

Unabhängige Expertenkommission "Kinder- und Jugendschutz in der digitalen Welt"

Hearing 2 – Technischer Jugendmedienschutz, Plattformgestaltung und Verantwortung digitaler Dienste – Frist für die Abgabe der schriftlichen Stellungnahme: 06.02.2026 –

LEITFRAGE Welche Kombination aus technischen, regulatorischen und pädagogischen Maßnahmen ist aus Ihrer fachlichen Sicht am wirksamsten, um Kinder und Jugendliche im digitalen Raum zu schützen und ihre Teilhabechancen zu stärken?

I Risiken, Schutzmaßnahmen und Gestaltung digitaler Dienste

- 1. Welche Maßnahmen sind aus Ihrer Sicht am wirksamsten, um die folgenden Risiken für Kinder und Jugendliche in digitalen Diensten zu verhindern, zu erschweren oder in ihren Folgen zu begrenzen:**

a. digitale sexualisierte Gewalt (z.B. Cybergrooming, Sextortion, Missbrauchsdarstellungen)

- Kontakt- und Auffindbarkeit einschränken durch Default-Einstellung auf „private“; Friktionsbehaltung von Fremdkontakt, Gruppeneinladungen nur mit Zustimmung) – senkt Opportunitäten für Grooming erheblich.
- Screenshot-/Download-Schutz bei Inhalten von Minderjährigen (wo immer technisch möglich) + „Forwarding“-Friktion bei sensiblen Medien zur Reduzierung von Weiterverbreitungsmöglichkeiten und quantitativer Absenkung von potenziellem Erpressungsdruck.
- Proaktive Erkennung illegaler Inhalte & Grooming-Muster durch Kombination aus Hash-Matching, Mustererkennung, Human Review, „code emoji“-Signale etc. in Risiko-Services kombiniert mit strikten Fehlalarm-/Beschwerdeprozessen, die benutzerfreundlich zu gestalten sind.
- Sicherheitsmechanismen in DMs: Warnhinweise bei Alters-/Beziehungsanbahnung, „nudge to report“-Mechanismus, Limitierung unbekannter DMs, sichere Medienannahme (z. B. Blur/Click-to-reveal) zur

Verbesserung des Schutzes vor unerwünschtem Kontakt/Grooming in den DSA-Maßnahmen erwähnt.

- **Spezialisierte „Child Safety“-Queues** (24/7 bei großen Diensten), Einrichtung von Hotlines mit eindeutigen Eskalationsprotokollen, schnelle Beweissicherung und Unterstützung für Betroffene. Bessere Verzahnung mit Strafverfolgung.
- **Schutz und Empowerment Ansatz des DSA umsetzen, u.a. durch transparente Rückmeldungen bei Meldungen + „Stay safe“-Ressourcen** zur Verbesserung des digitalen Kinder- und Jugendschutzes inklusive mentaler Gesundheit und Traumavorsorge (Sextortion-Checklisten, Notfallkontakte).

b. suchtbegünstigende/hochgradig bindende Gestaltungspraktiken (z.B. Mikrotransaktionen, Lootboxen/Zufallsmechaniken, Dark Patterns)

- **Deaktivierung/Begrenzung von Engagement-treibenden *features by default* für Minderjährige**, wie z.B. Autoplay, Push-Spam, „Streaks“, Read Receipts, ephemere „FOMO“-Mechanismen.
- **Finanzielle Schutzgeländer**: klare Preiswahrheit ohne „verschleiende“ virtuelle Währungen, obligatorische Kauf-Friktionen („Are you sure?“), Ausgabenlimits, Timeouts nach Käufen, Einrichtung von „Spending dashboards“ mit Eltern
- **Transparenzverpflichtung von Herstellern für Lootboxen und angebliche Zufallsmechanik zur Verbesserung des Minderjährigen-Schutzes** (Einrichtung von verbindlichen Warnhinweisen und Ausgaben Caps). Verbraucherschutz und Minderjährigen-Schutz sind zu verzahnen. Zu wünschen: Lootboxen grundlegend zu verbieten.
- **Verpflichtende Dark-Pattern-Audits & UX und Red Teaming** zur Aufdeckung von manipulative Defaults und „Confirmshaming“-Nachrichten, die emotionalen Druck aufbauen. Sog. „harmful commercial practices“ müssen durch einen stärker jugend-zentrierten Verbraucherschutz transparent gemacht und abgestellt werden.
- **Standardmäßige Nutzungs- und Pausen-Tools** wie Session Timer, „Take a break“-Nachrichten, Nachtmodus, Fokusmodus und **Erklärungen in jugendgerechter Sprache**.-Timer, „Take a break“

c. Hate Speech, Desinformation sowie verschwörungsideologische und extremistische Inhalte

