

Vorstudie

Sicher mit KI aufwachsen

Zwischen Algorithmen, Chatbots und Deepfakes:
Schutz, Teilhabe und Bildung für Kinder und Jugendliche

SOS KINDERDÖRFER WELTWEIT | TERRE DES HOMMES starke Kinder – gerechte Welt

Zusammenfassung

Diese Vorstudie wurde von Terre des Hommes und SOS-Kinderdörfer weltweit beauftragt und in enger Zusammenarbeit mit Ruben Hetfleisch (Klartext AI) erstellt. Ziel war es, eine praxisorientierte, evidenzbasierte Grundlage zu schaffen: Risiken und Chancen von Künstlicher Intelligenz für Kinder zu kartieren, relevante Stakeholder und Kinder-KI-Berührungspunkte zu identifizieren und Handlungsoptionen vorzuschlagen. Terre des Hommes und SOS-Kinderdörfer weltweit haben die Vorstudie als strategische Analyse angelegt, um aufzuzeigen, welche Sektoren bereits im Feld KI und Kinderrechte in welchem Maße aktiv sind, welche Lücken bestehen und welche Maßnahmen es zur Schließung dieser Lücken gibt. Die Vorstudie soll als Impuls und Ressource zum Teilen mit interessierten Stakeholdern dienen im Sinne des Sustainable Development Goals 17 – also zur vernetzten Lösung komplexer Probleme, zur Förderung von Partnerschaften, zur Bündelung von Ressourcen und zum gegenseitigen Kompetenzausbau. Neben dem Potenzial werden aber auch die Grenzen von NGOs klar benannt: Die Aufgaben lassen sich nicht allein lösen, es braucht verpflichtende Beiträge aus Politik und Tech-Industrie. Das Feld ist fragmentiert, Expertise und Maßnahmen sind über Forschung, Politik, Zivilgesellschaft und Tech-Branche verstreut, weshalb eine strukturelle Brücke zwischen Evidenz, Praxis und Regulierung nötig ist. Die Vorstudie hat verschiedene Zielgruppen im Blick: politische Entscheidungsträger:innen, Bildungseinrichtungen und Lehrkräfte, Eltern sowie NGOs, Tech-Akteure und Wissenschaft.

Methodisch basiert die Arbeit auf einer ausführlichen Literaturrecherche, halbstrukturierten Interviews mit internationalen Expert:innen aus Forschung, Zivilgesellschaft, Regierung und einigen Tech-Akteuren, einem strukturierten Expert:innen-Workshop sowie Brainstorming-Sessions mit Jugendlichen. Die Beteiligung der Tech-Branche war begrenzt, was die operativen Erkenntnisse einschränkt. Die Datenlage ist gemischt und entwickelt sich weiter, doch zeigten sich über die verschiedenen Methoden hinweg einheitliche Themen.

Kernerkenntnisse sind, dass KI bereits tief im Alltag junger Menschen verankert ist – durch Empfehlungssysteme, Personalisierung und

eingebettete Dienste – und dass dieselben Mechanismen, die Vorteile wie Zugänglichkeit, personalisiertes Lernen oder kreative Co-Creation ermöglichen, auch Manipulation, Datenschutzverletzungen, Ausschluss und neue Formen kommerzieller Ausbeutung begünstigen. Zu den wichtigsten identifizierten Risiken gehören Datenschutz und Profiling, die Exposition gegenüber schädlichen oder irreführenden Inhalten, algorithmische Verzerrungen mit disproportionalen Folgen für marginalisierte Gruppen, eine Abnahme der Lernhandlungsfähigkeit durch unreflektierte KI-Abhängigkeit sowie neue Formen der Aufmerksamkeitsgenerierung bei Kindern und die Monetarisierung ihrer Daten. Gleichzeitig bestehen bedeutende Chancen: Adaptive Lernhilfen, Barrierefreiheit (z. B. Text-zu-Sprache), Kreativitäts-Tools, die digitale Kompetenzen stärken, und Analyse-Tools für Fachkräfte. Doch um diesen Nutzen tatsächlich zu realisieren, sind menschliche Aufsicht und robuste Schutzmaßnahmen erforderlich.

Eine wiederkehrende Erkenntnis ist die Schlüsselrolle, die Erwachsene (Eltern, Lehrkräfte, Betreuungspersonen) einnehmen: sie sind entscheidend für eine sichere und nützliche Anwendung, werden aber oft unzureichend unterstützt und geschult; ein Defizit, das dringend angegangen werden muss. Auch auf Governance-Ebene besteht eine Lücke: Internationale rechtliche Vorgaben bieten einen Rahmen, doch eine kindgerechte Umsetzung findet noch nicht statt. Obwohl die Tech-Branche Design, Anreize und Ergebnisse aktiv gestaltet, externalisiert sie die Verantwortung an Nutzer:innen und hat die vulnerable Zielgruppe Kinder nicht im Blick.

Die Vorstudie ist bewusst als Einstieg konzipiert und soll NGOs, der Tech-Branche, der Wissenschaft und politischen Akteur:innen als Impuls dienen. Folgeinitiativen und Policy Making sind hochrelevant und erwünscht. Zentral ist dabei, wie in der vorliegenden Arbeit umgesetzt, dass Kinder und Jugendliche als aktive Mitgestalter:innen einbezogen und ihre Stimmen gehört werden. Die negativen Folgen von KI für die Lebensrealität von Kindern und Jugendlichen sind in ihrer Komplexität bisher kaum abschätzbar. Lösungen dürfen sich nicht auf strikte Verbote beschränken, sondern müssen einen systematischen Kompetenzaufbau für Kinder, Jugendliche, Eltern, Lehrkräfte und andere Beteiligte sowie rechtlich verankerte Schutzmaß-

nahmen kombinieren. Der Schutz von Kindern und Jugendlichen im KI-Zeitalter ist eine aktuelle Verpflichtung. Praktisches, rechtsbasiertes Handeln, kombiniert mit strategischer Interessenvertretung und sektorübergreifender Koordination, kann unmittelbare Schäden reduzieren und langfristig sicherere Systeme schaffen; die Vorstudie soll als Startpunkt und Aufruf dienen, Erwachsene zu befähigen, Kinder und Jugendliche zu stärken und institutionelle Brücken zu bauen, damit Politik und Tech-Industrie Verantwortung übernehmen.

Inhalt

Zusammenfassung 2

1. Einleitung 6

2. Methodik 8

3. KI-Klassen 11

4. Berührungspunkte junger Menschen mit KI 13

   4.1. Wie und wo Kinder und Jugendliche KI begegnen 13

   4.2. Chancen 18

   4.3. Risiken 22

   4.4. Schutzmaßnahmen 29

5. Stakeholder in der Verantwortung 33

6. Identifizierte Strategielücken und Handlungsfelder 37

   6.1. Strategische Lücken 37

   6.2. Handlungsimpulse 38

7. Fazit 45

   7.1. Zentrale Erkenntnisse 45

   7.2. Empfehlungen 47

8. Literaturverzeichnis 49

9. Anhang 59

Impressum
Autor: Ruben Hetfleisch | **Redaktion:** Aimy Ogaeri, Magdalena Firchau, Olivier Moliner, Helin Keles, Tanja Abubakar-Funkenberg, Jonas Schubert, Verena Milasta, Kaija Landsberg | **Bildnachweise:** portishead1/E+ via Getty Images, Privat (11) | **Grafik:** Ulrike Lang

Die Broschüre mit den zentralen Ergebnissen der Vorstudie finden Sie auf:
www.sos-kinderdoerfer.de/ki-kinderschutz
www.tdh.de/sicher-mit-ki

1. Einleitung

Das Thema KI und Kinder entzieht sich einer klaren Einordnung. Es ist weder ein Nischenthema noch ein rein politisches Anliegen, sondern ein gesamtgesellschaftliches Thema, das Bildung, Familienleben, soziale Interaktion und private Nutzung miteinander verbindet. KI ist kein isoliertes Phänomen: Sie durchdringt nahezu jeden Aspekt des Lebens von Kindern und Jugendlichen und beeinflusst die Kindheit nicht mehr nur am Rande. Sie prägt aktiv, was Kinder sehen, lernen, erschaffen, verfolgen und fühlen. In den Industrieländern sind fast alle Kinder und Jugendlichen online. Weltweit nutzten im Jahr 2023 zwei Drittel der Weltbevölkerung das Internet; 15- bis 24-Jährige sind die am stärksten vernetzte Kohorte.[1] KI treibt Empfehlungsfeeds, Klassenzimmertools und Familiendienste an. Etwa 57 % der 12- bis 19-Jährigen kennen ChatGPT und viele nutzen es für Schularbeiten und Unterhaltung.[2]

Die vorliegende Vorstudie zeigt die größten Risiken und Chancen von KI für Kinder und Jugendliche auf, erläutert Schutzmaßnahmen, identifiziert Interessengruppen und gibt abschließend Handlungsimpulse. Sie basiert auf einer gezielten Literaturrecherche, einem Brainstorming mit Jugendlichen im Alter von 12 bis 18 Jahren sowie einer Umfrage, einem Workshop und halbstrukturierten Interviews mit Expert:innen.

Die Analyse stützt sich auf eine kinderrechtsorientierte Perspektive. Die UN-Kinderrechtskonvention und der General Comment 25 verlangen, dass digitale Dienste so gestaltet und geregelt werden, dass das Wohl des Kindes im Mittelpunkt steht. Internationale Leitlinien legen den Schwerpunkt auf Folgenabschätzungen im Hinblick auf Kinderrechte, „Safety-by-Design“, altersgerechte Transparenz und sinnvolle Beteiligung. Das EU-Recht spiegelt diese Normen zunehmend wider: Der Digital Services Act (DSA) erlegt Plattformen, die für

1 ITU (2023): Global offline population steadily declines to 2.6 billion people in 2023.

2 Medienpädagogischer Forschungsverbund Südwest (2024): JIMStudie 2024.

Minderjährige zugänglich sind, bestimmte Pflichten auf. Der EU-KI-Gesetzentwurf schränkt Praktiken ein, die altersbedingte Schwachstellen ausnutzen, und stuft viele Anwendungen im Bildungsbereich sowie biometrische Anwendungen als risikoreich ein.

Die Erkenntnisse zu den Auswirkungen sind gemischt, deuten aber in eine klare Richtung. KI verstärkt bestehende Risiken – Empfehlungsalgorithmen erhöhen die Konfrontation mit schädlichen Inhalten, Deepfakes und „Nudification" erhöhen das Risiko für Rufschädigung und sexuellen Missbrauch und automatisierte Systeme können ausgeklügelte Formen von Grooming, Betrug und Überwachung ermöglichen. Langzeitstudien bringen eine intensivere Nutzung sozialer Medien im frühen Jugendalter mit stärkeren depressiven Symptomen in Verbindung. Cybermobbing sowie die Konfrontation mit sexualisierten Inhalten nehmen zu.

Gleichzeitig kann KI messbare Vorteile erzielen, wenn sie durchdacht eingesetzt wird: Adaptiver Nachhilfeunterricht, Text-zu-Sprache und Übersetzung verbessern den Zugang und Lernergebnisse, generative Tools erweitern kreative Möglichkeiten. Der Nutzen hängt maßgeblich von menschlicher Mitwirkung, strengen Datenschutzeinstellungen und einer strengen Evaluation ab. Diese Vorstudie empfiehlt daher Pilotprojekte mit Human-in-the-Loop-Designs, Leitlinien für Eltern und Pädagog:innen in einfacher Sprache sowie eine kurze Beschaffungscheckliste für kindzentrierte Tools, die mittelfristige Skalierung von Toolkits und gemeinsamen Monitoring- und Evaluierungs-Rahmenwerken sowie langfristige Investitionen in kindgerechte regulatorische Umsetzung und Forschung.

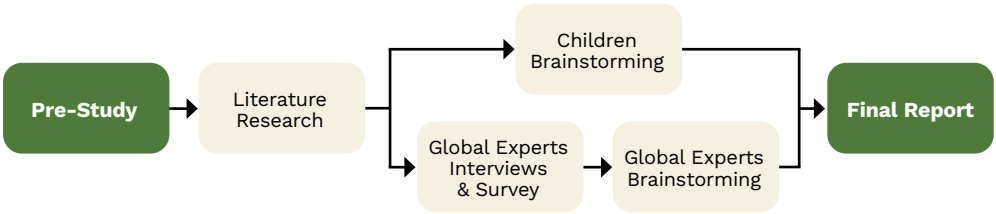
In der Praxis sollte KI, die sich an Kinder richtet, die Datenerhebung minimieren, standardmäßig einen hohen Datenschutz gewährleisten, Vorurteile bekämpfen, altersgerechte Erklärungen liefern und klare Melde- und Abhilfewege bieten. Die Umsetzung dieser Schutzmaßnahmen verläuft aber noch immer uneinheitlich.

Praktische, rechtsbasierte Maßnahmen – mit Schwerpunkt auf jugendlichen Stimmen und Erwachsenenkapazitäten – können unmittelbare Schäden verringern und im Laufe der Zeit sicherere Systeme

gestalten. Diese Systeme sind aber stets davon abhängig, dass verschiedene Sektoren – Technologiebranche, Zivilgesellschaft, Forschung und Politik – an einem Strang ziehen und gemeinsam darauf hinarbeiten, dass Kinder und Jugendliche sicher mit KI aufwachsen können.

2. Methodik

Diese Vorstudie nutzte einen mehrstufigen Prozess, der eine ausführliche Literaturrecherche mit Beiträgen von KI-Expert:innen und Jugendlichen vereint.



Die Literaturrecherche bietet eine umfassende Übersicht zu Forschung, Regulierungen, Risiken und Chancen für Kinder. Um Erkenntnisse zu gewinnen, wurden führende KI-Expert:innen aus Technologie, Regierung, Zivilgesellschaft und Forschung identifiziert und interviewt. Diese nahmen außerdem an einer strukturierten Umfrage und an einem Brainstorming-Workshop teil, um Ergebnisse und Lücken zusammenzuführen. Anschließende Workshops mit Jugendlichen erfassten gelebte Erfahrungen und stellten sicher, dass die Perspektiven junger Menschen in die Analyse einflossen.

Der Bericht ist benutzerfreundlich organisiert. Den Einstieg bildet ein Überblick über verschiedene KI-Klassen. Diesem folgt ein Abschnitt, in welchem erläutert wird, wie und wo Kinder und Jugendliche mit KI in Kontakt kommen, was die Risiken und Chancen dabei sind und welche Schutzmaßnahmen ergriffen werden sollten. Das daran anschließende Kapitel zur Stakeholder-Verantwortung klärt Akteure,

Rollen und Einfluss im Kind-KI-System. Die Ergebnisse des Expert:innen-Workshops sind zusammengefasst in Querschnittsthemen, Handlungsimpulsen und praktischen Ideen. Die Studie endet mit einem Fazit und konkreten Empfehlungen für Tech-Branche, Zivilgesellschaft, Forschung und Politik.

Literaturforschung: Die Literaturrecherche bildet die analytische Basis dieser Vorstudie. Mit Google Scholar und Perplexity wurden gezielte Zitatsuchen sowie umfassendere interdisziplinäre Abfragen durchgeführt, um bestehende Forschungsergebnisse, Klassifizierungsrahmen, Risiken, Chancen, Schnittstellen und Schutzmaßnahmen zu erfassen. Einflussreiche, repräsentative Arbeiten wurden priorisiert, um den Rahmen der Studie festzulegen und relevante Fachgebiete abzudecken. Die Ergebnisse wurden anschließend im Rahmen der Interviews validiert und auf Lücken überprüft. Das vollständige Literaturverzeichnis ist im Anhang aufgeführt.

