremistischen oder gewaltvollen Inhalten zu minimieren.

Digitale sexualisierte Gewalt, wie Cybergrooming, erfordert vor allem proaktive technische Lösungen. Plattformen müssen Safety-by-Design als Standard implementieren und Einstellungen vorhalten, die per default verhindern, dass die Konten von Kindern und Jugendlichen von Fremden kontaktiert werden können: Es sollten Standard-Privatsphäre-Einstellungen für Minderjährige gelten, die sicherstellen, dass ihre Profile und Inhalte nur für bestätigte Kontakte sichtbar sind. Plattformen sollten zudem verpflichtet werden, verdächtige Inhalte unverzüglich an Jugendschutzorganisationen zu melden, um eine schnelle Reaktion zu ermöglichen.

Wirksame Maßnahmen:

- standardmäßig private Profile und eingeschränkte Kontaktmöglichkeiten für Minderjährige, sodass Nachrichten oder Freundschaftsanfragen nur von bestätigten Kontakten möglich sind,
- Schutzmechanismen für Bilder und Videos, etwa erschwerte Weiterverbreitung, Screenshot-Blockaden oder Warnhinweise bei sensiblen Inhalten,
- automatische Unkenntlichmachung bzw. Filterung pornografischer oder sexualisierter Inhalte,
- niedrigschwellige, leicht auffindbare Melde- und Hilfefunktionen mit schneller Bearbeitung durch geschulte Moderationsteams,
- altersgerechte Sicherheitsinformationen und Warnhinweise.

Suchtfördernde Gestaltungspraktiken, die darauf abzielen, die Nutzungsdauer zu maximieren, sind ein zentraler Bestandteil des Geschäftsmodells von vielen Plattformen. Die EU-Kommission hat am 06.02.2026 bekannt gegeben, dass ein Verfahren gegen TikTok eingeleitet hat. Diesem liegt die vorläufige Feststellung zugrunde, dass TikTok suchterzeugende Designs verwendet, die gegen den Digital Services Act verstoßen. Diese und ähnliche Mechanismen, wie z.B. das gezielte Anzeigen von Magersuchtinhalten (“#skinnytok”) oder gefährliche Mutproben (“Tripod-Challenge”), schaden der psychischen und physischen Gesundheit von Kindern und Jugendlichen. Hier braucht es klare gesetzliche Regelungen: Features wie Endless Scrolling, Autoplay oder Push-Nachrichten müssen für Minderjährige standardmäßig deaktiviert sein und dürfen nur nach expliziter, informierter Zustimmung aktiviert werden. Die Erläuterungen diesbezüglich müssen niedrigschwellig zugänglich und leicht verständlich sein. Zudem müssen Transparenzpflichten für Algorithmen eingeführt werden, damit Nutzende und Eltern nachvollziehen können, warum bestimmte Inhalte empfohlen werden. Ebenso problematisch sind Mikrotransaktionen und Lootboxen, die bekannte Glücksspielstrategien ausnutzen. Diese sollten für Minderjährige grundsätzlich verboten werden, wie es in Ländern wie Belgien oder den Niederlanden bereits der Fall ist.

Wirksame Maßnahmen:

- Deaktivierung oder Einschränkung von Endlos-Scrollen, Autoplay und permanenten Push-Benachrichtigungen,
- Keine algorithmische Verstärkung schädlicher Inhalte, wie Mutproben oder Magersuchtinhalten,
- standardmäßig deaktivierte In-App-Käufe, Mikrotransaktionen und Zufallsmechaniken (z. B. Lootboxen),
- Empfehlungssysteme, die keine engagement getriebenen Inhalte priorisieren.

Die Verbreitung von Hate Speech, Desinformation und extremistischen Inhalten erfordert eine stärkere Moderation durch Plattformen. „Trust and Safety"-Teams müssen personell aufgestockt werden, um Hassrede und extremistische Inhalte frühzeitig zu erkennen und zu entfernen. KI-gestützte Moderation kann dabei helfen, schädliche Inhalte schneller zu identifizieren, ersetzt jedoch keine menschliche Prüfung durch geschultes und sensibilisiertes Personal. Diese gilt insbesondere für kontextsensitive Moderationsentscheidungen, um auch kodierte Sprache oder subtile Radikalisierungsstrategien zu erfassen. Ein transparenter Umgang mit gelöschten Inhalten und gesperrten Accounts ist dabei essenziell, um Vertrauen in die Moderationspraxis zu schaffen. Dies verlangt nicht zuletzt der DSA selbst in Art.17. Dabei müssen Moderationsentscheidungen auch altersgerecht begründet und erklärt werden.