- **Ranking- und Empfehlungslogik neu ausrichten:** weniger automatische, von der Plattform gesteuerte *engagementbasierte* Verstärkung von Feeds. Stattdessen mehr bewusste Kontrolle durch die Nutzer*innen. Für Minderjährige sollen sichere, voreingestellte Standard-Feeds gelten, die nicht primär durch algorithmisches Engagement getrieben werden, sondern durch etablierte Schutzkonzepte und Wahlmöglichkeiten.
- **Klare Umsetzungs-Regeln und abgestufte Maßnahmen:** Die *Policies* der Plattformen und *Enforcement-Tools* müssen darauf ausgerichtet sein, systemische Risiken deutlich zu reduzieren und abgestufte, nachvollziehbare Sanktionen mit nutzerfreundlichen Beschwerdewegen zu implementieren.
- **Die Sichtbarkeits- und Reichweitenbegrenzung problematischer Inhalte ist durch „Soft Interventions“ wie Warnhinweise** vor dem Teilen umfassender durchzusetzen („*Diese Quelle ist unbestätigt*“, „*Dieser Inhalt wurde als potenziell manipulativ gemeldet*“). Zusätzlich sollten Warnbanner zuverlässig bei Inhalten auftauchen, die z. B. medizinische oder klar erkennbare politische Falschinformationen enthalten. Dienliche Hinweise vor der Weiterverbreitung sind auszudehnen („Möchtest du den Artikel erst lesen?“). Die betreffenden Inhalte werden **nicht gelöscht**, aber weniger empfohlen oder gar nicht mehr algorithmisch verbreitet. Sie tauchen **nicht mehr im For-You-Feed** auf, werden aber im Profil weiterhin angezeigt. Es ist stärker darauf zu achten, dass es keine algorithmische Verstärkung von Themen gibt, die unter Gesichtspunkten des Kinder- und Jugendschutzes gesundheitsgefährdend sind (z. B. bedenkliche Selbstoptimierung, extreme Diäten, gefährliche Challenges).
- **Entfernung illegaler Inhalte** (darunter Gewaltverherrlichung, extremistische Inhalte, Hasskriminalität) sowie zuverlässige Entfernung von Inhalten, die klar gegen Plattformregeln verstoßen. Hier müssen die sog. Enforcement-Gaps geschlossen werden. Plattformen müssen gezwungen werden, illegale Inhalte nachweisbar, umfänglich und zeitnah zu entfernen.
- **Klare Maßnahmen gegen Regelverstöße:** Für Nutzer*innen, die wiederholt gegen Regeln verstoßen, braucht es konsequente Maßnahmen wie deutliche Verwarnungen, temporäre Sperrungen bis hin zu dauerhaften Kontolöschung bei systematischem Missbrauch der Plattform. Darüber hinaus sollten die Werbemöglichkeiten für Accounts deaktiviert werden, die wiederholt gegen Regeln verstoßen.
- **Transparenz- und Beschwerdemechanismen:** Nutzer*innen müssen klar informiert werden, warum sie eine Maßnahme erhalten haben.

Benutzerfreundliche Möglichkeiten zum Widerspruch sind, wie vom DSA gefordert, zu gewährleisten.

- **Spezifischer Schutz in Video-/Social-Media-Umgebungen:** Sicherstellung, dass die Richtlinie über audiovisuelle Mediendienste (Audiovisual Media Services Directive/ AVMSD) als zentraler Rechtsrahmen der Europäischen Union für audiovisuelle Inhalte (also **Fernsehen, Video on Demand-Dienste und Video-Sharing Plattformen** wie YouTube oder TikTok) zuverlässig und nachweisbar in nationales Recht umgesetzt, bzw. angewendet wird. Hier ist zu gewährleisten, dass Minderjährige wesentlich umfangreicher vor Hass und Anstiftung geschützt werden.
- **Ausweitung von Gegenrede-Elementen in-product** (Kontextboxen, verlässliche Quellen, „Learn more“)
- **Ausweitung und strukturelle Finanzierung des Safer-Internet Partner-Netzwerkes** in Deutschland unter Gewährleistung ausreichender Infrastrukturen für Beratungs- und Unterstützungsangebote, Hotlines **Partner. -Partner**
- Altersgerechte Medienkompetenz-Programme ab der Grundschule als Teil des Curriculums mit analoger Schulung des Fachpersonals.

d. Belästigung/Cybermobbing und weitere Interaktionsrisiken (z. B. Doxxing, Stalking, koordinierte Angriffe)

- **Ausweitung der „User Controls“ mit 1-Klick-Zugriff:**
Betroffene erhalten in Video- und Social-Media-Umgebungen sofort wirksame Werkzeuge, um eskalierende Situationen eigenständig zu stoppen – ohne sich erst durch komplexe Menüs klicken oder auf externe Moderation warten zu müssen. Dazu zählen schnelle Funktionen wie Blockieren und Stummschalten (Muting) ebenso wie feingranulare Einstellungen für Kommentare und Direktnachrichten (z. B. nur von bestätigten Kontakten). So wird verhindert, dass Belästigung, Hate oder Gruppendruck weiter „durchrutschen“. Besonders wichtig ist außerdem, dass Gruppeneinladungen nur mit ausdrücklicher Zustimmung möglich sind. Dadurch können Jugendliche nicht gegen ihren Willen in Gruppen gezogen werden, in denen Mobbing, Bloßstellung oder koordinierte Angriffe entstehen. Insgesamt verlagern solche 1-Klick-Controls die Kontrolle zurück zu den Nutzer*innen und reduzieren Mobbing-Dynamiken individuell, unmittelbar und ohne zusätzliche Hürden.

- **Technische Auffindbarkeit reduzieren:** Weniger *Discovery*-Mechanismen, die Inhalte oder Profile sichtbar machen, z. B. durch Empfehlungs-Feeds („For You“, „Explore“, „Reels“, „TikTok FYP“), Account-Vorschläge („Personen, die du vielleicht kennst“), Hashtag und Standortsuche-Empfehlungen auftaucht.
- **Nachhaltige Einschränkung der algorithmische Reichweitenverstärkung:** Bei Minderjährigen sollten diese Pfade **stark eingeschränkt** sein, damit Fremde Jugendliche nicht einfach „entdecken“ können, riskante Inhalte weniger leicht algorithmisch verstärkt werden, das Profil nicht in Suchmaschinen oder Plattform-Empfehlungen auftaucht, keine breite unbeabsichtigte Sichtbarkeit entsteht. Dies trägt zu einem stärkeren Schutz gegen Mobbing, Nachstellung, Belästigung und unerwünschter Ansprache bei.
- **Mehr Anti--Brigading- und Koordinationsschutz,** also Ausdehnung von Schutzmechanismen, die eingreifen, wenn sich digitale Angriffe nicht aus einer einzelnen Quelle ergeben, sondern **koordiniert, massenhaft und oft plötzlich** auftreten – etwa wenn viele Accounts gleichzeitig jemanden mentionen, melden oder mit aggressiven Kommentaren überziehen. Um solche „PileOn“-Momente und Cyberbullying-Wellen-Wellen einzudämmen, setzen Plattformen bereits technische Maßnahmen wie **Rate Limits**, also Obergrenzen für schnelle, massenhafte Interaktionen, sowie einen temporären **Slow Mode**, der das Kommentieren verlangsamt und Eskalationen direkt abbremsst. Hier sollten die Schwellenwerte von unabhängigen Kinderschutzexpert*innen konstant überprüft werden.
- **Überprüfbare Anwendung und Verfeinerung von sog. Verdachts--Clustern:** Wenn z. B. ungewöhnlich viele gleichzeitige Reports oder Mentions zu verzeichnen sind, sind automatische Schutzschirme zu aktivieren, wie etwa durch das zeitweilige Einschränken eingehender Interaktionen oder das Herausfiltern koordinierter Angriffe. So werden Nutzer*innen in Momenten hoher sozialer Anfeindung unmittelbar entlastet, während gleichzeitig verhindert wird, dass orchestrierte Übergriffe an Reichweite gewinnen oder sich weiter aufschaukeln.
-
- **Schnellere Reaktionszeiten und besserer Betroffenenschutz:** Plattformen müssen stärker verpflichtet werden nicht nur zügig auf Meldungen zu reagieren, sondern Betroffenen auch sofort zugängliche Sicherheits-Tools bereitzustellen – etwa Funktionen zur **Beweissicherung**, damit wichtige Inhalte für spätere Schritte gesichert werden können. Verpflichtende Bereitstellung eines „**Panic Button**“, um auf Knopfdruck persönliche Daten schützen zu können, Sichtbarkeit zu reduzieren und Interaktionen temporär zu blockieren. Stärkere Nachweisführung, ob Plattformen konsequent gegen **Doxxing** vorgehen, also persönliche Informationen schnell entfernen und die Weiterverbreitung unterbinden. Insgesamt sollen Betroffene in akuten Stresssituationen rasch entlastet und ihre Privatsphäre wirksam geschützt werden.