Interviews & Umfrage mit Expert:innen: Ergänzend wurden ausführliche qualitative Interviews mit international renommierten Expert:innen aus Technologie, Regierung, Zivilgesellschaft und Forschung durchgeführt. Teilnehmende Institutionen waren u. a. Snap Inc. und AI Austria (Technologie), jugendschutz.net und die Europäische Kommission (Regierung), Mtoto News, Common Sense Media, das Deutsche Kinderhilfswerk (Zivilgesellschaft) sowie die Ludwig-Maximilians-Universität München, RAND Europe, das Nederlands Jeugdinstituut, die Università Cattolica del Sacro Cuore und die London School of Economics (Wissenschaft). Die Expert:innen-Gruppe verfügt über ein fundiertes und vielfältiges Fachwissen, das sich über verschiedene Bereiche erstreckt. Ihr gehören führende Wissenschaftler:innen, hochrangige politische Berater:innen, Verantwortliche für Plattform-Sicherheit sowie Kinderrechtsaktivist:innen an, die gemeinsam die nationale und internationale Debatte über Kinder, Jugendliche und KI prägen. Viele von ihnen haben bei internationalen Gremien und Governance-Foren mitgewirkt oder diese beraten, darunter die Vereinten Nationen, die Europäische Kommission, UNICEF und nationale Ministerien, was die globale Relevanz und Glaubwürdigkeit der dargestellten Erkenntnisse unterstreicht. Die Gruppe besteht aus fünf Forschungsexpertinnen, drei Vertreter:innen der

Zivilgesellschaft, zwei Expert:innen aus staatlichen Institutionen und zwei Experten aus der Technologiebranche, wodurch eine ausgewogene, multiperspektivische Sichtweise auf Kinder und Jugendliche und KI gewährleistet wird. Details zu den Institutionen und Expert:innen sind im Anhang erläutert.

Mehrere große Technologieunternehmen lehnten Anfragen ab. Ein Mitarbeiter eines Tech-Unternehmens, der anonym bleiben möchte, sagte dazu: **„Das ist ein so sensibles Thema, und ich bin niemand, der auch nur in Erwägung ziehen würde, das Risiko einzugehen und darüber zu sprechen."**

Alle Interviews waren einstündige, halbstrukturierte Sitzungen und folgten einem Leitfaden, der den vier Säulen (technische, elterliche, pädagogische, politische Maßnahmen) der Vorstudie entspricht. Darüber hinaus wurden Expert:innen eingeladen, an Brainstorming-Sitzungen teilzunehmen, um die Kontinuität zwischen individuellen Erkenntnissen und kollaborativer Synthese sicherzustellen.

Brainstorming-Workshops mit Expert:innen: Die Erkenntnisse wurden in einem 1,5-stündigen Online-Brainstorming-Workshop in konkrete Maßnahmen umgesetzt. Der Workshop fand über Zoom statt und stützte sich auf die Ergebnisse der Literaturrecherche und der Interviews. Die Teilnehmenden nutzten eine angeleitete Ideenfindungsrunde in Miro, um politische, pädagogische und technische Maßnahmen für eine sicherere Nutzung von KI durch Kinder und Jugendliche zu erarbeiten. Aus der Sitzung gingen konkrete Ideen hervor, die am Ende dieser Vorstudie zu finden sind.

Brainstorming-Workshops mit Jugendlichen: Es wurden mehrstufige Online-Brainstorming-Sitzungen mit Jugendlichen durchgeführt, um vielfältige globale Perspektiven einzufangen. Die Rekrutierung erfolgte über Partnernetzwerke von Terre des Hommes in Lateinamerika, Afrika und Asien. Insgesamt nahmen 19 Teilnehmer:innen (im Alter von 12–18 Jahren; 10 Mädchen, 9 Jungen) aus Deutschland, Kolumbien, Venezuela, Frankreich, Indien, Italien, den Niederlanden, Ruanda, Großbritannien und der Ukraine teil. Die Sitzungen wurden simultan gedolmetscht und folgten einer altersgerechten Agenda. Die

Jugendlichen brachten sich durch Gespräche, Schreiben, Zeichnen, Emojis und Abstimmungen ein. Die Workshops dokumentierten ihre Erfahrungen und Gefühle gegenüber KI sowie wahrgenommene Risiken und praktische Ideen für ein sichereres, gerechteres Design.

3. KI-Klassen

Im folgenden Abschnitt soll ein grober Überblick über die verschiedenen Formen von KI gegeben werden.

Generative KI: Generative Modelle erzeugen auf Abruf Texte, Bilder, Audio- und Videodateien und nutzen zunehmend Werkzeuge (indem sie Suchanfragen oder Codes aufrufen, um Fehlinformationen zu reduzieren). Für Kinder ermöglichen sie Brainstorming, Schreibübungen und kreatives Arbeiten, bergen jedoch Risiken in Bezug auf Fehlinformationen, Plagiate, Sicherheitsgrenzen und Datenlecks. Die Akzeptanz unter jungen Menschen hat zugenommen, was weltweit zu Leitlinien geführt hat, die altersgerechte Schutzmaßnahmen, die Schulung von Pädagog:innen und klare Datenrichtlinien in den Vordergrund stellen. Zu den gängigen Anwendungen gehören Nachhilfeprogramme, kreative Filter, Hausaufgabenhilfe und kindgerechte Chatbots.

Empfehlungssysteme & Ranking: Empfehlungsalgorithmen steuern Autoplay, endlose Feeds und „Das könnte dir gefallen“-Empfehlungen, indem sie das Nutzerverhalten vorhersagen. Ohne strenge Sicherheitsvorkehrungen können sie unsichere oder altersunangemessene Inhalte anzeigen und aufmerksamkeitserregende Themen an schutzbedürftige Nutzer:innen herantragen. Empfehlungssysteme dominieren soziale Unterhaltung und den Handel und prägen auch Bildungsvorschläge. Sicherere Standardeinstellungen umfassen die Drosselung riskanter Empfehlungen und die Einschränkung von Autoplay für Kinderprofile.

Vorhersage- und Analysemodelle: Vorhersageanalysen prognostizieren anhand von Learning-Management-System-Protokollen und Interaktionsdaten Ergebnisse wie Lernfortschritt, Kursabschluss oder

-abbruch. Sind sie interpretierbar, geprüft und mit menschlichem Handeln verknüpft, ermöglichen sie zeitnahe, gezielte Unterstützung (Frühwarnungen, maßgeschneiderte Interventionen). Leistung und Fairness variieren je nach Kontext, sodass der Nutzen stark abhängt von der Transparenz, von der Aufsicht durch Lehrkräfte und von Überprüfungen auf Verzerrungen.

Computer Vision und Biometrie: Bildverarbeitungsmodelle erkennen Gesichter, Körperhaltungen und Objekte und können daraus Eigenschaften ableiten; biometrische Verfahren liefern zusätzlich Schätzungen zu Identität oder Alter. Probleme bei der Erkennung von Kindergesichtern und demografische Unterschiede bergen Risiken für die Prüfungsaufsicht, Anwesenheitserfassung und Verhaltensbewertung. Angesichts der Fehlerquoten und datenschutzrechtlicher Bedenken erfordern Anwendungen mit hohem Risiko „Privacy by Design“, strenge Einwilligungsverfahren, begrenzte Speicherdauer und, wo möglich, nicht-biometrische Alternativen.

Sprach- und Übersetzungstechnologie: ASR (Automatic Speech Recognition), TTS (Text to Speech) und Übersetzungsfunktionen verbessern die Barrierefreiheit, die freihändige Suche und die mehrsprachige Unterstützung, zeigen jedoch bei spontaner Kindersprache und nicht standardmäßigen Akzenten noch Schwächen. Diese Tools, die im Unterricht, in der Therapie und bei Heimassistenten nützlich sind, erfordern die Aufsicht durch Fach- bzw. Lehrkräfte sowie Datenschutzmaßnahmen, um gerecht und wirksam zu sein.

Robotik, Smart Toys & IoT (Internet of Things): Unter IoT versteht man ein Netzwerk aus physischen Objekten, die mit Sensoren und Software versehen sind und so Daten sammeln, austauschen und miteinander kommunizieren können, oft ohne menschliches Zutun. Sozialroboter und vernetztes Spielzeug können Interaktion, Therapie und Lernen fördern, enthalten jedoch Mikrofone, Kameras und Sensoren, die Risiken für die Privatsphäre und Sicherheit bergen. Es gibt vielversprechende Hinweise auf den Nutzen bei gezielten klinischen Anwendungen in Kombination mit menschlicher Anleitung, aber Langzeitstudien und starke Schutzmaßnahmen auf den Geräten sind unerlässlich, um Datenlecks und zu großes Vertrauen zu vermeiden.

Adtech und Profiling-Systeme: Adtech sammelt Verhaltenssignale, um Werbung gezielt auszurichten und zu messen, und erzielt damit hohe Umsätze bei jugendlichen Zielgruppen. Politische Änderungen (z. B. der EU-DSA) drängen darauf, auf Profiling basierende Werbung für Minderjährige zu verbieten und sicherere Standardeinstellungen vorzuschreiben. Im Hinblick auf Kinderrechte liegen die Prioritäten darin, die Datenerhebung zu minimieren, standardmäßig ein hohes Maß an Datenschutz zu gewährleisten und in kindgerechten Umgebungen kontextbezogene statt Profiling-Ansätze zu verwenden.

4. Berührungspunkte junger Menschen mit KI

4.1. Wie und wo Kinder und Jugendliche KI begegnen

Berührungspunkte zwischen Kindern und KI können sichtbar und unsichtbar sein und sind die Momente, in denen KI die Erfahrungen von Kindern prägt: direkte Tools (Chatbots, Lernbegleiter, Spiele), Umgebungserfassung (Smart-Lautsprecher, Kameras, Wearables), algorithmische Auswahl (Feeds, Suche, Empfehlungen) und Backoffice-Entscheidungssysteme (Schulanalysen, Sozialhilfebewertung). Die Exposition ist weit verbreitet und vielschichtig: Kinder und Jugendliche können aktive Nutzer:innen, passive Beobachter:innen oder Datensubjekte sein – oft ohne zu wissen, wie Systeme mit ihren Daten umgehen oder ihre Optionen beeinflussen. Daraus folgt: Schnittstellenspezifische Risiken und Chancen lassen sich nur durch eine Kombination technischer Schutzmaßnahmen (Safeguards), pädagogischer Begleitung und institutioneller Kontrolle wirksam angehen. Expert:innen stellen fest, dass Kinder und Jugendliche KI über viele unterschiedliche, oft unsichtbare Schnittstellen wahrnehmen und nicht über einen einzigen Zugangspunkt. Die Kontaktfrequenz ist insgesamt hoch, die Sensibilität (persönliche Daten/emotionale Relevanz) ist sehr hoch, und der wahrgenommene Nutzen ist bescheiden – was die Ambivalenz gegenüber der Rolle von KI im Leben junger Menschen widerspiegelt. Die Brainstorming-Sitzungen der Jugend-

lichen brachten eine Reihe von KI-Schnittstellen zutage, die bereits fest in ihrem Alltag verankert sind. Anstatt KI abstrakt zu diskutieren, verwiesen die Jugendlichen durchweg auf konkrete Tools, Apps und Situationen, in denen sie direkt oder indirekt mit KI interagieren.

Bildung und EdTech gehören zu den prominentesten und potenziell vorteilhaftesten Berührungspunkten. Adaptive Tutoren, Schreibassistenten und Lernanalysen können das Lerntempo personalisieren, die Barrierefreiheit verbessern (Text-to-Speech, Untertitel) und in Verbindung mit einer Vermittlung durch Lehrkräfte die Lernergebnisse steigern – insbesondere für mehrsprachige, leistungsschwächere und eingeschränkte Lernende. Gleichzeitig konzentrieren sich die Risiken auf die akademische Integrität (generative Texte, mangelhafte Prüfungsaufsicht), Datenprofilierung und -speicherung sowie voreingenommene Modelle, die Ungleichheiten verstärken können. Wirksame Schutzmaßnahmen umfassen den Einsatz unter Aufsicht von Lehrer:innen, strenge Datenbeschränkungen, klare Opt-out-Möglichkeiten sowie transparente Berichts- und Einspruchsmechanismen. Metaanalysen und randomisierte Studien belegen Verbesserungen, doch die Effekte hängen stark von der Qualität der Implementierung ab; in der Praxis nutzen Schüler:innen Tools oft bewusst für Hausaufgaben und Recherchen, während im Hintergrund Profiling und Analysen laufen. Expert:innen betonen, dass Anwendungen im schulischen Bereich – Hausaufgabenhilfe, Nachhilfe, Schreibhilfen und Lernanalysen – weit verbreitet sind, oft außerhalb des formalen Unterrichts stattfinden und manchmal ohne institutionelle Begleitung erfolgen. Jugendliche bestätigen den pragmatischen Nutzen: **„Wir finden, KI ist echt hilfreich für die Hausaufgaben. Man braucht nicht stundenlang zu suchen“**, berichtet ein Mädchen (15, Italien). Während ein anderes (12, Venezuela) erklärte, dass KIs helfen, das Schulleben zu organisieren: **„Ich benutze es als eine Art Organisator.“** Diese Aussagen unterstreichen, dass der wahrgenommene Nutzen stark von Kontext, Begleitung und Transparenz abhängt.

Soziale Medien und Unterhaltungsplattformen sind die häufigsten Berührungspunkte mit KI. Empfehlungsalgorithmen, kreative Filter und integrierte Chat-Funktionen lenken Aufmerksamkeit, Stimmung und Überzeugungen – oft, ohne dass sich die Nutzer:innen dessen

bewusst sind – und können alltägliche Suchanfragen zu immer riskanteren oder auf Aussehen fokussierten Inhalten lenken. Synthetische Medien wie Deepfakes erschweren zusätzlich die Einordnung von Authentizität. Die zufällige Konfrontation mit sexualisierten Inhalten über Feeds nimmt zu, was rechtlichen Druck hinsichtlich sichererer Standardeinstellungen nach sich zieht. Kinder und Jugendliche sind hier sowohl aktiv Teilnehmende als auch häufig Zuschauende. Sie erleben diese Schnittstelle als beeindruckend, aber ambivalent und reagieren häufig mit Skepsis und manueller Überprüfung der Ergebnisse. **„Die meisten Leute nutzen ChatGPT, weil es die beliebteste KI ist … Ich sehe viel KI auf Pinterest“**, so ein Mädchen (14, Niederlande). Schutzmaßnahmen sollten sich darauf konzentrieren, die Personalisierung für Minderjährige einzuschränken, sensible Themen herunterzustufen, Krisenhinweise einzuführen und transparente Kontroll- und Meldemöglichkeiten bereitzustellen.

Smart Toys, Robotik und IoT-Geräte bieten spielerische und therapeutische Potenziale, sind derzeit jedoch eine kleinere, altersspezifische Schnittstelle. Solche Geräte enthalten oft Mikrofone, Kameras und Sensoren, die Datenschutz- und Sicherheitsrisiken bergen; prominente Vorfälle (z. B. My Friend Cayla, CloudPets) zeigen, wie unsichere Cloud-Praktiken Sprachdaten und Familiengespräche offenlegen können. Die Literatur empfiehlt Standardmaßnahmen wie lokale Verarbeitung, standardmäßig deaktivierte Mikrofone, starke Verschlüsselung, minimale Speicherdauern, das Verbot verhaltensbasierter Werbung und klare Transparenz für Eltern hinsichtlich Aufzeichnung und Nutzung. Jugendliche erwähnen vernetztes Spielzeug in ihren Diskussionen eher selten, was darauf hinweist, dass diese Schnittstellen heute weniger zentral, aber dennoch relevant für bestimmte Altersgruppen bleiben.