Wirksame Maßnahmen:

- altersgerechte Kuratierung von Inhalten und Beschränkung der Sichtbarkeit jugendgefährdender oder extremistischer Beiträge,
- Anpassung von Empfehlungssystemen, sodass Qualität und Relevanz höher gewichtet werden als reine Interaktionsraten,
- transparente und einfach zugängliche Meldefunktionen,
- schnelle Moderation und kindgerecht formulierte, nachvollziehbare Entscheidungen,
- ergänzende Informations- und Medienkompetenzangebote.

Cybermobbing und Interaktionsrisiken, wie Doxxing oder Stalking, erfordern Echtzeit-Interventionsmechanismen, die koordinierte Angriffe erkennen und Betroffenen schnelle Schutzoptionen bieten. Anonyme Meldemechanismen sollten sicherstellen, dass Nutzende ohne Angst vor Repressalien Hilfe suchen können. Bei Selbstgefährdungsrisiken, wie Suizid, Essstörungen oder gefährlichen Challenges, müssen Plattformen proaktive Hilfsangebote bereitstellen, wenn Nutzende nach schädlichen Inhalten suchen. Algorithmen dürfen keine „Rabbit Holes" für selbstschädigendes Verhalten schaffen.

Wirksame Maßnahmen:

- private Standardeinstellungen und begrenzte Sichtbarkeit persönlicher Informationen,
- einfache Block-, Mute- und Einschränkungsfunktionen für einzelne Kontakte oder Gruppen,
- niedrigschwellige Meldesysteme,
- konsequente Sanktionierung belästigender und übergriffiger Accounts,
- Warnhinweise bei einschlägigen Suchanfragen oder Beiträgen,
- Einblendung von Hilfs- und Beratungsangeboten,
- schnelle Entfernung gefährlicher Challenges,
- geschulte Moderation für besonders sensible Themen,
- algorithmische Begrenzung selbstgefährdender Inhalte.

Die **kommerzielle Ausnutzung** von Kindern durch Tracking, Profiling oder irreführende Werbung muss durch strikte Datenschutzregeln unterbunden werden. Kinder dürfen nicht für personalisierte Werbung getrackt werden, und Influencer:innen müssen Werbung unmissverständlich kennzeichnen. Eltern benötigen klare Informationen, wenn Bilder ihrer Kinder öffentlich geteilt werden.

Wirksame Maßnahmen:

- Datensparsamkeit („Privacy by Design“) und Erhebung nur notwendiger Daten,
- keine personalisierte Werbung oder Profilbildung für Minderjährige,
- Deaktivierung von Tracking- und Drittanbieter-Cookies,
- transparente Informationen zur Datenverarbeitung,
- einfache Möglichkeiten zur Löschung von Daten und Accounts,
- Schutz vor irreführender oder manipulativer Werbung.

Was aus der Fülle an wirksamen Maßnahmen deutlich wird: die **Gestaltung digitaler Dienste** spielt eine entscheidende Rolle für den Schutz von Kindern. Default-Einstellungen sollten dabei höchste Privatsphäre-Standards voreinstellen, etwa durch private Profile für unter 16-Jährige. Empfehlungssysteme müssen so gestaltet sein, dass sie keine schädlichen Inhalte an Minderjährige ausspielen. Parental Controls sollten einfach bedienbar sein, ohne die Kinder zu überwachen. Erläuterungen und Einstellungen müssen einfach erklärt und niedrigschwellig zugänglich sein, damit auch Kinder und Jugendliche sie verstehen.