- **Zeitnahe nationale Umsetzung von EU-Aktionsplan gegen Cyberbullying** (für 2026 angekündigt) unter Federführung oder Beteiligung des BMBFSFJ zur Sicherstellung, dass insbesondere Kinder und Jugendliche stärker geschützt werden.

e. Selbstgefährdungsrisiken (z. B. Selbstschädigung/Suizid, Essstörungen, gefährliche Challenges),

- **Krisenprotokolle** sollten sicherstellen, dass bei Hinweisen auf akute Suizidalität oder Selbstverletzung sofort **prominent Hilfsangebote eingeblendet** werden. Gleichzeitig muss eine **priorisierte Moderation** greifen, die solche Inhalte und Situationen schneller prüft und entschärft. Je nach Rechtslage und Plattform -Policy können ergänzend „**Wellbeing Checks-Checks**“ aktiviert werden, um betroffene Personen unmittelbar zu schützen und geeignete Unterstützung einzuleiten.
- **Stärkere Überprüfung von Downranking-Mechanismen:** Schädliche Inhalte, die Selbstgefährdung fördern – etwa Anleitungen, Challenges oder riskante Trends – sollten konsequent herabgestuft (Downranking) oder vollständig entfernt werden, um ihre Reichweite unmittelbar zu begrenzen. Ergänzend braucht es wirksame Such- und Hashtag-Interventionen: Problematische Suchanfragen dürfen nicht zu gefährlichen Inhalten führen, sondern sollten gezielt auf sichere Informationen, Präventionsangebote und Hilfsstellen verweisen. Inhalte mit Risiken für Wohlbefinden und psychische Gesundheit müssen als zentraler Bestandteil des Schutzrahmens für Minderjährige besonders streng reguliert werden.
- **Empfehlungssysteme gegen „Rabbit Holes“:** Empfehlungssysteme sollten der Entstehung von „Rabbit Holes“ aktiv entgegenwirken, indem sie Inhalte stärker diversifizieren und wiederholte riskante Themenpfade automatisch begrenzen. Nach intensiven oder sensiblen Konsumphasen können kurze Pausen oder sanfte Unterbrechungen eingesetzt werden, um ein Abrutschen in einseitige und potenziell schädliche Feed-Muster zu verhindern. So bleibt die Nutzung ausgewogener und sicherer – insbesondere für Minderjährige.
- **Niedrigschwellige und lokal verfügbare Hilfsangebote** – etwa Hotlines, Chats oder regionale Beratungsstellen – sollten für Kinder, Jugendliche und Eltern jederzeit gut auffindbar sein. Ergänzend braucht es **zielgruppenspezifische Informationsressourcen**, etwa für Eltern oder Peers, die Orientierung in belastenden Situationen geben. Nach einem Report sollten Plattformen zudem **nachgelagerte Unterstützung** anbieten, damit Betroffene nicht allein gelassen werden und schnell passende Hilfe erhalten.

- f. kommerzielle Ausnutzung sowie Daten- und Privatsphäre-Risiken (z. B. Profiling/Tracking, irreführende Werbung, Scams)?

- Stärkere Umsetzungskontrollen für **Art. 28(2) DSA**: Für Minderjährige dürfen bereits jetzt **keine profilbasierten Werbeanzeigen** ausgeliefert werden. Dies muss stärker überprüft werden. Gleichzeitig sollten Plattformen eine „**known-minor**“-**Logik** umsetzen, die Minderjährige zuverlässig erkennt, jedoch **keine zusätzlichen Daten sammelt**, wie es **Art. 28(3) DSA** verlangt.
- **Datenschutz by default (Datenminimierung, hohe Privatsphäre Settings)**: Datenschutz sollte standardmäßig („by default“) so ausgestaltet sein, dass nur die unbedingt notwendigen Daten erhoben werden und die Privatsphäre-Einstellungen automatisch auf einem hohen Schutzniveau liegen. Ergänzend benötigen Kinder und Jugendliche transparente, altersgerechte Erklärungen, um nachvollziehen zu können, welche Daten zu welchem Zweck genutzt werden und wie sie ihre Privatsphäre eigenständig schützen können.
- **Strikte Anti Scam- und Werbeintegrität**: Plattformen sollten besonders risikoreiche Werbekund*innen durch verbindliche KYC-Prüfungen verifizieren und betrügerische Muster frühzeitig automatisiert erkennen. Ergänzend sind gezielte „Payment-Friction“-Mechanismen sowie deutlich sichtbare Warnhinweise bei verdächtigen Nachrichten, Angeboten oder Links notwendig, um Manipulation, Abzocke und andere schädliche kommerzielle Praktiken wirksam zu verhindern.
- **Stärkere Umsetzungskontrollen für AVMSD Werbeschutz für Minderjährige**: Der Werbeschutz für Minderjährige nach der AVMSD verpflichtet Anbieter dazu, Werbung so zu gestalten, dass weder die Unerfahrenheit oder Leichtgläubigkeit junger Menschen ausgenutzt noch Minderjährige in gefährliche Situationen gebracht werden. Zudem gelten klare Einschränkungen für gezielt an Minderjährige gerichtete Werbung, etwa bei Alkohol oder anderen besonders riskobehafteten Produkten. Konsequente Umsetzungskontrollen sind notwendig, um sicherzustellen, dass audiovisuelle Werbung Kinder und Jugendliche nicht manipuliert, überfordert oder in ihrer besonderen Schutzbedürftigkeit beeinträchtigt.
- **Einfacher Refund-/Dispute-Prozess für Minderjährige/Eltern**: Plattformen sollten klare, niedrigschwellige Prozesse bereitstellen, die es Minderjährigen und ihren Eltern ermöglichen, fehlerhafte Käufe oder Betrugsfälle schnell und unkompliziert zu klären. Dazu gehören verständlich geführte Abläufe ohne unnötige Hürden sowie die sofortige Sperrung verdächtiger Scam-Konten. Ergänzend ist eine transparente Kommunikation erforderlich, die nachvollziehbar erklärt, was passiert ist und welche nächsten Schritte folgen. So wird der