Berührungspunkte im häuslichen und familiären Umfeld – Sprachassistenten, Smart-TVs, Eltern-Apps und Kameras – führen schon früh zu einer Selbstverständlichkeit im Umgang mit KI. Den natürlichen und komfortablen Aspekt von KI zeigte ein Mädchen (14, Ruanda) so auf: **„Wenn du es leid bist zu tippen, kannst du sagen: ‚Meta, kannst du meiner Mutter eine Nachricht für mich schreiben?‘ Und es wird es tun.“** Diese Geräte können zwar Komfort bieten und

informelles Lernen fördern, zeichnen aber auch Interaktionen in der Umgebung auf, wozu Kinder nicht sinnvoll zustimmen können. Praktische Schutzmaßnahmen bestehen in Kindermodi, regelmäßigem Löschen von Sprachaufzeichnungen, deaktivierten Kaufoptionen, Hardware-Stummschaltungen und Familienvereinbarungen, die Geräte als öffentliche Mikrofone behandeln. Lokale Verarbeitung und strenge Aufbewahrungsfristen sind zu bevorzugen, um Datenschutzbeeinträchtigungen zu verringern. Die Expert:innen betonten, dass Kinder zu Hause oft durch das Beobachten des Verhaltens ihrer Eltern mit KI in Berührung kommen, noch bevor sie eigene Geräte besitzen.

Kinder und Jugendliche begegnen Adtech und persuasivem Design in Streamingdiensten, Suchmaschinen, E-Commerce und Apps. **Handel und Werbung** sind omnipräsent und werden von Jugendlichen meist als Hintergrund wahrgenommen – Plakatwerbung, Werbung für Handyspiele und gesponserte Beiträge. Jugendliche erkennen die Rolle der KI bei der Gestaltung kommerzieller Darbietungen an, erleben diese jedoch meist passiv. Eine Teilnehmerin (14, Deutschland) sagte: **„Ich habe auch Werbung für Marken gesehen, die von KI erstellt wurde … und auch einfach Werbung für Handyspiele mit KI.“** Forschungsergebnisse belegen manipulative Mechanismen (irreführende Aufforderungen, charakterbasierte Anstöße) und Profiling, das sich an Jugendliche richtet. Politische Kurswechsel wie das Verbot profilbasierter Werbung für Minderjährige verschärfen die Vorschriften, doch bestehen weiterhin Lücken bei der Durchsetzung. Zu den Risiken zählen verdecktes Influencer-Marketing, Lootbox-Mechanismen und dienstübergreifende Überwachung. Eine bewusst eingeschränkte Personalisierung könnte stattdessen altersgerechte Bildungsinhalte hervorheben. Zu den praktischen Schutzmaßnahmen gehören kindgerechte Modi ohne Werbung, die Deaktivierung personalisierter Werbung, Medienkompetenz in Bezug auf „gesponserte“ Inhalte und Beschaffungsfilter, um aggressiv monetarisierte Produkte zu vermeiden.

Die Schnittstellen **im Gesundheitswesen und in der Therapie** sind für Kinder weniger sichtbar, auch wenn KI in der Verwaltung oder im klinischen Bereich eingesetzt wird. KI unterstützt pädiatrische Vorsorgeuntersuchungen, Entscheidungshilfen und häusliche Interven-

tionen. Kontrollierte Studien sind vielversprechend, doch die Ergebnisse sind uneinheitlich und erfordern eine pädiatrische Validierung. Modelle, die auf Daten von Erwachsenen trainiert wurden, können bei Kindern weniger gut funktionieren; medizinische KI kann altersbedingte Fehlalarme erzeugen. Im Bereich der psychischen Gesundheit können Large-Language-Model(LLM)-Chatbots die Psychoedukation unterstützen, jedoch Risikohinweise übersehen. Der Einsatz sollte unter Aufsicht von Kliniker:innen mit Prüfpfaden und klaren Eskalationsregeln erfolgen. Das sichere Modell ist „Human-in-the-Loop“, also menschliche Einbeziehung: Kliniker:innen behalten die Entscheidungshoheit, KI liefert nachvollziehbare Vorschläge und Familien erhalten Zusammenfassungen in einfacher Sprache sowie eine strenge Verwaltung der Gesundheitsdaten. Die Expert:innen waren sich einig, dass der direkte Kontakt von Kindern und Jugendlichen mit KI im medizinischen Bereich nach wie vor begrenzt ist, wobei Interaktionen hauptsächlich über allgemeine Chatbots und nicht über spezielle medizinische Systeme stattfinden. Jugendliche nutzen KI gelegentlich situativ für Erste-Hilfe-Hinweise, doch sie sehen solche Tools nicht als Ersatz für klinische Betreuung.

Der Einsatz **im öffentlichen Dienst und in der Verwaltung** (Schulverwaltung, Sozialhilfe, Prüfungsüberwachung) kann Effizienzgewinne bringen, wirft aber erhebliche Fragen zur Transparenz, Rechenschaftspflicht und zur Diskriminierung vulnerabler Gruppen auf. Pandemiebedingte Fernüberwachung und biometrische Experimente zeigten, dass marginalisierte Schüler:innen oft unverhältnismäßig betroffen sind. Bewährte Praktiken umfassen kindgerechte algorithmische Folgenabschätzungen, eine menschliche Überprüfung bei Entscheidungen mit hohen Risiken, die Benachrichtigung der Familien und Rechtsbehelfsmöglichkeiten, den Einsatz des am wenigsten in die Privatsphäre eingreifenden wirksamen Instruments, die Validierung für Kindergruppen sowie Erklärungen zur Datennutzung in verständlicher Sprache. In diesen Systemen bestimmen institutionelle Schutzmaßnahmen und menschliche Aufsicht die Ergebnisse. Kinder sind hier hauptsächlich Zuschauende und betroffene Personen, KI-Schnittstellen im öffentlichen Dienst kamen nicht zur Sprache und administrative oder staatliche Anwendungen wurden nicht als Teil ihrer Lebenserfahrung erwähnt.

In Summe zeigen Literatur, Expert:innen und Jugendliche ein konsistentes Bild: Kinder kommen über vielfältige, oft unsichtbare Schnittstellen mit KI in Kontakt; Häufigkeit und Sensibilität sind hoch, der wahrgenommene Nutzen jedoch abhängig vom Kontext. Schnittstellenspezifische Maßnahmen müssen daher technische Standards (z. B. lokale Verarbeitung, minimale Datenspeicherung), organisatorische Regeln (Pädagogik, menschliche Aufsicht) und regulatorische Vorgaben (altersgerechte Standardeinstellungen, Werbebeschränkungen, Durchsetzungsmechanismen) verbinden. Nur so lassen sich potenzielle Vorteile, personalisierte Bildung, barrierefreie Zugänge und organisatorische Entlastung, in sichere und gerechte Praxis überführen; andernfalls drohen Überwachung, Profiling, akademische Gefährdungen und die Verstärkung bestehender Ungleichheiten.

4.2. Chancen

KI kann Kindern echte Vorteile bieten, wenn sie als Ergänzung und nicht als Ersatz für menschliche Unterstützung eingesetzt wird. Die Befunde der Literatur, die Einschätzungen der Expert:innen und die Erfahrungen der Jugendlichen ergänzen sich dabei. Die Erkenntnisse zeigen, dass die größten Fortschritte dort erzielt werden, wo Konzeption, Begleitung durch Erwachsene, institutionelle Unterstützung und eine starke Steuerung gegeben sind. Ohne diese Voraussetzungen ist der Nutzen ungleichmäßig verteilt und an Bedingungen geknüpft. Insgesamt erscheint das Versprechen der KI für Kinder zwar real, aber fragil. Es hängt von gezielter Pädagogik, sicherem Design, strenger Evaluation und Investitionen in die Kompetenzen Erwachsener ab.

Inklusive Bildung und Personalisierung bieten die deutlichsten Chancen. Adaptive Lernassistenten, angepasste Lernrhythmen und Barrierefreiheitsfunktionen wie Text-to-Speech, Untertitel oder anpassbare Layouts führen dann zu messbaren Lernerfolgen, besonders für leistungsschwächere, mehrsprachige und Lernende mit Einschränkungen, wenn sie auf den Lehrplan abgestimmt sind und von Lehrkräften begleitet werden. Metaanalysen und randomisierte Studien zeigen Verbesserungen, doch die Resultate hängen stark von der Qualität der Umsetzung ab, wobei vom tatsächlichen Niveau der Lernenden ausgegangen und Barrierefreiheit standardmäßig integriert

werden muss. Clemens Wasner, KI-Experte und führende Stimme für die gesellschaftlichen Auswirkungen Künstlicher Intelligenz, bringt die technische Seite auf den Punkt: **„Es wird nie wieder einen Text geben, den du nicht verstehst."** Expert:innen betonen, dass KI Lehrkräfte mit angemessener Ausbildung ergänzen – nicht ersetzen – sollte und dass Jugendliche KI bereits strategisch für Schulaufgaben nutzen. Frank Woithe, leitender Experte bei jugendschutz.net, fasst pragmatisch zusammen: **„Kinder und Jugendliche sind nicht dumm, sie wissen genau, wie sie mit minimalem Zeit- und Arbeitsaufwand eine gute Hausaufgabe zusammenbasteln können."** Jugendliche sahen dies genauso und betonten vor allem den Nutzen im Alltag: Zeitersparnis bei den Hausaufgaben, Tools zur Kreativitätsförderung, praktische Informationen statt abstrakter Vorteile auf Systemebene. Einige beschrieben KI als eine Möglichkeit, tiefere oder schnellere Erklärungen für Unterrichtsthemen zu erhalten. Ein 14-jähriges Mädchen aus Ruanda beschrieb es so: **„Es gibt viele forschungsbasierte Themen … manchmal schlage ich sie nach und der Begriff liefert mir nicht viele Ergebnisse, wenn ich ihn auf herkömmliche Weise suche, aber KI liefert mir tiefgreifendere oder aufschlussreichere Antworten."**

Kreativität und Ausdruck bilden einen weiteren besonders wertvollen Bereich. Generative Tools senken technische Barrieren für die Ideenfindung und Iteration, während die Urheberschaft der Schüler:innen gewahrt bleibt, wenn Aufgaben schrittweise aufgebaut werden. Es ist belegt, dass Kinder am meisten profitieren, wenn Erwachsene sie zur Reflexion und Kritik anregen. Kreative Sessions, die für den Unterricht geeignet sind und über Altersbeschränkungen, Datenschutzmaßnahmen sowie strukturierte Abläufe der gemeinsamen Gestaltung verfügen, tragen dazu bei, authentische Stimmen zu bewahren und vor übermäßiger Abhängigkeit zu schützen. Jugendliche beschreiben diese Erfahrung oft als befähigend: **„KI kann einen Drachen erschaffen, auf jede beliebige Weise"**, berichtet ein Mädchen (14, Ruanda). Sie sagt auch: **„KI könnte cool sein, um Schulen und Klassenzimmer interaktiver zu machen. Der Unterricht kann dadurch mehr Spaß machen."** Diese Einschätzung teilten auch die Expert:innen. Generative Tools senken technische Hürden und geben jungen Menschen die Freiheit, sich auf Ideen und das Erzählen von Geschichten zu konzentrieren – vorausgesetzt, die KI wird als Assis-

tent und nicht als Autor verstanden. Expert:innen wiesen darauf hin, dass diese Vorteile von zugänglichen, bildungsgerechten Tools und unterstützenden Strukturen abhängen, um die Eigenständigkeit von Schüler:innen und ihren Lernerfolg zu gewährleisten.

Anwendungen im Bereich **Sicherheit und Wohlbefinden** sind vielversprechend, erfordern jedoch Vorsicht. KI kann Muster von Cybermobbing, Grooming oder Selbstverletzung aufdecken, Mitarbeitenden helfen, sich auf die dringendsten Fälle zu konzentrieren und die Belastung von Moderator:innen und Berater:innen durch traumatisches Material verringern. Bei richtiger Anwendung unterstützt sie die Fürsorgepflicht, ohne dabei die Verhältnismäßigkeit und ein ordnungsgemäßes Verfahren aus den Augen zu verlieren. Bei unsachgemäßer Anwendung kann sie abschreckende Wirkungen, Fehlalarme und inakzeptable Datenlecks verursachen, die das Bildungsklima untergraben. Unabhängige Evaluierungen, eng gefasste Datenumfänge und eine transparente Kommunikation mit Schulen und Familien erhöhen die Legitimität und Wirksamkeit. Auch die Expert:innen bewerten die Möglichkeiten (Moderation, Altersüberprüfung, Früherkennung) als vielversprechend, aber an Bedingungen geknüpft: nützlich, um „die Nadel im Heuhaufen zu finden“, aber Transparenz, Rechenschaftspflicht und strenge Aufsicht sind notwendig, um Fehlklassifizierungen oder Übergriffe zu vermeiden. Jugendliche berichten hier vor allem von situationsbezogenen Vorteilen: schnelle Beruhigung oder sofortige Anleitung, wenn keine Erwachsenen zur Stelle waren, nicht jedoch als Ersatz für professionelle Hilfe. Ein Beispiel einer Jugendlichen (14, Deutschland) bezog sich auf praktische Erste-Hilfe-Ratschläge, die in einer beängstigenden Situation eingeholt wurden: **„Also, ich weiß, einmal wurde meine Freundin von ihrer Katze gekratzt und es blutete. Sie hatte Schwierigkeiten, ihr Handgelenk zu bewegen, und fragte ChatGPT, was sie dagegen tun könne, weil sie allein zu Hause war. Also fragte sie, was sie tun solle und wie sie die Blutung stoppen und die Wunde desinfizieren könne.“**

Einige Jugendliche nutzten KI als Gesprächsraum oder schnelle Orientierungshilfe, wenn kein Erwachsener verfügbar war. Dazu gehörte emotionale Unterstützung sowie allgemeine Beruhigung. Diese Nutzungen wurden jedoch zurückhaltend beschrieben und nicht als Ersatz für menschliche Fürsorge oder professionelle Hilfe dargestellt. Der wahrgenommene Nutzen lag eher in der Unmittelbarkeit und Zugänglichkeit als in der Zuverlässigkeit oder Autorität. Die Jugendlichen erkannten diesen unterstützenden Aspekt zwar an, gingen aber über diese konkreten Beispiele hinaus nicht auf KI als Instrument für das Wohlbefinden ein. Eine andere Teilnehmerin (16, Indien) betonte die emotionale Wirkung kreativer Anerkennung: **„Es stärkt dein Selbstbewusstsein. Wir fühlen uns gut und wertgeschätzt."**

Anwendungsfälle im Bereich **Gesundheit und Therapie** zeigen gezieltes Potenzial (Sprach- und Leseübungen, Unterstützung sozialer Kompetenzen und von Ärzt:innen angeleitete Interventionen zu Hause), doch die Ergebnisse sind uneinheitlich. Viele Systeme wurden mit erwachsenen Daten trainiert und liefern bei Kindern teils schlechtere Ergebnisse. Pädiatrische Validierung, Aufsicht durch Therapeut:innen, klare Eskalationswege und strenge Regeln für den Umgang mit Gesundheitsdaten sind daher unerlässlich. Diese Tools sollten professionelle Betreuung ergänzen, nicht ersetzen. Bei den Expert:innen stießen Anwendungen im o. g. Bereich auf die größte Skepsis. Sie lehnen Allzweck-Chatbots als Ersatz für professionelle Betreuung weitgehend ab und sehen sie allenfalls als niederschwellige Ergänzung in Versorgungslücken.