Plattformen müssen regelmäßige **Risikoanalysen** durchführen und öffentlich machen, wie es der Digital Services Act (DSA) in Artikel 34 vorschreibt. Ermitteln die Plattformen ein Risiko müssen sie diesem mit angemessenen, verhältnismäßigen und wirksamen Risikominderungsmaßnahmen begegnen, Art.35 DSA. Zu diesen Risiken gehören u.a. auch die Ausübung der Unions-Grundrechte und nachteilige Auswirkungen auf die Gesundheit von Minderjährigen. Zu bewerten ist hierbei explizit auch die Gestaltung von Empfehlungssystemen, der Moderationssysteme und datenbezogene Praktiken. Ein aktuelles Beispiel für systemische Risiken und die rechtlichen Anforderungen zur Risikominderung liefert die formelle, vorläufige Feststellung der Europäischen Kommission, dass die Social-Media-Plattform TikTok in ihrer derzeitigen Ausgestaltung gegen Pflichten aus dem Digital Services Act (DSA) verstößt. Nach einjähriger Untersuchung hat die Kommission festgestellt, dass Kernfunktionen der App — etwa „infinite scroll“, autoplay, push-Notifications und ein stark personalisiertes Empfehlungs-Algorithmus Design — nicht ausreichend im Rahmen der DSA-Pflichten bewertet wurden und durch ihre sucht- und zwanghaften Nutzungseffekte signifikante Risiken für die physische und psychische Gesundheit von Nutzerinnen und Nutzern, insbesondere Minderjährigen, begründen, ohne dass angemessene, effektive Minderungs- und Schutzmaßnahmen erkennbar wären. Die Kommission verlangt daher strukturelle Änderungen im Design, darunter etwa wirksamere screen-time-breaks und Anpassungen im Empfehlungssystem, um den gesetzlich geforderten Schutz gegenüber jungen Menschen und vulnerablen Gruppen sicherzustellen.

Unabhängige Audits der Algorithmen durch externe Stellen, wie Jugendschutzorganisationen, sind notwendig, um Manipulation und Radikalisierung zu verhindern.

II. Plattformverantwortung, Regulierung und Durchsetzung

Die Verantwortung der Plattformen gegenüber Kindern und Jugendlichen muss klar definiert und überprüfbar gestaltet werden. Der DSA und der Jugendmedienschutz-Staatsvertrag (JMStV) bieten zwar starke Instrumente, werden jedoch oft nicht konsequent durchgesetzt. Plattformen müssen verbindliche Sorgfaltspflichten erfüllen, etwa durch regelmäßige Risikoanalysen und Transparenzberichte. Melde- und Abhilfewege für Nutzende müssen klar kommuniziert und leicht zugänglich sein. Sanktionen bei Verstößen sollten verschärft werden, um Plattformen zur Einhaltung der Regeln zu bewegen.

Die **Umsetzung bestehender Vorschriften** muss verbessert werden. Der DSA verlangt von Plattformen, Risiken für Minderjährige zu analysieren und zu mindern (Art.34 Abs.1 d)). Zusätzlich zu dieser Verpflichtung sieht Art.28 DSA vor, dass Online-Plattformen geeignete und verhältnismäßige Maßnahmen ergreifen, um die Privatsphäre und Sicherheit von Minderjährigen zu gewährleisten, sowie für ihren Schutz zu sorgen. Die Kommission hat im Juli 2025 Leitlinien zum Jugendschutz veröffentlicht. Diese Leitlinien dienen als Anhaltspunkt für die Plattformen, wie sie ihre Dienste und etwaige Altersverifikationsmechanismen konform mit Art.28 ausgestalten können. Die Umsetzung dieser Leitlinien ist nicht verpflichtend – jedoch können sie als Auslegungshilfe herangezogen werden, ob ein Verstoß gegen Art. 28 DSA vorliegt. Dies muss durch die Aufsichtsbehörden kritisch geprüft werden.

Plattformen generieren Einnahmen, indem sie die Daten ihrer Nutzenden sammeln und monetarisieren. Um ihre Gewinne zu maximieren, sind sie daher bestrebt, die Zahl ihrer Nutzenden und deren Interaktion mit der Plattform zu steigern. Darum nutzen Plattformen süchtig machenden Designelementen und verstärken reichweitenstarke Inhalte algorithmisch – auch wenn diese Inhalte oft kontrovers, unwahr oder geradezu schädlich sind. Starke Jugendschutzauflagen, die ein solches Plattformdesign verbieten, steht somit dem

Profitinteresse der Plattformen diametral gegenüber. Dies schafft einen fortwährenden Anreiz für Plattformen, lästige Vorschriften zu umgehen. Um langfristig die Sicherheit von Kindern und Jugendlichen online zu gewährleisten, müssen daher alternative Plattformmodelle gezielt gefördert werden, die dezentral organisiert sind und ohne überwachungsbasierte Werbung oder süchtig machende Algorithmen auskommen (z.B, Fediverse, Eurosky, NYZZU). Neben der Bereitstellung von finanziellen Mitteln ist auch die Einführung von gesetzlich verankerten Portabilitätsrechten erforderlich, um Nutzenden den Wechsel zwischen Plattformen zu erleichtern.