Verbraucherschutz gestärkt und der DSA-Gedanke eines kindgerechten, sicheren digitalen Umfelds konsequent umgesetzt.

2. Welche Rolle spielen dabei konkrete Gestaltungsentscheidungen von Diensten (z.B. Default-Einstellungen, Empfehlungs- und Rankingsysteme, Sharing /Kontaktfunktionen, Interface-Design, Parental-Control-Funktionen)?

- Konkrete Gestaltungsentscheidungen tragen maßgeblich dazu bei, wie wirksam Minderjährige im digitalen Raum geschützt sind. Voreinstellungen („Defaults“) – etwa restriktive Sichtbarkeit oder sichere Privatsphäre-Standards – können Risiken von Beginn an nachweislich reduzieren. Empfehlungs- und Rankingsysteme steuern maßgeblich, welche Inhalte Minderjährige erreichen, und müssen technisch so ausgestaltet sein, dass schädliche oder ausnutzende Inhalte nicht aktiv verstärkt werden (vgl. AVMSD-Vorgaben zum Jugend- und Werbeschutz).
- Darüber hinaus beeinflussen Sharing-, Kontakt- und Interface-Funktionen unmittelbar, in welchem Ausmaß junge Nutzer*innen potenzieller Belästigung, Manipulation oder kommerzieller Ausnutzung ausgesetzt sind. Es muss stärker sichergestellt werden, dass insbesondere die altersbedingte Unerfahrenheit und Leichtgläubigkeit Minderjähriger nicht gezielt ausgenutzt werden. Designentscheidungen bestimmen damit wesentlich das tatsächliche Schutzniveau im digitalen Alltag junger Nutzer*innen und stellen folglich einen zentralen Hebel zur Stärkung des Kinder- und Jugendmedienschutzes dar.

3. Wie können Diensteanbieter algorithmisch verstärkte Risiken (z.B. durch KI-basierte Empfehlungen und Profiling) erkennen, bewerten und wirksam mindern?

- Diensteanbieter müssen verpflichtet werden, algorithmisch verstärkte Risiken im Rahmen systematischer Risiko-Audits umfassend zu prüfen. Dazu gehört insbesondere, unabhängigen Forschungszugang zu Empfehlungs- und Profilingssystemen zu ermöglichen, einschließlich der Analyse eskalierender „Rabbit-Hole“-Dynamiken sowie potenzieller Bias in algorithmischen Entscheidungsprozessen.
- Darüber hinaus sind kontinuierliches Monitoring und die Umsetzung präventiver technischer Maßnahmen erforderlich. Dazu zählen unter anderem die stärkere Diversifizierung von Empfehlungen, Begrenzungen für wiederholt ausgespielte Risikoinhalte, der Einsatz von Alters- und Kontextsignalen sowie sicherheitsorientierte Default-Priorisierungen. Entscheidend ist eine Kombination aus transparenter Modellsteuerung, Human-in-the-Loop-Prüfungen, abgestuften Interventionsmechanismen, kontinuierlicher

unabhängiger Expert*innenberatung sowie einem umfassenden, kostenfreien Zugang für Forschende.

4. Benennen Sie nach Möglichkeit Prioritäten (Top-3-Maßnahmen) und Kriterien/Indikatoren, wie die Wirksamkeit von Maßnahmen überprüft wird/werden sollte.

Priorität / Maßnahme	Kurzbeschreibung	Kriterien / Indikatoren zur Wirksamkeitsprüfung
1. Sichere Default- & Empfehlungslogiken	Restriktive, altersgemäße Standardeinstellungen, nicht-profilbasierte Empfehlungen und granulare Risikofilter für Minderjährige. Umsetzung der AVMSD	<i>Expositionsverringering:</i> Die Sichtbarkeit schädlicher Inhalte bei Minderjährigen geht zurück, was durch unabhängige Forschung regelmäßig nachzuweisen ist <i>Compliance Nachweise:</i> Die Umsetzung von Jugendschutzstandards wird dokumentiert und von unabhängigen Expertengremium beurteilt, deren Empfehlungen umzusetzen sind. -<i>Nachweise:</i>
2. Systeme zur Erkennung & Eindämmung algorithmischer Risiken	Fortlaufende, umfassende Risiko Audits, konstantes Monitoring von Risikopfaden, Erkennung von Clustern und „Rabbit Holes“, zuverlässige Diversifizierung der Feeds durch eine Kombination durch algorithmische Maßnahmen und Moderation sowie Human-in-the-loop Überprüfungen. -Audits,	<i>Messbare Verringerung riskanter Verstärkung:</i> Es werden deutlich weniger Risikoinhalte im Feed wiederholt, was durch unabhängige Forschung nachzuweisen ist. <i>Wirksamkeit von Interventionsmechanismen:</i> Der Umfang von aktiv ausgelösten Begrenzungen wird

		nachgewiesen, eskalierende Inhalte werden deutlich reduziert. Beides wird durch unabhängige Forschung nachgewiesen.
3. Transparenz & kindgerechte Kommunikation	Die automatisierte Kommunikation muss altersgerecht sein. Auf Risiken wird leicht verständlich eingegangen. Schutzfunktionen sind leicht auffindbar und zu bedienen.	<i>Verständlichkeit:</i> Die Nutzerfreundlichkeit der Sprache und Funktionen wird regelmäßig von Kommunikationsexperten und Jugendbeiräten geprüft. <i>Nutzung von Schutzfunktionen:</i> Die Anwendung der Schutzfunktionen ist nachvollziehbar zu dokumentieren.