Digitale Kompetenz und kritisches Denken sind wesentliche langfristige Chancen. Eine altersgerechte KI-Kompetenz – die die Funktionsweise von Modellen, ihre Grenzen, Verzerrungen, Verifizierungs- und Herkunftsprüfungen abdeckt – sollte fächerübergreifend verankert werden, anstatt als einmalige Unterrichtseinheit behandelt zu werden. Praktische, handlungsorientierte Sequenzen und fundierte berufliche Weiterbildung der Lehrkräfte sind entscheidend, um im Unterricht Normen für die Zuordnung, akzeptable Unterstützung und Verifizierung zu etablieren und damit die akademische Integrität zu schützen.

Die **Befähigung von Eltern und Lehrkräften** ist noch unzureichend entwickelt, aber von entscheidender Bedeutung. Gut konzipierte Co-Pilot-Funktionen und Dashboards für Lehrkräfte können die Routinearbeit verringern und die formative Beurteilung ausweiten, während klare, leicht verständliche Leitlinien ihnen helfen, KI-Nutzung zu vermitteln. Um Vertrauen und Gerechtigkeit zu wahren, müssen solche Tools mit eingebautem Datenschutz, transparenten Bewertungskriterien, menschlicher Aufsicht und finanziell geförderter Fortbildung konzipiert werden. Auch die Expert:innen hoben die Bedeutung digitaler Kompetenz, kritischen Denkens und der Befähigung von Lehrkräften und Eltern hervor, betonten aber, dass sie nur dann langfristige Vorteile bringen können, wenn sie in Lehrpläne, Fortbildungen und Leitungsstrukturen eingebettet sind. Die Erfahrungen der Jugendlichen zeigen, dass der wahrgenommene Nutzen oft in der Unmittelbarkeit, Zugänglichkeit und emotionalen Unterstützung liegt, nicht zwingend in der Zuverlässigkeit oder Autorität der Systeme. Wichtig sind strukturelle Bedingungen: Ohne pädagogische Begleitung, geeignete Governance und finanzierte Personalentwicklung bleibt der Nutzen fragmentiert und kann Ungleichheiten verstärken. Expert:innen warnen vor unrealistischen Erwartungen und betonen, dass Potenzial nur durch gezielte Gestaltung, Sicherheitsvorkehrungen und die Positionierung von KI als Unterstützung zu realisieren ist.

Zusammengefasst ist das Chancenpotenzial von KI für Kinder groß, aber an Bedingungen geknüpft. Realisierbar wird es nur durch integrative Maßnahmen: pädagogisch sinnvolle Implementierung, kindgerechte Datenschutz- und Sicherheitsstandards, systematische Evaluationen mit Gleichstellungsindikatoren sowie nachhaltige Investition in die Befähigung von Lehrkräften und Eltern. Fehlen diese Voraussetzungen, drohen die Vorteile ungleich verteilt zu bleiben oder sogar neue Schäden und Ungleichheiten zu verursachen.

4.3. Risiken

In diesem Kapitel bezieht sich der Begriff „Risiko“ („Möglichkeit eines Schadens“) auf die Möglichkeit, dass KI-Systeme negative Auswirkungen auf die Sicherheit, das Wohlergehen oder die Rechte eines Kindes haben. Risiko (engl. risk) unterscheidet sich von Schaden

(engl. harm): Schaden ist das eingetretene negative Ergebnis, Risiko ist die Wahrscheinlichkeit und potenzielle Schwere solcher Ergebnisse unter Unsicherheit. Hier werden Risiken im Hinblick auf potenzielle Schäden in den Bereichen Inhalt, Kontakt, Verhalten, Entwicklung, Gerechtigkeit, Überwachung, Daten und Recht betrachtet. Aus der Fachliteratur, Expert:innen-Interviews, der Umfrage und den Jugendworkshops ergibt sich eine einheitliche Schlussfolgerung: KI verursacht selten völlig neue Gefahren, sondern verstärkt bestehende – in Bezug auf Ausmaß, Geschwindigkeit, Personalisierung und Undurchsichtigkeit. Expert:innen stuften die Exposition und die Schädlichkeit als hoch ein, während Schutzmaßnahmen insgesamt als unzureichend bewertet werden; daraus ergibt sich eine erhebliche Schutzlücke.

Content Harms (inhaltsbezogene Schäden) sind besonders präsent. Kinder begegnen KI-generierten Falschinformationen, Deepfakes sowie sexualisierten oder gewalttätigen Inhalten in Feeds, Suchmaschinen, Chats und kreativen Tools, oft ohne die Möglichkeit, Herkunft oder Authentizität zu prüfen. Kurzvideos und KI-gestützte Funktionen sind bei 8- bis 17-Jährigen weit verbreitet und Deepfakes werden immer überzeugender, während die schulische Beratung damit kaum Schritt hält. Empfehlungssysteme können grenzwertige Inhalte ohne aktive Suche anzeigen und Tools wie Gesichtsaustausch oder „Nudification“ senken die Hemmschwelle für Belästigung oder Erpressung. Besonders risikoreich sind Apps, die synthetische sexuelle Bilder von Minderjährigen erzeugen, oder Deepfake-Imitationen und Bildgeneratoren, die Filter umgehen. Auch die befragten Expert:innen teilten die Einschätzung, dass KI bekannte Gefahren verstärkt. Viraj Doshi, Platform Safety Lead bei Snapchat, merkte an, dass Sicherheitsteams nun „unmittelbaren Bedrohungen“ und „sexualisierten Inhalten“ Vorrang einräumen, während Martin Fischer vom Deutschen Kinderhilfswerk warnte: **„Deepfakes, insbesondere jetzt Nudefakes von Kindern, künstlich erzeugte Darstellungen von Missbrauch, sind ein großes Problem.“** Sonia Livingstone, eine der weltweit einflussreichsten Wissenschaftler:innen für Kinderrechte im digitalen Raum, bezeichnete diese Versäumnisse als systemisch und stellte fest: **„Das sind keine Fehler – es sind Designentscheidungen.“**

Die Jugendlichen boten in den Workshops eine differenzierte, erfahrungsbasierte Sicht auf Risiken der KI, die zwar mit den Kategorien der Studie übereinstimmt, jedoch den Schwerpunkt auf Unmittelbarkeit, Kontrollverlust und konkret erlebte Schäden legt. Am meisten sorgten sie sich um irreführende oder manipulative Inhalte aus Such- und Chatsystemen – Antworten, die glaubwürdig wirken, aber falsch sein können – sowie um realistisch wirkende Fake-Bilder und Nachrichten, die Unsicherheiten verschleiern. Ein 12-jähriges Mädchen aus der Ukraine sagte dazu: **„Das ist nicht sehr sicher."**

Contact Harms (durch Kontakt verursachte Schäden) haben durch die Möglichkeiten zur Anonymisierung, Reichweitensteigerung und Erzeugung synthetischer Identitäten zugenommen. Täter:innen nutzen Stimmklone, synthetische Profile und mehrsprachige Automatisierung für Grooming, Identitätsbetrug, Sextortion und organisierte Betrugsmaschen; viele Vorfälle haben inzwischen globale und industrielle Strukturen. Betroffene berichten von schwerwiegenden emotionalen Schäden. KI ermöglicht skalierbares Grooming durch massenhaft erzeugte Deepfakes, hartnäckige Chatbots und automatisierte Übersetzungen, während die Meldesysteme der Plattformen und die Sicherheit in Direktnachrichten weiterhin lückenhaft sind, was gefährliche blinde Flecken schafft. Rebecca Wald, Expertin am Nederlands Jeugdinstituut, betont, dass Kinder oft nicht zuverlässig zwischen Bots und Menschen unterscheiden können, wodurch Manipulation und Ausbeutung erleichtert werden. Auch die jugendlichen Workshop-Teilnehmer:innen stellten fest, dass KI Menschen sehr überzeugend imitieren kann und dadurch Vertrauen und zwischenmenschliche Beziehungen untergräbt. Sie brachten dies mit Angst vor emotionaler Manipulation und Ausbeutung in Verbindung. Ein Junge (18, Großbritannien) formulierte es so: **„Ich glaube, es besteht auch im sozialen Bereich die Gefahr, dass Menschen zu gut imitiert werden. Ich denke, dass es ein ziemliches Problem sein kann, wenn zwischenmenschliche Beziehungen ersetzt werden."**

Conduct Harms (verhaltensbezogene Schäden) entstehen, da generative KI den Missbrauch unter Gleichaltrigen verstärkt: Identitätsdiebstahl, Bildmanipulation und koordinierte Belästigung beschleunigen Mobbing sowie Demütigungen und führen zu Reputations

und psychischen Schäden, die über den Online-Raum hinausreichen. Empfehlungsalgorithmen können Schlupflöcher schaffen, die schutzbedürftige Nutzer:innen wiederholt Inhalten aussetzen, die zu Selbstverletzung oder Essstörungen führen. Die akademische Integrität wird durch unvollkommene KI-Textdetektoren untergraben, die Fehlalarme und Verzerrungen erzeugen. Zu den Hochrisikoszenarien gehören gefälschte „durchgesickerte" Nacktfotos, Bot-Kampagnen gegen jugendliche Creator und ungerechtfertigte Sanktionen auf der Grundlage unzuverlässiger Detektoren. Auch die Expert:innen schätzen KI-gestützte Demütigung und Mobbing als am schädlichsten ein. Rebecca Wald bringt die Konsequenz knapp auf den Punkt: **„Kinder können online nicht sicher hinfallen und wieder aufstehen – es wird ihnen zu schwer gemacht, ihre Fehler zu korrigieren."**

Die befragten Jugendlichen sorgten sich allerdings mehr um die **kognitive Entwicklung und psychologische Folgen.** Viele befürchteten, dass eine übermäßige Abhängigkeit von KI die eigenen Fähigkeiten und unabhängiges Denken schwächen würde. Eine Teilnehmerin (14, Ruanda) befürchtete: **„Die Menschen werden immer abhängiger und nutzen ihr eigenes kritisches Denken nicht mehr. Sie verlassen sich auf KI-Tools."** Entwicklungsbezogene Schäden stehen in Verbindung mit intensiver Online-Nutzung sowie mit Gestaltungsmerkmalen (Endlos-Scrollen, variable Belohnungen), welche entwicklungsbedingte Anfälligkeiten ausnutzen. Langzeitbeobachtungen zeigen bei einigen Gruppen einen kleinen, aber signifikanten Zusammenhang zwischen längerer Bildschirmzeit und späteren depressiven Symptomen. Intensive Nutzung korreliert zudem mit Schlafstörungen und vermindertem Wohlbefinden. Problematische oder gestörte Nutzung betrifft zwar nur eine Minderheit, kann aber konzentriert auftreten. Es gibt zwar Gegenmaßnahmen wie Ruhezeiten, sanfte Unterbrechungen oder Dashboards für Erziehungsberechtigte, doch weitreichende kindzentrierte Bewertungen ihrer Wirksamkeit fehlen. Robbie Torney, Forscher zu KI-Sicherheit, bezeichnete solche Systeme als **„parasoziale Beziehungen, die Trost und Kameradschaft vortäuschen ohne die Reibung oder Vorteile echter Beziehungen"**, und warnte: **„Der Einsatz von KI-Begleitern bei Kindern ist ein bisschen wie ein unkontrolliertes medizinisches Experiment."** Expert:innen verwiesen auf weitreichende epistemische Risiken. Clemens Wasner

bemerkte: **„Ich sehe, dass KI die Art und Weise verändert, wie wir mit Informationen und dem Lernen an sich umgehen – das ist bei weitem das größte Risiko.“** Youmna Hashem, Spezialistin für verantwortungsvolle KI-Einführung, fügte hinzu: **„Ich kann mir kaum vorstellen, wie es wäre, in einer solchen Informationsumgebung aufzuwachsen.“**

Bedenken hinsichtlich **Gerechtigkeit und Voreingenommenheit** sind allgegenwärtig. Zugangsunterschiede (Geräte, Konnektivität, digitale Kompetenzen) führen zu ungleicher Nutzung und ungleichen Vorteilen. Voreingenommenheiten auf Modellebene und demografische Unterschiede – z. B. bei der Gesichtserkennung und Altersschätzung – bergen das Risiko von Fehlklassifizierungen mit realen Konsequenzen. Zugangsprobleme bleiben bestehen. Automatisierte Überwachungs- und Aufmerksamkeitserfassungs-Tools können bei neurodiversen Studierenden zu Fehlinterpretationen führen. Automatische Spracherkennung funktioniert bei kindlicher Sprache und nicht standardisierten Akzenten schlecht und Übersetzungen können kulturelle oder geschlechtsspezifische Stereotypen reproduzieren. Jugendliche bringen diese Sorgen zu Gerechtigkeit und Chancengleichheit auf Alltagsebene zum Ausdruck, wenn sie um Existenzgrundlagen fürchten. Mehrere Jugendliche befürchteten den Verlust von Arbeitsplätzen in kreativen und technischen Bereichen. Ein Teilnehmer (12, Lateinamerika) benennt die ökonomische Dimension: **„Viele Menschen verlieren durch die KI ihren Job. Ich bin ein Arbeiter und die KI wird meine Arbeit machen. Ich erhalte mein Einkommen nicht mehr.“**

Überwachung und abschreckende Effekte entstehen durch allgegenwärtiges Tracking, biometrische Experimente und automatisierte Prüfungsüberwachung. Studien finden kaum Belege dafür, dass Fernüberwachung Schummeln verringert, jedoch deutliche Hinweise auf Probleme hinsichtlich Datenschutz, Voreingenommenheit und Validität: Fehlalarme, aufdringliche Überwachung und Raumscans können normales Verhalten unterdrücken und Kreativität einschränken. Diese Sicht teilten die Jugendlichen, welche die allgegenwärtige Nachverfolgung als invasiv und kreativitätshemmend beschrieben. Ein Mädchen (14, Niederlande) erinnert sich, dass, **„als die KI auf Snapchat neu war, viele Leute darüber gesprochen haben, dass sie**

nicht wirklich sicher sei. Bei Snapchat gibt es eine Funktion, bei der du teilen kannst, wo du bist. Und einigen Leuten, die sie ausgeschaltet hatten, wurde von der KI gesagt, wo sie sind, obwohl sie diese Funktion nicht aktiviert hatten.“ Einige biometrische Anwendungen im schulischen Bereich stießen auf Widerstand seitens der Regulierungsbehörden. Neue EU-Vorschriften (DSA-Leitlinien, 2025) drängen Plattformen zu mehr Datenschutz und sichereren Standardeinstellungen für Minderjährige, doch die Umsetzung bleibt weiterhin fraglich.