III. Altersgrenzen, Altersverifikation und spezifische Schutzräume

Ein Verbot digitaler Dienste, insbesondere sozialer Medien, wie es derzeit in Australien für unter 16-jährige erprobt wird, ist nicht erstrebenswert und darf nur das letzte Mittel sein. Hierbei muss nicht nur das Recht von Kindern und Jugendlichen auf soziale Teilhabe mit dem Recht auf Schutz vor Gewalt, Sucht und Radikalisierung abgewogen werden. Auch die Rechte anderer, erwachsener Nutzender müssen mitgedacht werden, die ebenso z.B. von Altersverifikationsmaßnahmen betroffen sein können. Bevor ein solches Verbot erwogen wird, müssen daher alle anderen Optionen ausgeschöpft werden. **Vorzugswürdig sind sichere Plattformen, auf denen Kinder und Jugendliche alle Vorzüge der Unterhaltung, des Wissenstransfers und der sozialen Teilhabe ausschöpfen können.** Würden alle sehr großen Online-Plattformen, die als soziale Netzwerke eingestuft werden, die Vorgaben der Leitlinien der EU-Kommission zu Art. 28 Abs. 4 DSA erfüllen, wären bereits viele Gefahren für Kinder und Jugendliche adressiert. Die Einführung eines Verbots würde die Anwendung von Altersverifikationsmaßnahmen wohl erforderlich machen, damit das Alter der Nutzenden verlässlich festgestellt werden kann. Aktuell beschränken sich derartige Überprüfungen oft auf das Anklicken einer Bestätigung älter als 18 Jahre zu sein.

Auch ohne ein Verbot digitaler Dienste könnte sich jedoch die Erforderlichkeit der Einführung einer Altersverifizierung ergeben. Dies gilt vor allem dann, wenn Plattformen Kindern und Jugendlichen andere Funktionalitäten zur Verfügung stellen und Inhalte anzeigen wollen als Erwachsenen. Die Frage, wie dies datensparsam ausgestaltet werden kann, ist hier von maßgeblicher Bedeutung. Gemäß Art. 28 Abs. 3 DSA sind Diensteanbieter nicht verpflichtet diesbezüglich zusätzliche personenbezogene Daten zu erheben. braucht es einen **harmonisierten Rechtsrahmen für verbindliche Altersverifikation auf europäischer Ebene.** Nationale Alleingänge, wie sie beim NetzDG zu beobachten waren, sollten vermieden werden, da sie zu Rechtsunsicherheiten führen.

Auch hier gilt: Vor allem in sozialen Netzwerken ist eine sichere Ausgestaltung für alle Nutzenden und nicht nur für Kinder und Jugendliche vorzugswürdig. Hierfür braucht es jedenfalls eine konsequente Umsetzung der geltenden Regulierung. Auch neue Phänomene, wie z.B. KI-generierte Inhalte, müssen hier mitgedacht werden.

Gelingt es nicht, die Vorgaben zum Schutz von Kindern und Jugendlichen, sowie der Minimierung systemischer Risiken zeitnah durchzusetzen, müssen Altersgrenzen für die Nutzung digitaler Dienste erwogen werden. Dies

legen aktuelle Forschungsergebnisse nahe, die u.a. die Nationale Akademie der Wissenschaft Leopoldina eindrücklich dargelegt hat.[2]

Altersgrenzen sollten je nach dem Inhalt, dem Design und den Sicherheitsvorkehrungen einer Plattform variieren. Die Festlegung von Altersgrenzen muss mit wissenschaftlicher Expertise erfolgen. Rechtlich betrachtet sind Kinder ab 14 Jahren beschränkt geschäftsfähig. Erst ab 16 Jahren können Jugendliche in die Verarbeitung ihrer personenbezogenen Daten laut DSGVO ohne Zustimmung der Erziehungsberechtigen einwilligen. Es ist davon auszugehen, dass diese Vorgabe derzeit flächendeckend missachtet wird. Die meisten großen sozialen Netzwerken legen die Grenze zur Anmeldung in ihren allgemeinen Geschäftsbedingungen auf 14 Jahre fest.