II Plattformverantwortung, Regulierung und Durchsetzung

1. Wie sollte die Verantwortung der Plattformbetreiber gegenüber Kindern und Jugendlichen rechtlich und praktisch konkretisiert werden (z.B. Sorgfaltspflichten, Risikoanalysen, Melde- und Abhilfewege, Transparenzpflichten)?

- Plattformbetreiber sollten klar definierte Sorgfaltspflichten gegenüber Minderjährigen erfüllen. Dazu zählen verpflichtende Risikoanalysen für Inhalte, algorithmische Systeme und Kontaktfunktionen sowie die Bereitstellung schneller, leicht zugänglicher Melde- und Abhilfewege, die auch für Kinder und Jugendliche verständlich sind. Ergänzend sind regelmäßige Transparenzberichte zu den Funktionsweisen von Empfehlungssystemen, Werbung und Datenverarbeitung erforderlich, ebenso wie robuste „Protection-by-Default“-Einstellungen, die ein höheres Schutzniveau ohne aktives Zutun der Minderjährigen gewährleisten.
- Um die Integrität und Wirksamkeit dieser Berichte sicherzustellen, müssen Plattformen unabhängigen Forschenden umfassenden Zugang zu relevanten Daten und Systemen gewähren. Zudem ist sicherzustellen, dass ausreichende

finanzielle Mittel für die unabhängige Prüfung bereitstehen und zeitnah freigegeben werden. Nur so können Risikoberichte rechtzeitig veröffentlicht werden und verhindern, dass Schäden erst nach gesellschaftlichen oder politischen Eskalationen sichtbar werden – wie etwa beim deutlichen Anstieg antisemitischer Inhalte seit der Pandemie und nach dem Terrorangriff der Hamas am 7. Oktober 2023.

2. Wie lässt sich die Verantwortung der Plattformbetreiber überprüfbar ausgestalten (z.B. Audits, Reporting, unabhängige Aufsicht, Beteiligung von Kindern und Zivilgesellschaft)?

- Die Verantwortung der Plattformbetreiber muss durch verpflichtende unabhängige Audits zu Sicherheit, Moderation und algorithmischen Risiken überprüft werden. Die Ergebnisse müssen in regelmäßigen, standardisierten Reports öffentlich gemacht werden. Damit die Datenbasis der unabhängigen Aufsicht nicht zu schmal oder einseitig ist, muss der Zugang und die Finanzierung unabhängiger Forschungsteams gewährleistet werden. Die Nationalen Digital Services Koordinatoren sind mit ausreichend Personal zu besetzen, die hinreichende Gelder für Fortbildungen haben. Die Ausbildung digital-forensischer Expert*innen ist durch die Förderung von Lehrstühlen und universitären Ausbildungsgängen zu gewährleisten, die verschiedenen Fachbereiche einbeziehen. Die systematische Beteiligung von Jugendlichen, Fachexpert*innen, Forschung und Zivilgesellschaft ist eine wesentliche Voraussetzung für die Überprüfung, ob die Plattformen ihrer gesellschaftlichen und rechtlichen Verantwortung gerecht werden.

3. Wie kann die Umsetzung bestehender Vorschriften – insbesondere der Vorgaben zum Schutz Minderjähriger nach Artikel 28 DSA samt Leitlinien der Kommission – sowie des JMStV und JuSchG verbessert werden?

- Die Umsetzung bestehender Schutzvorgaben kann durch eine zuverlässige Altersverifikation deutlich verbessert werden, da reine Selbstauskünfte hierfür nicht ausreichen. Zudem sollten sichere Standardeinstellungen sowie konsequente „Safety-by-Design“-Ansätze verpflichtend umgesetzt werden. Ergänzend ist eine stärkere Kontrolle algorithmischer Empfehlungssysteme erforderlich, um spezifische Risiken für Minderjährige wirksam zu reduzieren. Auch die Durchsetzung lässt sich durch eine engere behördliche Zusammenarbeit stärken, insbesondere zwischen der EU-Ebene, den Medienanstalten und der Stelle zur Durchsetzung von Kinderrechten in digitalen Diensten (KidD) bei der Bundeszentrale für Kinder- und Jugendmedienschutz (BzKJ). Darüber hinaus sind umfassende Fort- und Weiterbildungsangebote für

Ermittlungs- und Strafverfolgungsbehörden, Richter*innen, Nachrichtendienste sowie in der juristischen Ausbildung notwendig.

- Wesentlich ist zudem der Ausbau akademischer Ausbildungsangebote in den Bereichen Cybersicherheit, hybride Demokratiegefährdungen und Kinder- und Jugendmedienschutz. Ergänzend müssen Melde- und Hilfsangebote für Betroffene systematisch ausgebaut und dauerhaft finanziert werden. Schließlich ist eine regelmäßige Aktualisierung von Risikoanalysen unerlässlich, um neue Gefährdungslagen frühzeitig zu erkennen und ihnen wirksam zu begegnen.

4. Gibt es aus Ihrer Sicht Regelungslücken im Bereich des technischen Kinder- und Jugendmedienschutzes, und welche Instrumente wären geeignet, diese wirksam zu schließen?