Sicherheits und Datenschutzrisiken betreffen Bildungssysteme und kindorientierte Anwendungen, die sensible Daten (Identifikatoren, Gesundheitsdaten, Noten) speichern. Lieferkettenverstöße, routinemäßige Weitergabe an Dritte, eingebettete Software-Development-Kits und mangelnder Datenschutz in beliebten Apps für Jugendliche führen zu anhaltenden Risiken durch Datenlecks und missbräuchliche Wiederverwendung. Die Regulierungsbehörden verschärfen die Standards, doch die Durchsetzung ist uneinheitlich, sodass Familien und Schulen unterschiedlich gut geschützt sind. Jugendliche äußerten konkrete Befürchtungen zu Datenerhebung, -speicherung und -missbrauch und stellten fest, dass Systeme Minderjährige oft nicht erkennen, was zu altersunangemessenen Reaktionen und unzureichendem Schutz führt. Auch die befragten Expert:innen hoben die Risiken in Bezug auf Gerechtigkeit, Überwachung und Daten hervor: Voreingenommene Trainingsdaten und eine undurchsichtige Profilerstellung können Stereotypen verstärken und marginalisierte Kinder benachteiligen, während die allgegenwärtige Datenerhebung Überwachung normalisiert und Meinungsäußerung einschränkt. Sie wiesen darauf hin, dass Kinder keine nennenswerte Kontrolle über ihre Daten haben und dass Regulierungslücken und eine schwache Durchsetzung unsichere Designs und Missbrauch ermöglichen. Die Gefährdung wurde als sehr hoch, die Schädlichkeit als noch höher, die Schutzmaßnahmen hingegen als schwach eingestuft. Wie Viraj Doshi zusammenfasste: **„Wenn es Risiken bei KI gibt, ist das ein gesamtgesellschaftliches Problem, das gesamtgesellschaftliche Lösungen erfordert.“** Frank Woithe pflichtete ihm bei: **„Wir werden derzeit von den technologischen Entwicklungen völlig überfordert.“**

Recht und Compliance: Das regulatorische Umfeld befindet sich im Wandel. Das EU-KI-Gesetz (2024) verfolgt einen Ansatz, der Praktiken verbietet, die Schwachstellen (einschließlich des Alters) ausnutzen, und Pflichten für risikoreiche Systeme auferlegt. Der DSA und neue Leitlinien führen zusätzliche Pflichten für Plattformen im Hinblick auf Minderjährige ein und präzisieren sicherere Standardeinstellungen. Nationale Kodizes setzen die Standards noch höher. Dennoch bestehen weiterhin Compliance-Lücken wie die Genauigkeit der Altersüberprüfung, grenzüberschreitende Datenübermittlungen oder die Nachvollziehbarkeit von Entscheidungen, die Kinder betreffen. Auch die Durchsetzungskapazitäten bleiben ungelöst, sodass Folgenabschätzungen in Bezug auf Kinderrechte, strenge Sorgfaltspflichten der Anbieter und überprüfbare Schutzmaßnahmen eingeführt werden müssen. Die Expert:innen betonten wiederholt regulatorische Lücken. Viraj Doshi wies auf weitverbreitete Missverständnisse hinsichtlich geltender Gesetze hin und verwies auf Untersuchungen, aus denen hervorgeht, dass vielen nicht bewusst ist, dass es illegal ist, mithilfe von KI-Technologie gefälschte sexuelle Darstellungen einer Person zu erstellen. Marta Kuljon (Europäische Kommission) betonte, dass die bestehenden Vorschriften und politischen Instrumente nicht in vollem Umfang umgesetzt werden, insbesondere nicht länderübergreifend. Sonia Livingstone argumentierte, dass eine schwache Durchsetzung die Kombination aus unsicherem Design, unkritischer Akzeptanz und böswilligem Missbrauch begünstige. Die jungen Menschen sprachen nicht direkt über rechtliche oder regulatorische Rahmenbedingungen, diese kamen jedoch durch ihre Erfahrungen mit unzureichendem Schutz zur Sprache. Sie wiesen auf Lücken in der Governance und bei der Einhaltung von Vorschriften hin.

Insgesamt verstärkt KI bereits vorhandene Risiken in den Bereichen Inhalte, Kontakt, Verhalten, Entwicklung, Gerechtigkeit, Überwachung, Datenschutz und Recht. Daraus folgt die dringende Notwendigkeit stärkerer kindzentrierter Schutzanforderungen, einer besseren Durchsetzung, altersgerechter Standardeinstellungen und kombinierter technischer, regulatorischer sowie pädagogischer Maßnahmen zur Erkennung und Verhinderung von Missbrauch.

4.4. Schutzmaßnahmen

Dieses Kapitel fasst praktische Safeguards zusammen, die Kinder vor KI-bedingten Gefahren schützen sollen. Literatur, Expert:innen-Einschätzungen und Jugendliche liefern ein gemeinsames, konsistentes Bild: Es braucht vier sich ergänzende, parallel wirksame Ansätze – technische, elterliche, pädagogische und politische Maßnahmen –, die thematisch verwoben und schnittstellenspezifisch umgesetzt werden müssen. Kein einzelner Schritt reicht, nur ihr Zusammenspiel schafft robuste, kindgerechte Sicherheit.

Technische Schutzmaßnahmen setzen Kinderrechte in technische und produktbezogene Praktiken um. Nicht verhandelbare Grundlagen sind Safety-by-Design und Privacy-by-Design, also Konzepte, die sicherheitstechnische und datenschutzrechtliche Aspekte bereits im Entwicklungsprozess von Produkten und Systemen berücksichtigen. Diese müssen in Produktentwicklung, Bereitstellung und Wartung verankert sein. Die Literatur empfiehlt datenschutzorientierte Standardeinstellungen, Datenminimierung, altersgerechte Nutzererfahrungen und einen abgestuften Zugriff auf Funktionen, der an eine zuverlässige Altersüberprüfung gebunden ist. Außerdem sind mehrschichtige Inhalts- und Verhaltensschutzmaßnahmen, die menschliche Expertise in automatisierte Entscheidungsprozesse (Human-in-the-Loop) integrieren, ebenso zentral wie kinderrechtliche Folgenabschätzungen und Maßnahmen wie Red Teaming (Methode des ethischen Hackings, bei der die Perspektive eines Angreifers eingenommen wird, um Sicherheitslücken in Systemen und Prozessen zu identifizieren). Unabhängige Audits und Herkunftssignale für synthetische Medien sind erforderlich, um Fehlinformationen, Kontakt und Verhaltensmissbrauch, Überwachung und Profiling zu reduzieren und implizite Vorurteile aufzudecken. Die Umsetzung erfordert eine Koordination zwischen Produktentwicklung, Vertrauens- und Sicherheitsabteilungen, unabhängigen Prüfer:innen, Regulierungsbehörden und Vertreter:innen der Zivilgesellschaft und Jugend. Expert:innen betonen, dass technische Kontrollen nur dann wirksam sind, wenn sie von standardisierten Prüfzyklen, kompatiblen Normen und unabhängiger Bewertung begleitet werden. Zugleich ist auf Nebeneffekte zu achten, etwa übermäßige Filterung, datenschutzpro-

blematische Altersverifikation oder unsichere Herkunftsnachweise. Robbie Torney argumentierte, dass **„Technikprodukte, die für Kinder entwickelt werden, auch wirklich für Kinder konzipiert sein sollten und nicht nur angepasste Tools für Erwachsene“**, und wies dabei auf die Risiken hin, die mit dem Einsatz von Allzwecksystemen ohne kindgerechte Tests verbunden sind. Jugendliche forderten wiederholt stärkere Alterskontrollen, klare, aktive Zustimmungsflüsse und kinderfreundliche Modi. Eine Teilnehmerin (16, Indien) sagte: **„Es sollten Einwilligungen erforderlich sein. KI sollte nicht einfach automatisch funktionieren. Sie sollte erst die Erlaubnis der Nutzer:innen einholen. Und es braucht strengere Altersgrenzen.“** Statt undurchsichtiger Automatisierung wollen Jugendliche sichere Suchfilter, transparentes Systemverhalten und klare Grenzen sensibler Ausgaben.

Elterliche Begleitung ergänzt die Sicherheitsvorkehrungen der Plattformen, indem sie den digitalen Alltag der Kinder mitgestaltet. Die Expert:innen sehen Eltern und Betreuungspersonen als tägliche Vermittler:innen der KI-Nutzung, die dabei unterstützt werden sollten, Technologien gemeinsam mit ihren Kindern zu erleben und zu diskutieren, anstatt sie einzuschränken. Wirksame Ansätze verbinden gemeinsame Nutzung und offene Gespräche, gemeinsam vereinbarte Familienregeln, den pragmatischen Einsatz von Tools (Zeitlimits, Dashboards) sowie klare Wege zur Hilfesuche und Meldung von Vorfällen. Eine dialogische, von Herzlichkeit geprägte Begleitung fördert digitale Kompetenzen und mindert Risiken, während pauschale Einschränkungen riskantes Verhalten in den Offline-Bereich verlagern und das Lernen behindern können. Gute Begleitung vermittelt gesunde Gewohnheiten, erklärt Grundlagen (Chatbots, Empfehlungslogik, Herkunftshinweise) und nutzt die Jugend-Dashboards der Plattformen, um Daten und Einstellungen zu überprüfen. Die Begleitung sollte verhältnismäßig und altersgerecht sein. Sie verringert Schäden in Bezug auf Entwicklung, Kontakte und Fehlinformationen, darf jedoch nicht überwachen und so Vertrauen untergraben. In der Praxis sollten Eltern die tägliche Begleitung übernehmen, unterstützt durch Schulen, Gesundheits- und Sozialdienste, Plattformen und NGOs. Jugendliche formulieren dies als geteilte, ermöglichende Verantwortung und nicht als bloße Kontrolle. Junge Menschen fordern keine Verbote, sondern eine geführte, gemeinsame Nutzung: Eltern, die gute Praxis

vorleben, Grenzen diskutieren und sich engagieren. Dieser Ansatz reduziert Angst und normalisiert Gespräche über KI: **„Es liegt an den Eltern, sie zu führen [...] Und Eltern sollten bei ihren Kindern sein und sicherstellen, dass sie nicht nach unsicheren Dingen oder Inhalten suchen, die sie nicht konsumieren sollten“**, sagte ein Jugendlicher (17, Italien).

Bildung und Medien-/KI-Kompetenz fungieren als präventive Schutzmaßnahmen: Eine altersgerechte, praxisorientierte KI-Kompetenz-Schulung – die die Funktionsweise von Modellen, deren Grenzen und Verzerrungen, die richtige Formulierung von Eingabeaufforderungen sowie Verifizierungs- und Herkunftsprüfungen umfasst – vermittelt Kindern, Eltern und Lehrkräften die Fähigkeit, Ergebnisse zu bewerten und ihre Handlungsfähigkeit zu bewahren. Die berufliche Weiterbildung von Lehrkräften und schulische Ansätze sind unerlässlich; es ist sinnvoll, praktische Verifizierungsaufgaben mit klaren Richtlinien für KI-gestützte Arbeit, inklusiven Barrierefreiheitsfunktionen und eindeutigen Bewertungsrichtlinien zu kombinieren. UNESCO/UNICEF empfehlen eine menschenzentrierte, altersgerechte Integration, Datenminimierung und Offline-Ausweichmöglichkeiten. Medienkompetenz reduziert Fehlinformationen und schützt die akademische Integrität, doch Lehrpläne müssen vor übermäßiger Abhängigkeit und strafenden Erkennungsmechanismen schützen; es bedarf einer ausgewogenen Pädagogik und restaurativer Prozesse. Die Verantwortung liegt bei Bildungsministerien, Schulleitungen und EdTech-Anbietern und sollte gesellschaftlich unterstützt werden, um das Lernen über den Unterricht hinaus aufrechtzuerhalten. Dem stimmten auch die jugendlichen Workshop-Teilnehmenden zu. Ein britischer Junge (18) betonte: **„Zu sagen, dass Kinder keine KI nutzen dürfen, ist nicht unbedingt sinnvoll [...] man sollte ihnen schon von klein auf etwas darüber beibringen und es in den Unterricht integrieren.“** Die Jugendlichen legten großen Wert darauf, KI-Kompetenz in die Schulcurricula zu integrieren – Lehrverifikation, Datenschutz und kritische Nutzung –, damit Schüler:innen lernen, Ergebnisse zu bewerten und ihre Handlungsfähigkeit zu behalten. Sie sahen die Nutzung in der Schule dann als positiv an, wenn Lehrkräfte geschult werden. Insgesamt lehnten sie das Verbot zugunsten strukturierten Lernens und kritischer Praxis ab.

Richtlinien und Rahmenbedingungen legen die durchsetzbare Grundlage fest, die KI mit Kinderrechten in Einklang bringt. Internationale Leitlinien (GC 25), DSGVO und nationale Kodizes verlangen nicht nur Folgenabschätzungen zu Kinderrechten, sondern auch Datenschutz durch Standards und Transparenz, während das EU-KI-Gesetz Verpflichtungen für KI-Systeme, die als besonders risikoreich eingestuft werden, hinzufügt und ausbeuterische Praktiken verbietet. Nationale Jugendschutzregeln und -behörden wie die Bundeszentrale für Kinder- und Jugendmedienschutz reagieren auf Risiken und Beschwerden. Implementierungshandbücher unterstützen Risikomanagement, Audits und Post-Marktüberwachung, aber eine schlecht gestaltete Alterskontrolle oder die fragmentierte Einhaltung der Altersbeschränkung können dabei kontraproduktiv sein. Junge Menschen erwarten Regeln und eine Rechenschaftspflicht von Unternehmen und Regierungen, waren in den Workshops aber skeptisch, dass Entscheidungsträger:innen zuhören würden. Sie forderten, dass Erwachsene, die Regeln festlegen, ein praktisches Verständnis von KI benötigen, und baten darum, bei Politik- und Produktgestaltung konsultiert zu werden. Ihre Haltung verbindet pragmatische Forderungen nach konkreten Schutzmaßnahmen mit dem Wunsch nach Inklusion. Eine wirksame Politik verbindet daher durchsetzbare Aufgaben und unabhängige Aufsicht mit mandatierter Beteiligung junger Menschen und klaren, praktischen Leitlinien für Betreiber:innen.

Zusammenfassend lässt sich festhalten: Wirksamer Schutz von Kindern gegenüber KI ist mehrdimensional und konditional. In der Praxis müssen die vier Säulen (technische Sicherheitsvorgaben, elterliche Begleitung, pädagogische Befähigung, durchsetzbare Politik) ineinandergreifen. Beispiele zeigen die Notwendigkeit der Verzahnung: Datenschutzfreundliche Standardeinstellungen sind wirkungslos, wenn Lehrkräfte nicht wissen, wie sie sie erklären; sichere Altersmodi können umgangen werden, wenn Eltern und Schulen nicht sensibilisiert sind. Nur verzahnt lassen sich Risiken wie Fehlinformation, Kontaktmissbrauch, Überwachung, Profiling und bildungsbezogene Schäden wirksam mildern, ohne die Chancen für Inklusion, personalisiertes Lernen und kreative Teilhabe zu untergraben. Jugendliche bringen dabei konkrete, handlungsnahe Erwartungen ein: erklärbares Systemverhalten, aktive Zustimmungsmechanismen, altersgerechte

Modi statt pauschaler Verbote und eine Stimme in Politik und Produktdesign. Sie fordern Bildung, die sie befähigt, Ergebnisse zu verifizieren, und Eltern, die mit ihnen gemeinsam Regeln entwickeln, statt zu überwachen. Expert:innen und die Fachliteratur stützen diese Forderungen, indem sie auf die praktische Umsetzung, unabhängige Evaluation und die Notwendigkeit einer Jugendbeteiligung insistieren.

5. Stakeholder in der Verantwortung

Dieser Abschnitt gibt einen Überblick über die Akteure, die für die sichere Nutzung von KI durch Kinder verantwortlich sind, und betont, dass die Verantwortung nicht bei einer einzigen Institution liegt, sondern ein komplexes Zusammenspiel aller Beteiligten ist.

Kinder und Jugendliche sind zentrale Akteure. In Deutschland begegnen sie KI sowohl direkt (Chatbots, Spiele, Smart Toys) als auch indirekt (Empfehlungssysteme, automatisierte Moderation). Als Rechteinhaber:innen gemäß der UN-Kinderrechtskonvention und dem deutschen Jugendmediengesetz sollten ihre Erfahrungen als Leitlinie für Gestaltung, Steuerung und Bewertung dienen. Die nationale Debatte verdeutlicht, wie KI-Risiken (Deepfakes, Identitätsmissbrauch) und Schutzmöglichkeiten für Akteure im Jugendschutz verschiebt, während Anbieter oft schnelle Innovation über die Sicherheit von Kindern stellen, was zu Lücken bei der Durchsetzung und Gestaltung führt. Jugendliche sollten sinnvoll in Konsultationen, Mitgestaltung und Bewertungen von KI, die von Minderjährigen genutzt wird, einbezogen werden und altersgerechte digitale Resilienz, Meldemöglichkeiten sowie Einwilligungsfähigkeiten erlernen. Auch gegenseitige Unterstützung und von Jugendlichen geleitete Feedback-Kanäle tragen dazu bei, aufkommende Gefahren frühzeitig aufzudecken. Beispiele hierfür sind das *klicksafe*-Jugendpanel, das bei Kampagnen und Materialien berät, die Online-Beratung von Jugendlichen für Jugendliche bei *JUUUPORT* sowie die Programmier- und Datenkompetenzprogramme von *Jugend hackt* – Modelle, die Partizipation mit Unterstützung und Kompetenzaufbau verbinden.