Technische Verfahren zur Altersüberprüfung müssen sich strikt an die Grundsätze der Datenminimierung halten und Diskriminierung oder Profiling vermeiden. Der Missbrauch des Jugendschutzes als Vorwand für eine umfassendere Datenerhebung muss ausdrücklich verhindert werden. Zu diesem Zweck sollte die Altersüberprüfung über spezialisierte, unabhängige Drittanbieter (z.B. Yoti-ID) oder Wallet-Lösungen (z.B. EUDI-Wallet) und nicht durch die Plattformen selbst erfolgen. Die Überprüfung sollte sich auch nicht ausschließlich auf Ausweisdokumente stützen, um die Inklusion von Personen ohne oder mit vorläufigen Dokumenten zu gewährleisten. Es sollten daher stets mehrere Optionen zur Verifizierung zur Auswahl stehen.

Bereits jetzt zeigen alternative Plattformen, dass durch Safety-by-Design-Ansätze geschützte digitale Räume geschaffen werden können, in denen Kinder und Jugendliche nur mit Gleichaltrigen interagieren und Inhalte konsequent moderiert werden. Diese spezifischen Schutzräume kommen ohne öffentliche Profile oder manipulative Algorithmen aus und geben auch keine Daten an Dritte weiter, um die Privatsphäre der Nutzenden zu wahren. Da es wichtig ist, dass kinderfreundliche digitale Schutzräume auch sonst keine kommerziellen Interessen verfolgen, sollten diese öffentlich unterstützt werden.

IV. Chancen, Teilhabe und spezielle Problemfelder

Ein wirksamer Jugendmedienschutz schafft sichere Lernumgebungen, die digitale Kompetenz und Kreativität fördern. Kinder und Jugendliche, die sich online frei ausdrücken können, ohne Hass oder Manipulation ausgesetzt zu sein, lernen den sicheren und selbstbewussten Umgang mit Algorithmen, Diskursdynamiken und Meinungsvielfalt. Sie entwickeln Kompetenzen für kritisches Denken und Medienbewertung. Dies ermöglicht ihnen soziale Medien und Messengerdienste als Instrumente zu begreifen, die ihnen helfen, am sozialen Leben teilzuhaben, sich zu informieren und die eigene Meinung auszudrücken.

Online-Communities bieten auch wichtige Räume, in denen Kinder und Jugendliche bei moderater Nutzung Zugehörigkeit erfahren können – insbesondere dann, wenn sie offline isoliert sind. Digitale Teilhabe kann damit ein Ausgleich sein für Benachteiligungen im lokalen Umfeld. Dies gilt auch für den Peer-to-Peer-Wissenstransfers (via Tutorials, Erklärvideos, Diskussionsforen etc.) durch den jungen Menschen sich frühzeitig ihren Interessen widmen können und etwaige Bildungsdefizite in ihrem Umfeld ausgleichen können.

[2] Brailovskaia, J., Buchmann, J., Hertwig, R., Metzinger, T., Montag, C., Sadeghi, A.-R., Schneider, S., Spiecker gen. Döhmann, I. & Waldherr, A. (2025): Soziale Medien und die psychische Gesundheit von Kindern und Jugendlichen. Diskussion Nr. 40, Halle (Saale): Nationale Akademie der Wissenschaften Leopoldina.

Die Erziehung von Kindern und Jugendliche zu einem gesunden Umgang mit sozialen Medien erfordert die Anstrengung sowohl von Eltern als auch von Schulen. **Die Plattformen sind in der Pflicht, suchtfördernde Designs zu vermeiden** und die Nutzungszeiten von Minderjährigen zu durch Nudges wie Pausenerinnerungen oder Nutzungsdaueranzeigen zu reduzieren. Parental Controls sollten einfach bedienbar sein und es Eltern erlauben, die Nutzungszeit ihrer Kinder zu reglementieren, ohne die Kinder zu überwachen.

Über HateAid

HateAid gGmbH

Greifswalder Straße 4

10405 Berlin

E-Mail: kontakt@hateaid.org

hateaid.org

Sitz der Gesellschaft: Berlin

Registergericht: Amtsgericht Charlottenburg

Handelsregisternummer: HRB 203883 B

Umsatzsteuer-Identifikationsnummer: DE322705305

Lobbyregisternummer: R001880

Geschäftsführung: Anna-Lena von Hodenberg, Josephine Ballon