- Im technischen Kinder- und Jugendmedienschutz bestehen weiterhin erhebliche Regelungslücken. So fehlt es bislang an einer verlässlich umgesetzten Altersverifikation: Plattformen wie TikTok und Instagram erlauben eine Anmeldung ab 13 Jahren, verlangen jedoch lediglich die Eingabe eines Geburtsdatums, ohne eine tatsächliche Altersprüfung vorzunehmen. Gleichzeitig zeigen Beispiele wie Roblox, dass selbst bei eingeführten Altersnachweisen – etwa dem seit 2026 eingesetzten Gesichtsscan – technische Fehler und Umgehungsmöglichkeiten bestehen bleiben, sodass der Schutz Minderjähriger nicht ausreichend gewährleistet ist.
- Auch im Bereich algorithmischer Risiken mangelt es an wirksamer Regulierung, die der Entstehung problematischer Content-Spiralen systematisch entgegenwirkt. Zwar setzen mehrere EU-Mitgliedstaaten bereits strengere Vorgaben für Plattformen wie TikTok, Instagram und YouTube um, doch fehlen bislang einheitliche Standards. Safety-by-Design-Ansätze müssen verbindlich vorgeschrieben werden; sichere Default-Einstellungen sind freiwilligen Filtersystemen klar vorzuziehen. Darüber hinaus müssen technische Anti-Grooming-Mechanismen dringend flächendeckend implementiert werden. Parallel dazu ist ein Ausbau der Kapazitäten für eine wirksame Strafverfolgung erforderlich, um digitale Gefährdungen von Kindern und Jugendlichen konsequent zu adressieren.
- Während für klassische Formen des Glücksspiels Altersgrenzen und klare Schutzmechanismen gelten, existiert für digitale Äquivalente bislang kaum eine vergleichbare Schutzinfrastruktur. Es mangelt an verbindlichen Regeln für manipulative Designelemente wie Dark Patterns oder Glücksspielmechaniken, etwa Lootboxen.

5. Welche Rolle spielen starke Verschlüsselungslösungen (z.B. Ende-zu-Ende-Verschlüsselung) für den Schutz von Kindern – und welche Spannungsfelder ergeben sich mit Überwachungs- und Interventionsmöglichkeiten?

- Starke Ende-zu-Ende-Verschlüsselung schützt Kinder, indem sie ihre private Kommunikation wirksam vor unbefugten Zugriffen von außen absichert. Gleichzeitig erschwert sie es Sicherheitsbehörden, Inhalte einzusehen und sexualisierte Gewalt gegen Kinder im digitalen Raum aufzudecken. In diesem Spannungsfeld setzt die geplante EU-Initiative zur sogenannten „Chatkontrolle“ an, die vorsieht, private Nachrichten bereits vor der Verschlüsselung zu scannen. Ein solcher Ansatz würde die Sicherheitsarchitektur von Diensten wie Signal oder WhatsApp faktisch unterlaufen.
- Kritiker*innen warnen, dass derartige Eingriffe zu flächendeckender Überwachung führen und das Vertrauen in verschlüsselte Kommunikation insgesamt schwächen könnten. Dies würde nicht nur die Privatsphäre von Erwachsenen beeinträchtigen, sondern auch neue Risiken für Kinder schaffen, etwa durch eine erhöhte Anfälligkeit für Datenmissbrauch oder Sicherheitslücken. Damit bleibt ein grundlegendes Spannungsfeld zwischen dem Schutz der Privatsphäre und der staatlichen Fähigkeit zur Intervention bei schweren Straftaten bestehen.

6. Welche Rolle sollen Betriebssysteme und App-Stores für den technischen Jugendmedienschutz übernehmen (z.B. geräteweite Jugendschutzeinstellungen, einheitliche Altersstufen-/Label-Standards, Default- und Kaufbeschränkungen) und welche regulatorischen Instrumente wären dafür geeignet?

7. Welche wirtschaftlichen und betrieblichen Implikationen gehen mit einem umfassenden Jugendmedienschutz einher?

III Altersgrenzen, Altersverifikation und spezifische Schutzräume

1. Halten Sie verbindliche Altersgrenzen für bestimmte digitale Angebote für erforderlich, und wenn ja: Nach welchen Kriterien sollten diese festgelegt werden?

- Altersgrenzen, welche sich an den Kinder- und Jugendschutzrichtlinien der USK-Freigabe orientieren sowie mit dem Vorsorgeprinzip begründet werden können, sind erforderlich. Grundsätzlich sollte gelten, dass alle sozialen Medien bis 16 Jahren nur mit vorgeschriebener Zustimmung der Erziehungsberechtigten nutzbar sind. Gleichzeitig sollten soziale Plattformen je nach Altersgruppe

unterschiedlich gestaltet werden bei den algorithmischen Vorschlägen, bei der Verfügbarkeit von gefährdenden Inhalten, bei der Möglichkeit von Kontaktforderungen durch Fremde. Dies sollte durch ein Verbot von personalisierter Werbung sowie durch die Unterbindung besonders suchterzeugender Funktionen wie Push-Nachrichten und endloses Scrollen erfolgen. Da es baldmöglichst eine Lösung braucht, empfehlen wir zunächst klare Zugangsbeschränkungen nach Alter, welche sofort greifen. Diese fungieren als notwendige Interimslösung, bis die Plattformen entsprechende Anpassungen vorgenommen haben, die Implementierung des DSA von unabhängigen Instanzen kontinuierlich überprüft werden kann und Angebote nach Altersstufen zuverlässig diversifiziert sind.