Eltern, Betreuer:innen und Pädagog:innen sind die erste Instanz, wenn es darum geht, Kinder bei der Nutzung von KI zu unterstützen und zu schützen. Sie legen Erwartungen fest, wählen Tools aus, fördern sichere Praktiken, reagieren auf Vorfälle und leiten Beschwerden an Behörden oder Hotlines weiter. Beratungs- und Aufklärungsdienste stellen bereits praktische Materialien zum Thema KI für das Familienleben und Schulen bereit, darunter Informationen zur Deepfake-Aufklärung und zum kritischen Umgang mit Inhalten. Die KMK-Leitlinien von 2024 bieten eine gemeinsame Position für Länder und Schulen hinsichtlich der sicheren, angemessenen Nutzung im Unterricht und der Unterstützung durch Lehrkräfte. Eltern und Lehrkräfte vermitteln Familien Beschwerdemechanismen für illegale Inhalte oder bei Kontakt mit schädlichen Inhalten. Sie sollten vorbildliches, datenschutzbewusstes Verhalten vorleben, die altersgerechte Nutzung beaufsichtigen, bei Bedarf Einwilligungen dokumentieren, KI-Kompetenz mithilfe vertrauenswürdiger Ressourcen vermitteln, wissen, wie man Schäden meldet, und Bedenken bezüglich Plattformen oder Produkten über offizielle Kanäle melden. Zu den praktischen Hilfsmitteln gehören leicht verständliche Elternratgeber (*Elternguide.online*), nationale Kampagnen (*SCHAU HIN!*) und Unterrichtsmaterialien von *klicksafe*, die politische Vorgaben in die tägliche Praxis umsetzen.

Die deutsche **Zivilgesellschaft und NGOs** bieten umfangreiche Aufklärungs-, Beratungs-, Forschungsübersetzungs- und Beschwerdedienste an und sollten eng mit Regulierungsbehörden und Tech-Industrie zusammenarbeiten, um durch KI verstärkte Schäden zu verhindern und darauf zu reagieren. Organisationen stellen Unterrichtsmaterialien, von Jugendlichen geleitete Beratungsangebote und Hotlines bereit, die mit internationalen Netzwerken zur Entfernung von Inhalten verbunden sind, und helfen so Familien und Schulen, Risiken und Sicherheitseinstellungen schnell und glaubwürdig zu interpretieren. Sie leiten zudem Fälle und Erkenntnisse aus der Praxis an Regulierungs- und Branchenprozesse weiter und beeinflussen so Politik und Praxis. Beispiele hierfür sind die gemeinsame Internet-Beschwerdestelle von *eco* und *FSM* (über *INHOPE*) sowie *jugendschutz.net*, das Plattformrisiken überwacht und jährliche Trendberichte veröffentlicht.

Deutsche **Hochschulen und Forschungsinstitute** erheben Daten zu Online-Erfahrungen von Minderjährigen und bewerten KI-Tools, um Politik und Praxis zu unterstützen. Institute wie das Hans-Bredow-Institut (HBI) liefern repräsentative „EU Kids Online"-Daten für Deutschland, während Projekte wie SIKID Rechtswissenschaften, Governance und Entwicklungswissenschaften miteinander verbinden, um die Sicherheit in digitalen Umgebungen zu untersuchen. Forschende setzen die Ergebnisse in Leitfäden für Schulen und Ressourcen für Eltern um, entwickeln kindzentrierte Bewertungsmethoden und ethische Rahmenbedingungen und beraten Ministerien und Regulierungsbehörden. Zu den praktischen Ergebnissen zählen die nationalen Analysen des HBI, die politikorientierte Forschung des SIKID-Konsortiums und die technischen Machbarkeitsstudien des Fraunhofer SIT zum Jugendschutz.

Industrie- und Technologieunternehmen in Deutschland entwickeln und betreiben zahlreiche KI-Systeme für Kinder, bieten Technologien zur Altersüberprüfung und Moderation an und betreiben Hotlines sowie Programme zur Förderung der Medienkompetenz. Sie unterliegen den Bestimmungen des Jugendmedienschutz-Staatsvertrages (JMStV), des DSA, der DSGVO und des künftigen KI-Gesetzes und müssen Maßnahmen zu Safety-by-Design, datenschutzfreundliche Standardeinstellungen, eine zuverlässige Altersüberprüfung sowie klare Meldeverfahren integrieren. Selbstregulierungsorganisationen setzen die Pflichten der Plattformen um und unterstützen die Bearbeitung von Beschwerden, während Anbieter von Identitätsprüfungsdiensten die von den Regulierungsbehörden geforderten technischen Kontrollen durchführen. Die Bundesnetzagentur kann in ihrer Funktion als Koordinatorin für digitale Dienste die Durchsetzung von Vorschriften und die Bearbeitung von Beschwerden gegen Plattformen, die den Schutz von Minderjährigen vernachlässigen, koordinieren. Beispiele hierfür sind *IDnow* (Altersüberprüfung), *FSM* (Branchenbeschwerden und Aufklärungsarbeit), *eco* (Beschwerdeportal) und die *Deutsche Telekom* (Sensibilisierungs- und Schulprogramme).

Staatliche Institutionen: Bundes- und Landesbehörden regeln und setzen das Jugendmedienrecht, den Datenschutz und die Plattform-Governance durch. Durch die neuen Aufgaben im Rahmen des DSA

wird ihr Zuständigkeitsbereich erweitert, um KI-bedingte Risiken in großem Maßstab anzugehen. Die Bundeszentrale für Kinder- und Jugendmedienschutz legt die Pflichten des DSA im Hinblick auf Kinderrechte aus und koordiniert die Zusammenarbeit mit den Landesbehörden und Hotlines. Die Bundesnetzagentur betreibt ein Verbraucherportal und koordiniert grenzüberschreitende Maßnahmen gegen systemische Plattformrisiken, die Minderjährige betreffen. Die Kommission für Jugendmedienschutz und die Landesmedienanstalten bewerten Maßnahmen zur Altersüberprüfung und zum technischen Jugendschutz für Rundfunk und Telemedien, während der Bundesbeauftragte für den Datenschutz und die Informationsfreiheit und die Datenschutzbehörden der Länder die DSGVO durchsetzen und kinderspezifische Leitlinien veröffentlichen. Die Kultusministerkonferenz (KMK) und die Bildungsministerien geben Schulen Leitlinien zur Unterstützung von Lehrkräften, zum Datenschutz und zur sicheren Nutzung im Unterricht. Gemeinsam sind diese Stellen für die Aufsicht, die Durchsetzung, technische Standards und Leitlinien zum Schutz von Kindern in digitalen Diensten verantwortlich.

Bildungseinrichtungen vermitteln KI-Kompetenz, gewährleisten einen sicheren und gerechten Einsatz im Unterricht und wählen konforme Tools aus. Die KMK-Leitlinien helfen dabei, KI in den Unterricht zu integrieren und gleichzeitig die Integrität der Leistungsbewertung, die Transparenz und den Datenschutz zu wahren. Von den Einrichtungen wird erwartet, dass sie datenschutzfreundliche Standardeinstellungen verwenden und Regeln für KI-gestützte Arbeiten festlegen. Die Schulung des Personals, die Einbeziehung der Eltern und die Beteiligung der Schüler:innen sind unerlässlich. Forschungspartnerschaften helfen bei der Bewertung von Tools und der Einbettung ethischer Grundsätze in Pilotprojekte und Beschaffungsprozesse. Praktische Beispiele sind die KMK-Leitlinien 2024, Schul-Forschungs-Kooperationen mit HBI/SIKID und die Unterrichtsmaterialien von *klicksafe*.

6. Identifizierte Strategielücken und Handlungsfelder

6.1. Strategische Lücken

Dieses Kapitel erläutert die in der Vorstudie aufgedeckten strategischen Lücken.

1. Fragmentierung und Fehlen einer zentralen Anlaufstelle: Die Arbeit zu KI und Kindern ist auf Forschung, Politik, Nichtregierungsorganisationen und die Tech-Branche verteilt, wobei es nur selten zu sektorübergreifender Zusammenarbeit kommt. Erkenntnisse und Praxis laufen nicht zusammen, was den Lernprozess verlangsamt und zu Doppelarbeit führt. Hier braucht es eine dauerhafte Koordinierungsstelle, die Erkenntnisse bündelt, praktische Erfahrungen austauscht und diese in gemeinsame Agenden und operative Leitlinien umsetzt.

2. Diskrepanz zwischen Gefährdung und Schutz: Kinder und Jugendliche sind bereits in hohem Maße den Risiken der KI ausgesetzt, während Schutzmaßnahmen hinterherhinken. Aktuelle Maßnahmen sind oft reaktiv, fragmentarisch oder legalistisch und können mit dem Tempo und den Entwicklungstrends der KI nicht Schritt halten. Diese Lücke sollte geschlossen werden, indem koordinierte, risikobasierte Ansätze gefördert und Prävention (sicherere Gestaltung, Beschaffungsstandards), Aufklärung und zugängliche Abhilfemaßnahmen miteinander verbunden werden.

3. Unzureichende Einbindung von Kindern, Eltern und Lehrkräften: Kinder, Eltern und Pädagog:innen sind zentrale Akteure, bleiben aber an den Rand gedrängt oder unterqualifiziert und -finanziert. Jugendliche bieten praktische Einblicke und sollten aktive Partner:innen statt passive Konsument:innen sein. Eine sinnvolle Beteiligung sollte institutionalisiert und praktische Unterstützung für Eltern und Lehrkräfte (Toolkits, Schulungen, schnelle Anleitungen) skaliert werden, damit Schutzmaßnahmen wirken.

4. Mangelnde Verantwortung und Rechenschaftspflicht der Technologiebranche: Tech-Unternehmen prägen den KI-Umgang von Kindern, lagern die Verantwortung jedoch häufig aus. Mechanismen zur Rechenschaftspflicht für vorhersehbare, durch die Gestaltung bedingte Schäden sind schwach ausgeprägt. Hier fehlen glaubwürdige Vermittler:innen. Stakeholder sollten sich für „Child Rights by Design" einsetzen, Transparenz und unabhängige Prüfungen einfordern und Multi-Stakeholder-Foren zur Rechenschaftspflicht einberufen, in denen die Stimmen der Jugendlichen Gehör finden.

Interessierte und kompetente Stakeholder aus unterschiedlichen Sektoren sollten zusammenarbeiten und sich als handlungsorientierter globaler Knotenpunkt – organisierend, vermittelnd und umsetzend – positionieren, indem sie Erkenntnisse bündeln, partizipative Pilotprojekte durchführen, Leitlinien für Beschaffung und Sicherheit erstellen, Kapazitäten von Partner:innen ausbauen und Akteure aus Industrie und Politik an Rechenschaftsstandards messen. Eine Kombination von kurzfristigen praktischen Maßnahmen (Pilotprojekte, Toolkits, Material für Eltern) mit mittelfristigem Kapazitätsaufbau und langfristiger Arbeit in den Bereichen Politik und Rechenschaftspflicht kann strukturelle Lücken schließen und eine kindgerechte KI-Praxis in großem Maßstab vorantreiben.

6.2. Handlungsimpulse

Der Brainstorming-Workshop führte dazu, dass Interview- und Literaturergebnisse in konkrete Ideen für Folgemaßnahmen mündeten. Die Expert:innen waren sich einig, dass Risiken und Chancen zwar gut beschrieben sind, es aber an koordinierten, kindzentrierten Interventionen fehlt. Im Folgenden werden konkrete Vorschläge, starke sektorübergreifende Handlungsimpulse und die vielversprechendsten Follow-up-Initiativen vorgestellt, die für interessierte Leser:innen und Stakeholder einen Anstoß für eigenes Engagement bieten können.

Von Wissenslücken zu strukturierten Leitlinien: Mehrfach wurde der Mangel an zugänglichen, vertrauenswürdigen Leitlinien für Eltern und Betreuungspersonen betont; Verantwortung werde oft auf Familien abgewälzt, ohne sie zu befähigen. Eltern gehören zu den wichtigs-

ten, aber am wenigsten unterstützten Akteuren beim Umgang von Kindern mit KI. Von ihnen wird erwartet, dass sie als Gatekeeper und Vorbilder fungieren, doch oft fehlt es ihnen an praktischen, kindgerechten Anleitungen. Als unmittelbare und wirkungsvolle Maßnahme wurde ein Elternhandbuch vorgeschlagen, das in klarer, nicht-technischer Sprache verfasst ist und auf realen Anwendungsfällen aus der Vorstudie basiert. Ziel ist nicht, Eltern zu KI-Expert:innen zu machen, sondern ihr Selbstvertrauen zu stärken, gemeinsam mit ihren Kindern KI zu erleben, zu diskutieren und zu reflektieren. Das Handbuch sollte erklären, was KI ist, wo junge Menschen ihr begegnen, welche Risiken, Chancen und Schutzmaßnahmen bestehen, und einfache Praktiken für den Dialog und die gemeinsame Entscheidungsfindung aufzeigen, anstatt strikte Verbote zu erlassen. Diese Idee fand breite Unterstützung: Viraj Doshi erklärte die Notwendigkeit eines solchen Handbuches: **„Ich verstehe, dass Eltern überfordert sind. Angesichts der rasanten Entwicklung der KI-Technologie kann es für Familien eine große Herausforderung sein, diese Tools vollständig zu verstehen und sich darin zurechtzufinden. Unser Ansatz besteht darin, klaren, leicht zugänglichen Informationsquellen wie dem Family Center Vorrang einzuräumen, um nützliche Unterstützung zu bieten.“** Giovanna Mascheroni, Soziologin für digitale Medien, stimmte zu: **„Wir brauchen Bildung nicht nur für Kinder, sondern auch für Eltern, Lehrer und Pädagogen im Allgemeinen.“**

Die Expert:innen forderten Aufklärungskampagnen, die ausgewogene, reflektierte Gespräche über KI normalisieren – ohne moralische Panik oder naiven Optimismus –, sowie Materialien, die die Grenzen und Unsicherheiten von KI sowie Meldemöglichkeiten erläutern.

Konkrete Ziele: Erstellung eines weltweit anwendbaren Elternhandbuchs, das auf Forschungsergebnissen und jugendlichen Stimmen basiert; Übersetzung technischer Risiken in praktische Alltagsratschläge für Familien; Verbreitung des Handbuchs über Schulen, NGOs und Elterngruppen.

Bildung als struktureller Hebel: Bildung ist sowohl die vielversprechendste Chance für den sicheren Umgang mit KI als auch einer der am wenigsten koordinierten Einstiegspunkte in den alltäglichen KI-

Gebrauch von Kindern. Von Lehrkräften wird bereits ein kompetenter Umgang mit KI erwartet, doch Lehrpläne, pädagogische Beratung und berufliche Weiterbildung hinken hinterher. Bildung wurde als klarster Weg zur Skalierung identifiziert: In Klassenzimmern kann die sichere KI-Nutzung erprobt, bewertet und etabliert werden, während Lehrkräfte darin geschult werden, deren Auswirkungen zu vermitteln. Expert:innen empfahlen ein standardisiertes Bildungstoolkit (auf nationaler/EU-Ebene), das zentrale Lernsequenzen, Fortbildungsmodule für Lehrkräfte und Monitoring-Indikatoren bereitstellt. Dieses modulare, kindzentrierte KI-Kompetenz-Toolkit soll gemeinsame Kernprinzipien festlegen, gleichzeitig lokale pädagogische Flexibilität ermöglichen, altersgerechte Module zu KI-Grundlagen, Fehlinformation, Datenschutz und Einwilligung, emotionalen Grenzen, ethischer Nutzung und zukünftigen Fähigkeiten enthalten und die Stimmen von Kindern und Jugendlichen sowie Beispiele aus dem Alltag einbinden. Die Unterstützung von Lehrkräften steht im Mittelpunkt: Das Toolkit muss Lehrerleitfäden, praxisnahe Unterrichtsaktivitäten und Bewertungsleitlinien enthalten, damit Lehrkräfte die KI-Nutzung selbstbewusst vermitteln können. Die Einbettung des Toolkits in Lehrpläne würde eine Kompetenz-Nachverfolgung ermöglichen und Fragmentierung verringern. Entscheidend ist, dass dies als langfristige Kompetenz vermittelt wird, die mit der kognitiven und sozioemotionalen Entwicklung der Kinder wächst, und nicht als einmalige Lektion. Außerdem wurde eine zentrale Landingpage vorgeschlagen, auf der geprüfte Ressourcen zur KI-Kompetenz gebündelt werden, um Doppelarbeit zu vermeiden und Akzeptanz zu stärken. Joseph Atulo von Mtoto News, einer Plattform, die direkt mit Kindern und Jugendlichen in digitalen Räumen arbeitet, erklärte die Relevanz eines derartigen Toolkits: **„Ein Teil unserer Arbeit ist es, Kindern die Chancen von KI zu zeigen, sie aber auch auf die Risiken aufmerksam zu machen – besonders wenn die erhaltenen Inhalte nicht für sie geeignet sind.“**

Konkrete Ziele: Schaffung eines modularen, altersabhängigen KI-Kompetenz-Toolkits für Schulen; Befürwortung der Aufnahme in verpflichtende digitale Kompetenzlehrpläne; Bereitstellung von Lehrerschulungsressourcen bei Wahrung der Autonomie im Klassenzimmer; Überwachung und Berichterstattung von KI-Kompetenzergebnissen in Pilotregionen zur Unterstützung des Skalierungswachstums.

Gestaltung gesellschaftlicher Normen und öffentlicher Erwartungen: Expert:innen sprachen sich dafür aus, gesellschaftliche Normen durch koordinierte öffentliche Beteiligung zu verändern – idealerweise im Rahmen einer Kampagne, die Fragen wie „Was sollte KI für Kinder leisten?“ in den Mittelpunkt stellt, anstatt lediglich vor Gefahren zu warnen. Dabei sollten möglichst Regulierungsbehörden, NGOs, Pädagog:innen und die Tech-Branche einbezogen werden mit dem Ziel, gemeinsame Erwartungen hinsichtlich „Safety-by-Default“, akzeptabler Nutzungsformen und der Kinderrechte in digitalen Diensten zu formulieren.

Evidenzlücken durch Forschung und Pilotprojekte schließen: Aus der Symbiose fundierter Pilotprojekte mit Forschungsarbeiten entstanden nennenswerte Ideen wie ein EdTech-Forschungsprogramm zur Erprobung von KI-Tutorien unter kontrollierten, kindzentrierten Bedingungen sowie sorgfältige, von Therapeut:innen begleitete Studien zu KI-gestützten Instrumenten für die psychische Gesundheit. Die Expert:innen betonten, dass Pilotprojekte eine solide Monitoring- und Evaluierungskomponente, Gleichstellungsindikatoren und unabhängige Replikationsstudien beinhalten müssen, um als Basis für politische Entscheidungen und Beschaffungsmaßnahmen zu dienen.

Rechenschaftspflicht, Audits und Jugendbeteiligung: Ein immer wiederkehrendes Thema war systemische Rechenschaftspflicht. Die Vorschläge umfassten unabhängige Sicherheitsaudits für KI mit transparenten Kriterien, auf Einwilligung basierende Datenpools für Kinderdaten zum ethischen Modelltraining sowie auf Jugendliche zugeschnittene Sicherheitstools. Ein Schwerpunkt lag auf der Kontinuität der Beteiligung: „Rapid Response Councils“, „Creativity Labs“ und Jugendbeiräte würden Kindern einen dauerhaften Einfluss auf Produktgestaltung, Reaktion auf Vorfälle und politische Prioritäten ermöglichen, anstatt sie nur einmalig zu konsultieren.

Auf dem Weg zu struktureller und globaler Ausrichtung: Mehrere Ideen zielen auf eine Abstimmung auf Systemebene ab: die Einbindung von Folgenabschätzungen zu Kinderrechten in Regulierungsprozesse im Bereich KI, die Durchführung von „Child Rights by Design“-Workshops für die Tech-Industrie, die Einberufung einer globalen

Konferenz für politische Entscheidungsträger:innen und NGOs sowie die Einrichtung eines „Global Leadership Circle on Children & AI“. Eine solche strukturelle Abstimmung über Rechtsordnungen und Sektoren hinweg sei unerlässlich, um Pilotprojekte und Leitlinien in dauerhafte Schutzmaßnahmen umzusetzen.

Tech-Verantwortlichkeit & Child-AI-Roundtable: Die Vorstudie stellte eine strukturelle Lücke fest: Technologieunternehmen führen selten einen nachhaltigen, rechenschaftspflichtigen Dialog über die Auswirkungen der KI auf Kinderrechte. Um dies zu beheben, sollte ein wiederkehrender kuratierter Tech Circle eingerichtet werden, der ausgewählte Vertreter:innen der Branche, Kinderrechts-NGOs, Forschende und, wenn angemessen, Jugendliche zusammenbringt. Dies soll ein moderierter Rechenschaftsraum (keine PR) sein, in dem Unternehmen Designentscheidungen, Risikoabwägungen und Schutzmaßnahmen erklären und strukturiertes, rechtebasiertes Feedback erhalten. Die Sicherung der Tech-Teilnahme wird schwierig werden, aber Fortschritte werden nur mit aktiver Teilnahme aus dem Technologiesektor möglich sein. Dies unterstrich Robbie Torney: **„Ohne Technik am Tisch ist systemischer Wandel unmöglich.“** Ein weiterer Experte pflichtete ihm bei, indem er es als obligatorisch ansah, Tech-Unternehmen in die Verantwortung zu nehmen: **„Schaffen Sie eine strukturierte Rechenschaftspflicht über freiwillige Aussagen hinaus.“**

Konkrete Ziele: 2- bis 3-mal pro Jahr einen moderierten Runden Tisch zum Thema Kinder und KI veranstalten; von teilnehmenden Unternehmen verlangen, kindgerechte Funktionen, Schutzmaßnahmen und bekannte Fehlerquellen offenzulegen; strukturierte Mechanismen zur Rechenschaftspflicht schaffen, die über freiwillige Ansprüche hinausgehen (öffentliche Zusammenfassungen, Folgemaßnahmen); das Forum nutzen, um Produktstandardwerte, interne Governance und Sicherheitspraktiken im Laufe der Zeit zu beeinflussen.

Globaler Rapid Response Council für Kinder & Jugendliche zu KI: KI entwickelt sich viel schneller als Forschung, Politik oder Schutzzyklen und schafft eine anhaltende Kluft zwischen Echtzeit-Exposition und Schutzkapazität. Um diese Lücke zu schließen, sollte ein

kleiner, agiler Rapid Response Council eingerichtet werden, der aus jungen Menschen, Kinderrechtsorganisationen und Fachexpert:innen besteht, um neue Produkte, Funktionen und Vorschriften schnell zu bewerten und umsetzbare Leitlinien zu liefern, solange die Probleme noch flexibel sind. Der Rat würde Geschwindigkeit und Relevanz über lange Forschungszeiträume hinweg priorisieren, die Perspektiven von Kindern in der Governance verstärken und die langfristige Bewertung durch sofortige, evidenzbasierte Empfehlungen ergänzen. Joseph Atulo betonte dabei die Relevanz von Jugendbeteiligung: **„Wir setzen uns immer dafür ein, dass Kinder in diesen Räumen, in Podiumsdiskussionen, politischen Diskussionen und Diskussionen über KI sein sollten, weil sie direkt von den getroffenen Entscheidungen betroffen sind."**

Konkrete Ziele: Schaffung eines ständigen, jugendorientierten KI-Beratungsgremiums für Kinder; schnelle Beiträge zu neuen Technologien, Produkten und Regulierungsentwürfen liefern; prägnante und umsetzbare Empfehlungen an Programme, Regulierungsbehörden und Industrie liefern; Erkenntnisse des Rates nutzen, um langfristige Forschungs- und Pilotagenden zu fördern und zu priorisieren.

Zentrum für Kinderrechte und KI (Global Coordination Hub): Obwohl es starke Expertise in Forschung, NGOs, Politik und der Technikbranche gibt, gibt es derzeit keine neutrale Instanz, die Erkenntnisse, Praxis und Koordinierung im Bereich Kinderrechte und KI zusammenführt, was zu Doppelarbeit, langsamem Lernfortschritt und verhaltenen Reaktionen auf Systemebene führt. Der Vorschlag sieht die Einrichtung eines „Center for Child & AI Rights" als globale Koordinierungsstelle vor, die bestehende Akteure miteinander vernetzt, anstatt sie zu ersetzen. Frank Woithe unterstreicht dies: **„In all diesen einzelnen Bereichen ist es jetzt wirklich wichtig, dass die Gruppen, die in diesem Bereich arbeiten, viel stärker miteinander verbunden werden und sich mehr miteinander austauschen."** Das Zentrum könnte Forschungsergebnisse bündeln, Erkenntnisse in praktische Leitlinien umsetzen, Dialoge mit verschiedenen Interessengruppen einberufen, gemeinsame Ressourcen bereitstellen und die Kontinuität über Projekte und Regionen hinweg gewährleisten. Durch eine systematische Zusammenarbeit würde das Zentrum das kollektive

Lernen beschleunigen, Doppelarbeit reduzieren und die Kinderrechte im Zuge der zunehmenden Verbreitung von KI im Mittelpunkt halten.

Konkrete Ziele: Schaffung eines Koordinations- und Wissenszentrums mit Fokus auf KI und Kinderrechten; sicherstellen, dass Kinderrechte in KI-Design, -Implementierung und -Aufsicht integriert sind; Vernetzung von Forschung, NGOs, Regierungen und Industriepartnern in einem nachhaltigen Forum; Toolentwicklung und -evaluation.

Insgesamt zeigen die Ergebnisse eine klare Verschiebung von der Problemdiagnose hin zur Lösung. Die Ideen sind ambitioniert und praktisch begründet und zielen darauf ab, empirische Lücken zu schließen, Governance-Versäumnisse zu beheben und die Perspektiven der Jugendlichen zu verankern. Die Erkenntnisse bestätigen, dass die Vorstudie als Ausgangspunkt für koordinierte, wirkungsvolle Folgemaßnahmen dienen kann – vorausgesetzt, Jugendstimmen bleiben zentral, Erwachsene werden befähigt und eine strukturelle Brücke zwischen Forschung, Praxis, Politik und Tech-Branche wird gebaut.

7. Fazit

Dieses Kapitel fasst die auf der Grundlage von Literatur, Jugendworkshops, Expert:innen-Interviews und Umfrage gewonnenen Erkenntnisse zusammen und leitet daraus konkrete Empfehlungen für die Praxis ab. Eine zentrale Erkenntnis ist, dass die Arbeit auf verschiedene Sektoren verteilt ist, was zu Doppelarbeit, Wissenslücken und schwachen Verbindungen führt, insbesondere zwischen Zivilgesellschaft/Forschung und Technologiebranche, sodass koordinierte, sektorübergreifende Maßnahmen erforderlich sind.

7.1. Zentrale Erkenntnisse

1. KI ist bereits fest im Leben von Kindern verankert.
KI prägt Bildung, soziale Medien, Suchmaschinen, Unterhaltung und kreative Tools – oft unsichtbar –, sodass junge Menschen routinemäßig mit ihr interagieren, bevor sie sie benennen können.

2. Risiken und Chancen sind eng miteinander verflochten.
Dieselben Gestaltungsmerkmale, die Personalisierung, Kreativität und Barrierefreiheit ermöglichen, bergen zugleich die Gefahr von Fehlinformationen, Voreingenommenheit, Abhängigkeit und kommerzieller Ausbeutung.

3. Die Risikobelastung übersteigt die Schutzmaßnahmen.
Expert:innen bewerten die Exposition von Kindern und den potenziellen Schaden als hoch, während technische, institutionelle und regulatorische Schutzmaßnahmen weiterhin unzureichend sind.

4. KI ist eine gesamtgesellschaftliche Herausforderung.
Wirksame Maßnahmen erfordern koordiniertes Handeln in den Bereichen Technologieentwicklung, Bildung, Familienpraxis, Geschäftsmodelle und Regulierung – und keine isolierten Lösungen.

5. Jugendliche sind aktive, reflektierte Stakeholder.
Die Workshops zeigten, dass junge Menschen konkrete Risiken und praktische Lösungen erkennen können und von Erwachsenen erwarten, dass diese Schutzmaßnahmen und Orientierung bieten.

6. Bildung ist die vielversprechendste, aber fragilste Chance.
Anpassungsfähige Tutoren und Barrierefreiheitstools können Lücken schließen, wenn Lehrervermittlung, Beschaffungsstandards und Evaluierung vorhanden sind; eine unstrukturierte Nutzung birgt die Gefahr, das Lernen zu untergraben und Chancenungleichheiten zu vertiefen.

7. KI-Begleitung und emotionale Abhängigkeit entwickeln sich zu hochriskanten Themen.
Chatbots können Intimität ohne Rechenschaftspflicht simulieren; sie dürfen professionelle Betreuung oder menschliche Beziehungen nicht ersetzen.

8. Eltern und Lehrkräfte spielen eine entscheidende Rolle, verfügen aber über zu wenige Ressourcen.
Da Erwachsene die Nutzung und den Schutz vermitteln, führen Investitionen in Kompetenzaufbau, klare Schulrichtlinien und die Unterstützung von Betreuungspersonen zu unmittelbaren, skalierbaren Schutzmaßnahmen.

9. Das Fachwissen ist fundiert, aber fragmentiert; es gibt weiterhin Evidenzlücken.
Es mangelt an langfristiger, wirkungsorientierter Forschung, einheitlichen Messgrößen und anwendungsorientierten Studien, die Erkenntnisse mit praktikablen Maßnahmen verknüpfen.

10. Eine strukturelle Brücke zum Technologiesektor fehlt.
Die nachhaltige, transparente Zusammenarbeit mit großen Technologieunternehmen ist kaum vorhanden. Es bedarf eines verbindenden, handlungsorientierten Zentrums, um Standards, Audits und gemeinsame Sicherheitsdaten aufeinander abzustimmen.

Insgesamt lässt sich aus diesen Punkten ableiten, dass es Koordinierung und operative Umsetzung braucht: Erkenntnisse bündeln, partizipative Pilotprojekte durchführen, Erwachsene befähigen und auf eine verantwortungsvolle Praxis hinwirken – dabei kurzfristige Maßnahmen mit mittel- und langfristiger Kapazitäts- und Politikarbeit verbinden.