- Plattformen müssen stärker verpflichtet werden, sichere Standardeinstellungen, wirksame Moderationsstrukturen, klar verständliche Meldewege sowie transparente Schutzmechanismen bereitzustellen. Dabei ist insbesondere sicherzustellen, dass die Umsetzung der geltenden Rechtsvorschriften personell ausreichend ausgestattet ist und unabhängige Forschung dauerhaft, unbürokratisch und flexibel finanziert wird, um mit der schnellen technologischen und inhaltlichen Entwicklung Schritt halten zu können. Solange diese Voraussetzungen nicht erfüllt sind, können höhere Altersgrenzen als vorübergehende Schutzmaßnahme sinnvoll sein.
- Bestehende Risiken für die mentale Gesundheit und Persönlichkeitsentwicklung müssen ebenso adressiert werden wie Gewaltverherrlichung, problematische Nutzungsanreize durch In-App-Käufe oder exzessive Spielmechaniken. Darüber hinaus sind Faktoren wie Cybermobbing, Suchtgefahren und algorithmisch verstärkte Radikalisierung systematisch zu berücksichtigen. Politisch bedarf es hierfür eines einheitlichen Kriterienmodells, das Inhalte, Interaktionen und mentale Belastungen integriert und dabei die bestehenden Maßstäbe des Kinder- und Jugendmedienschutzes sowie des Datenschutzes zugrunde legt. Anbieter, die keine ausreichenden Schutzmaßnahmen umsetzen, sollten durch höhere Altersfreigaben sanktioniert werden. Altersbewertungen müssen angesichts der dynamischen technologischen Entwicklung regelmäßig überprüft und flexibel angepasst werden.
- Gleichzeitig ist ausdrücklich festzuhalten, dass Altersgrenzen aufgrund ihrer leichten Umgehbarkeit keine Lösung für sich darstellen. Zu jedem Zeitpunkt müssen die durch die UN-Kinderrechtskonvention verankerten Rechte auf digitale Teilhabe und auf Schutz in ein ausgewogenes Verhältnis gebracht werden. Nationale Altersbeschränkungen sollten daher stets mit einer überzeugenden, umfassenden medienpädagogischen Strategie verbunden sein. Diese muss Kindern, Eltern und Fachkräften bereits ab dem Grundschulalter ermöglichen, altersbezogene und allgemeine Risiken zu erkennen, geeignete Schutzmaßnahmen anzuwenden sowie Eigenverantwortung und einen gesunden digitalen Medienumgang zu entwickeln.

- **Reine Verbote bergen die Gefahr, Kinder und Jugendliche orientierungslos zurückzulassen.** Medienpädagogische Angebote sind daher eine kosten- und personalintensive gesamtgesellschaftliche Aufgabe. Sie müssen langfristig, generationenübergreifend und niedrigschwellig angelegt sein und unabhängig von Legislaturperioden national verankert werden.

2. Wie sollte eine altersangemessene Altersverifikation rechtlich und praktisch ausgestaltet werden?

- Wir plädieren für ein zweistufiges Modell der Altersverifikation, weisen jedoch darauf hin, dass erst nach einer Pilotierungsphase abschließend beurteilt werden kann, ob sichere, kinder- und jugendschutzgerechte, manipulationsresistente und zugleich datensparsame Lösungen verfügbar sind:
 1. **Attributbasierter Identitätsnachweis über Wallet-Standards** (z. B. „über 13“, „über 16“, „über 18“), verpflichtend nach den Vorgaben von eIDAS 2.0.
 2. **Technische Age-Assurance-Methoden** (z. B. verifizierte Geräte-IDs oder elterliche Freigaben) für Kontexte, in denen ein formaler Identitätsnachweis unverhältnismäßig wäre.
- Auf diese Weise könnte ein System entstehen, das Minderjährige wirksam schützt, Plattformen Rechtssicherheit bietet und zugleich dem Prinzip der Datenminimierung entspricht. Dienste erhalten ausschließlich das erforderliche Altersattribut – nicht jedoch Name, Adresse oder Ausweisnummer –, während Eltern und Jugendliche die Kontrolle über ihre Daten behalten.
- Wir verweisen in diesem Zusammenhang auf die Ergebnisse des *Global Age Assurance Standards Summit* (2025), auf dem unterschiedliche technische Verfahren wie biometrische Schätzungen, Altersabschätzung und attributbasierte Freigaben diskutiert wurden. Zwar zeigen sich Fortschritte bei Genauigkeit und Effizienz, zugleich bestehen jedoch weiterhin relevante Schwachstellen, darunter Umgehungsmöglichkeiten und Fehlklassifikationen.

Zusammenfassend lässt sich festhalten:

- Wallet-basierte Verfahren könnten einen gangbaren Weg darstellen, da sie datensparsam und interoperabel sind.
- Bislang liegen jedoch keine abschließenden Wirksamkeitsnachweise vor.
- Eine verlässliche Altersverifikation setzt hohe Datenschutz- und Sicherheitsstandards sowie eine hohe Resistenz gegenüber Manipulation voraus. Entsprechende Pilotprojekte müssen daher transparent und unabhängig evaluiert werden.
- **Altersverifikationssysteme gehören in sichere Händen.** Sie müssen so gestaltet sein, dass der Schutz junger Menschen im digitalen Raum gewährleistet

wird, ohne dabei unnötige oder flächendeckende Erhebungen sensibler personenbezogener Daten zu erzeugen. **Dafür braucht es datenschutzkonforme, gemeinwohlorientierte und unabhängige Verifizierungsinstanzen**, die nicht den Geschäftsinteressen privater Anbieter unterliegen.

3. Welche Formen der Altersabsicherung (Age-Assurance wie Schätzung, Verifikation, Token-, Wallet-basierte Nachweise) halten Sie für geeignet und wie können Over Collection, Diskriminierung und Umgehbarkeit vermieden werden?

Siehe vorherige Antwort. Hier ist die fachspezifische Forschung gefragt.

4. Braucht es aus Ihrer Sicht speziell gestaltete digitale Räume für Kinder und Jugendliche mit begrenztem oder ausgeschlossenem Zugang für Erwachsene, und wie ließen sich solche Räume technisch, rechtlich und organisatorisch realisieren?

Nein. Erwachsene sollten immer in der Lage sein, einen umfassenden Kinder- und Jugendschutz in Online-Räumen zu gewährleisten. Gefahren können mitunter auch in digitalen Räumen entstehen, in denen nur Gleichaltrige kommunizieren.

IV Chancen, Teilhabe und spezielle Problemfelder

1. Welche Chancen für Information, Bildung, Beteiligung und Freizeit ergeben sich durch technischen Kinder- und Jugendmedienschutz, wenn er beispielsweise „by design“ in digitale Dienste integriert wird?

- Safety-by-Design Vorgaben müssen zu jeder Zeit durch den Zugang unabhängiger, umfassender und agiler Forschung sowie hinreichender Ermittlungskapazitäten abgesichert werden. Dieser notwendige Ansatz ersetzt indessen nicht unabhängige, systemisch verankerte medienpädagogische Angebote, bei denen die Zivilgesellschaft als Scharnier eine entscheidende Rolle spielen sollte. Junge Menschen benötigen geschützte Freiräume im Sinne von attraktiven, sicheren und sinnvollen Alternativen zu den aktuellen digitalen Plattformen, in denen ein Großteil ihres sozialen Lebens stattfindet. Einschränkungen ohne Alternativen führen unweigerlich zu einem Vakuum: sozial, kulturell und bildungspolitisch.
- Studien wie die JIM-Studie (mpfs, 2025) zeigen deutlich, dass soziale Medien für Jugendliche weit mehr sind als Unterhaltung. Sie dienen der Identitätsentwicklung, der Verständigung mit Gleichaltrigen, dem Zugang zu

Informationen und der politischen Meinungsbildung. Damit sind sie niedrigschwellige Lernumgebungen und zentrale Erfahrungsräume – und diese Funktionen lassen sich nicht einfach abschalten, ohne ihnen angemessene Alternativen gegenüberzustellen.

- Die zentrale Frage lautet daher: **Welche digitalen Räume wollen wir als Gesellschaft für junge Menschen schaffen – und wer ist für sie verantwortlich?** Es gibt einen klaren Bedarf an sicheren, datenschutzkonformen Räumen für soziale Interaktion, Informationsaustausch und kreative Teilhabe.
- Bleiben attraktive, sichere und europäisch verankerte Alternativen aus, droht eine Abwanderung in unregulierte, kommerziell geprägte, dezentrale oder intransparente Plattformökosysteme. Die Dynamik des rasanten Wachstums neuer Dienste – wie beim Aufstieg von TikTok nach der Schwächung von Facebook – zeigt, wie schnell solche Verschiebungen erfolgen können. Kritisch wird es insbesondere dann, wenn ein Verantwortungsvakuum entsteht und notwendiges Wissen nicht rechtzeitig in Handlung überführt wird.

2. Wie kann ein übermäßiger Medienkonsum von Kindern und Jugendlichen wirksam begrenzt werden, und welche Verantwortung tragen Anbieter hinsichtlich der Designentscheidungen, der Nutzungsdauer und des digitalen Engagements (Interaktionen)?

Siehe Antworten oben, unter anderem zu den Einschränkungen süchtig machender Designmechanismen.

- Hinzuzufügen ist, dass Altersbeschränkungen, medienpädagogische Angebote und/oder Safety-by-Design Mechanismen automatisch eine für die gesunde Entwicklung von Kindern- und Jugendlichen angemessene Balance zwischen analoger und digitaler Freizeitgestaltung schaffen. Deswegen bleibt die umfassende finanzielle Absicherung von Freizeitangeboten im Bereich Sport, Kunst, Musik, Natur und gesellschaftlichem/sozialem Engagement ebenso wichtig wie Resilienz-stärkende Programme, die die Eigenverantwortung und das Selbstbewusstsein junger Menschen stärken. Auch hier spielen zivilgesellschaftliche Angebote eine entscheidende Rolle. Dies gilt insbesondere für Schul-transformatorische Ansätze, die die Partizipation und aktive Mitgestaltung des Schulalltags von Kindern und Jugendlichen fördern.
- Medienbildung sollte deshalb von Beginn der Schullaufbahn an systematisch gefördert werden. Statt den Zugang zu Medien lediglich stark restriktiv zu regulieren, sollte der Schwerpunkt darauf liegen, junge Menschen zu befähigen, soziale Medien sachkundig und verantwortungsbewusst zu nutzen. Die Entwicklung von Medienkompetenz stellt somit den entscheidenden Faktor dar, um Gefahren zu reduzieren und Potenziale optimal zu nutzen. Dies umfasst unter anderem den kritischen Umgang mit sozialen Medien, das Wissen um deren

Chancen, aber auch um Risiken und die Interessen der Anbieter*innen und Content Creatoren dahinter, die Fähigkeit zur Erkennung von Desinformation sowie den bewussten Umgang mit Datenschutz.

3. Wie kann in Fällen von Sharenting und Family-/Kinder-Influencing einem Missbrauch von Bildern und personenbezogenen Daten technisch, regulatorisch und durch Medienbildung vorgebeugt werden?

- Eine wirksame Prävention gegen den Missbrauch von Kinderfotos in Sharenting- und Family-Influencing-Kontexten beruht auf drei zentralen Handlungsfeldern:
 - **A. Technische und gestalterische Schutzmaßnahmen:**
Dazu gehören strengere Privatsphäre-Voreinstellungen, der Einsatz automatischer Erkennungsmechanismen für sensible oder wiederholt hochgeladene Kinderbilder sowie vereinfachte, nutzerfreundliche Melde- und Löschrprozesse. So können Risiken frühzeitig erkannt und problematische Inhalte schneller entfernt werden.
 - **B. Klarere regulatorische Vorgaben:**
Die Veröffentlichung von Kinderfotos muss konsequent am Kindeswohl ausgerichtet sein, insbesondere wenn Eltern wirtschaftliche Interessen verfolgen. Hierfür braucht es strengere rechtliche Beschränkungen sowie eine stärkere und konsequentere Einbindung der zuständigen Aufsichtsbehörden.
 - **C. Stärkung von Medienbildung und digitaler Selbstbestimmung:**
Kinder dürfen nicht zur Ware in sozialen Netzwerken werden. Deshalb ist Medienbildung ein zentraler Bestandteil der Prävention. Eltern müssen besser verstehen, welche langfristigen Folgen das öffentliche Teilen privater Momente haben kann. Gleichzeitig sollten Kinder schrittweise befähigt werden, über ihre digitale Identität mitzubestimmen. Hierfür sind die bestehenden Regelungen zur digitalen Selbstbestimmung von Minderjährigen weiterzuentwickeln und zu stärken.