7.2. Empfehlungen

Die prioritären Empfehlungen folgen einer dreistufigen Logik aus kurz-, mittel- und langfristigen Maßnahmen. Kurzfristig sollten evaluative Pilotprojekte in Bildung, Jugendarbeit und humanitären Settings mit Human-in-the-Loop-Designs finanziert und durchgeführt, Mitentwicklungs- und Sicherheits-KPIs für Jugendliche eingebettet, eine KI-Beschaffungscheckliste für Anwendungen mit Kinderkontakt veröffentlicht und einfache Informationsmaterialien für Eltern und Pädagog:innen bereitgestellt werden. Mittelfristig gilt es, Train-the-Trainer-Programme zum Kapazitätsaufbau zu skalieren, effektive Toolkits zu verbreiten, gemeinsame Monitoring- und Evaluationsrahmenwerke mit Gleichberechtigungsindikatoren einzuführen und gezielte Expertenzirkel einschließlich der Tech-Industrie zu initiieren. Langfristig sollte für eine kindgerechte Implementierung von Vorschriften eingetreten und langfristige Forschung zu Entwicklungsfolgen gefördert, sollten validierte Modelle kontextübergreifend skaliert und dauerhafte Standards sowie Monitoring-Mechanismen verankert werden.

Technologieunternehmen müssen Kinderrechte-by-Design in allen Systemen verankern, die wahrscheinlich Minderjährige erreichen. Das bedeutet: routinemäßige kinderrechtliche Folgenabschätzung, altersgerechte Standardmaßnahmen, strenge Datenminimierung, zuverlässige Alterskontrolle und schnelle Minderung bei vorhersehbarem Missbrauch (Deepfakes, Grooming, bösartige Bots). Unternehmen sollten über bloße Compliance hinausgehen: Sicherheitsdaten mit Forschenden und Regulierungsbehörden teilen, unabhängige Audits finanzieren und strukturierte Multi-Stakeholder-Verantwortlichkeitsprozesse einschließen. Letztlich muss die Industrie die Verantwortung für designbedingte nachgelagerte Schäden übernehmen.

Regierungen sollten bestehende Regeln (Plattformsicherheit, Datenschutz, Jugendmediengesetz) durchsetzen und Kinderrechtsauswirkungen für KI, die Minderjährige betrifft, vorschreiben. Sie müssen in Basiskompetenzen investieren – KI-Kompetenz, Lehrkräfteausbildung, Elternführung – und zugängliche Materialien finanzieren. Stärkere grenzüberschreitende Zusammenarbeit, besser ausgestattete Aufsichtsbehörden, institutionalisierte Jugendbeteiligung an der

Politik und dauerhafte Koordinationsmechanismen sind ebenfalls erforderlich.

Wissenschaft und Forschung sollten reale, langfristige und evaluative Studien priorisieren, die die Praxis informieren: standardisierte Benchmarks, kindspezifische Validierung und Studien zu kumulativen Effekten. Die Ergebnisse müssen in brauchbare Leitlinien (Briefs, Tools, Lehrerressourcen) übersetzt werden. Forscher:innen sollten hohe ethische Standards für die Teilhabe von Kindern einhalten und disziplinübergreifend sowie mit der Zivilgesellschaft und Regulierungsbehörden zusammenarbeiten, damit die Evidenz bei Entscheidungsträger:innen ankommt.

NGOs und die Zivilgesellschaft sollten praktische, niedrigschwellige Unterstützungen – Eltern- und Lehrer:innen-Toolkits, Hotlines, schnelle Beratung und kulturell sensible KI-Kompetenz – ausweiten. Sie müssen reale Schäden systematisch dokumentieren und Fälle in politische und Durchsetzungskanäle einleiten. NGOs sollten sektorübergreifende Foren einberufen, die kontinuierliche Jugendbeteiligung stärken und Prinzipien in operative Werkzeuge (Beschaffungschecklisten, einfache Leitfäden) für den sofortigen Einsatz umsetzen.

Den Umgang mit KI sicher zu gestalten, ist eine gemeinsame Aufgabe: Die Tech-Branche muss auf Sicherheit setzen, Regierungen müssen Vorschriften durchsetzen und die entsprechenden Kapazitäten finanzieren, Forschende müssen verwertbare Erkenntnisse vorlegen, Nichtregierungsorganisationen müssen eine Brücke zwischen Praxis und Politik schlagen und dabei die Jugend in den Mittelpunkt stellen. Eine nachhaltige Koordinierung – gestützt auf Transparenz, unabhängige Überprüfungen und eine sinnvolle Jugendbeteiligung – ist der schnellste Weg von der Problemdiagnose hin zu messbarem Schutz und Nutzen für Kinder und Jugendliche.

8. Literaturverzeichnis

Adobe. (2024). *Adobe Express for Education Brings AI Designed to Be Safe for the Classroom to Indian Teachers and Students to Enhance Creative and Digital Literacy*. https://news.adobe.com/en/apac/news/pr/in/press-release-adobe-express-for-education-brings-ai-designed-to-be-safe-for-the-classroom-to-indian-teachers-and-students-to-enhance-creative-and-digital-literacy

AI Act Database. (2024). *EU Artificial Intelligence Act: Article 5 (Prohibited AI practices)*. https://artificialintelligenceact.eu/article/5/

AI4K12 Initiative. (2019). *K-12 AI Guidelines: The Five Big Ideas*. https://ai4k12.org/

Arizmendi, C. J. et al. (2022). *Predicting student outcomes using digital logs of learning behaviors: Review, current standards, and suggestions for future work*. https://pmc.ncbi.nlm.nih.gov/articles/PMC10556130/

Arnon, S. (2022). Association of Cyberbullying Experiences and Perpetration With Suicidality in Early Adolescence. *JAMA Network Open*. https://jamanetwork.com/journals/jamanetworkopen/fullarticle/2793627

Associated Press. (2025, März 12). *Schools use AI to monitor kids, hoping to prevent violence. Our investigation found security risks*. https://apnews.com/article/ai-school-chromebook-gaggle-goguardian-securly-25a3946727397951fd42324139aaf70f

Ayer, L., Boudreaux, B., Welburn Paige, J., Holmes, P., Blagg, T. L., & Mendon-Plasek, S. J. (2023). *Artificial Intelligence–Based Student Activity Monitoring for Suicide Risk: Considerations for K–12 Schools, Caregivers, Government, and Technology Developers* (RR-A2910-1). RAND Corporation. https://doi.org/10.7249/RRA2910-1

Bahmani, K. et al. (2022). *Face Recognition In Children: A Longitudinal Study (arXiv preprint)*. https://arxiv.org/pdf/2204.01760

Bark Technologies. (2023). What Being Online Was Like for Kids in 2023. *Bark´s Annual Report 2023*. https://bark.us/annual-report-2023

Better Internet for Kids (BIK). (2025, September 4). *In the spotlight: Safer Internet Centre Germany*. https://better-internet-for-kids.europa.eu/en/news/spotlight-safer-internet-centre-germany

BfDI. (2025, September 4). *BfDI welcomes EDPB guidelines on legitimate interest*. https://www.bfdi.bund.de/SharedDocs/Pressemitteilungen/EN/2024/12_EDSA-Leitlinien-berechtigtes-Interesse.html

BMBFSFJ. (2025a, September 4). *Initiative „SCHAU HIN! Was Dein Kind mit Medien macht“*. https://www.bmbfsfj.bund.de/bmbfsfj/themen/kinder-und-jugend/medienkompetenz/initiative-schau-hin-was-dein-kind-mit-medien-macht--73512

BMBFSFJ. (2025b, September 4). *Künstliche Intelligenz verschärft Risiken für Kinder und Jugendliche im Netz.* Jahresbericht von jugendschutz.net. https://www.bmbfsfj.bund.de/bmbfsfj/aktuelles/alle-meldungen/kuenstliche-intelligenz-verschaerft-risiken-fuer-kinder-und-jugendliche-im-netz-243642

Bradley, H. et al (2025). Voice assistant technology continues to underperform on children's speech. *JASA Express Letters*. https://pubs.aip.org/asa/jel/article/5/3/035201/3338215/Voice-assistant-technology-continues-to-underperform-on-children-s-speech

Bundesbeauftragte für den Datenschutz und die Informationsfreiheit (BfDI). (2025, September 4). *Basiswissen für Kinder und Jugendliche*. https://www.bfdi.bund.de/DE/Buerger/Basiswissen/KinderundJugendliche/KinderundJugendliche-node.html

Bundesnetzagentur. (2025a, September 4). *Bundesnetzagentur becomes central platform supervisory authority for Germany*. https://www.bundesnetzagentur.de/SharedDocs/Pressemitteilungen/EN/2024/20240514_DSC.html

Bundesnetzagentur. (2025b, September 4). *Digital Services Coordinator*. https://www.dsc.bund.de/

Bundesnetzagentur. (2025c, September 4). *Digital Services Coordinator publishes 2024 activity report.* https://www.bundesnetzagentur.de/SharedDocs/Pressemitteilungen/EN/2025/20250812_DSC_Taetigkeitsbericht.html

Bundeszentrale für Kinder- und Jugendmedienschutz (BzKJ). (2025). *Das Jugendschutzgesetz*. juschg-stand-09-04-2021-data.pdf

Cambridge Dictionary. (2025). *Opportunity Definition*. https://dictionary.cambridge.org/dictionary/english/opportunity

Center for Democracy & Technology. (2024). *Student Demands for Better Guidance Outpace School Supports to Spot Deepfakes*. https://cdt.org/insights/just-released-research-student-demands-for-better-guidance-outpace-school-supports-to-spot-deepfakes/

Center for Democracy and Technology. (2020). *How Automated Test Proctoring Software Discriminates Against Disabled Students*. https://cdt.org/insights/how-automated-test-proctoring-software-discriminates-against-disabled-students/

Children's Commissioner for England. (2025). *„Sex is kind of broken now": children and pornography*. Office of the Children's Commissioner for England. https://www.childrenscommissioner.gov.uk/resource/sex-is-kind-of-broken-now-children-and-pornography/

Coalition for Content Provenance and Authenticity (C2PA). (2024). *Content Credentials: C2PA Technical Specification, Version 2.1*. https://spec.c2pa.org/specifications/specifications/2.1/specs/_attachments/C2PA_Specification.pdf

Coghlan, S. et al. (2021). Good Proctor or "Big Brother"? Ethics of Online Exam Supervision Technologies. *Philosophy & Technology, 34*(4), 1581–1606. https://doi.org/10.1007/s13347-021-00476-1

Council of Europe. (2018). *Guidelines to respect, protect and fulfil the rights of the child in the digital environment*. https://rm.coe.int/guidelines-to-respect-protect-and-fulfil-the-rights-of-the-child-in-th/16808d881a

Deutsche Telekom AG. (2025, September 4). *#ShareWithCare deepfake awareness campaign*. https://www.telekom.com/en/company/details/share-with-care-telekom-raises-awareness-1041810

Diel, A., Lalgi, T., Schröter, I. C., MacDorman, K. F., Teufel, M. & Bäuerle, A. (2024). Human performance in detecting deepfakes: A systematic review and meta-analysis of 56 papers. *Computers in Human Behavior Reports*, *16*, 100538. https://doi.org/10.1016/j.chbr.2024.100538

eco – Association of the Internet Industry. (2025, September 4). *Eco Complaints Office network page*. https://international.eco.de/topics/policy-law/eco-complaints-office/network/

eco – Verband der Internetwirtschaft e. V. & FSM. (2026). *Internet-Beschwerdestelle*. https://www.internet-beschwerdestelle.de/en/

Elternguide.online consortium. (2025, September 4). *Künstliche Intelligenz – Elternguide.online*. https://www.elternguide.online/kuenstliche-intelligenz/

EUR-LEX. (2024). *EU AI Act*. https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng

European Commission. (2022a). *A Digital Decade for children and youth: the new European strategy for a better internet for kids (BIK+)*. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52022DC0212

European Commission. (2022b). *The EU's Digital Services Act*. https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/digital-services-act_en

European Commission. (2025a). *Commission publishes guidelines on the protection of minors*. European Commission. https://digital-strategy.ec.europa.eu/en/library/commission-publishes-guidelines-protection-minors

European Commission. (2025b). *Guidelines on the protection of minors under the Digital Services Act*. https://digital-strategy.ec.europa.eu/en/library/commission-publishes-guidelines-protection-minors

European Parliament. (2017, März 22). *Spy device in 'My Friend Cayla' doll*. https://www.europarl.europa.eu/doceo/document/E-8-2017-001901_EN.html

European Parliament. (2025). *Children and deepfakes*. https://www.europarl.europa.eu/RegData/etudes/BRIE/2025/775855/EPRS_BRI%282025%29775855_EN.pdf

Falcione, K. (2025). Psychopathology and Gaming Disorder in Adolescents. *JAMA Network Open*. https://jamanetwork.com/journals/jamanetworkopen/fullarticle/2836903

Federal Bureau of Investigation. (2024a). *The Financially Motivated Sextortion Threat*. https://www.fbi.gov/news/stories/the-financially-motivated-sextortion-threat

Federal Bureau of Investigation, I. (2024b). *2024 Internet Crime Report*. https://www.ic3.gov/AnnualReport/Reports/2024_IC3Report.pdf

Federal Trade Commission. (2025). *Children's Online Privacy Protection Rule* (Final Rule, 90 FR 16916). https://www.federalregister.gov/documents/2025/04/22/2025-05904/childrens-online-privacy-protection-rule

Fox, C. B. et al. (2021). Automatic Speech Recognition and Real-Time Transcription. *Journal of Speech, Language, and Hearing Research*. https://doi.org/10.1044/2021_JSLHR-21-00096

Fraunhofer SIT. (2025, September 4). *Digitaler Jugendschutz mit künstlicher Intelligenz – Machbarkeitsstudie*. https://www.sit.fraunhofer.de/de/presse/details/news-article/show/digitaler-jugendschutz-mit-kuenstlicher-intelligenz/

FSM. (2025, September 4). *FSM-Beschwerdestelle*. https://www.fsm.de/fsm/beschwerdestelle/

Diliberti, M. et al. (2024). *Using Artificial Intelligence Tools in K–12 Classrooms* (RR-A956-21). RAND Corporation. https://www.rand.org/pubs/research_reports/RRA956-21.html

Gentleman, A., Horton, H. & Milmo, D. (2026, 7. Januar). Grok AI still being used to digitally undress women and children despite suspension pledge. *The Guardian*. https://www.theguardian.com/technology/2026/jan/05/elon-musk-grok-ai-digitally-undress-images-of-women-children

Gorwa, R., Binns, R. & Katzenbach, C. (2020). Algorithmic content moderation: Technical and political challenges in the automation of platform governance. *Big Data & Society*. https://doi.org/10.1177/2053951719897945

Grother, P. (2022). *Face Recognition Vendor Test (FRVT) Part 8: Summarizing Demographic Differentials* (NIST IR 8429 ipd). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.IR.8429.ipd

Hanaoka, K., Ngan, M., Yang, J., Quinn, G. W., Hom, A., & Grother, P. (2024). *Face Analysis Technology Evaluation: Age Estimation and Verification* (NIST IR 8525). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.IR.8525

Han, S. et al. (2024). *Digital Proctoring in Higher Education: A Systematic literature review. International Journal of Educational Management, 38*(1), 265–285. https://doi.org/10.1108/IJEM-12-2022-0522

Have I Been Pwned. (2017). *CloudPets Data Breach*. https://haveibeenpwned.com/Breach/CloudPets

Hern, A. (2017, 28. Februar). *CloudPets stuffed toys leak details of half a million users*. https://www.theguardian.com/technology/2017/feb/28/cloudpets-data-breach-leaks-details-of-500000-children-and-adults

Ho, S. & Burke, G. (2022, April 29). *An algorithm that screens for child neglect raises concerns*. https://apnews.com/article/child-welfare-algorithm-investigation-9497ee937e0053ad4144a86c68241ef1

Hua, S. B. Z. et al. (2025). Lack of children in public medical imaging data points to growing age bias in biomedical AI. *Current Problems in Diagnostic Radiology (preprint/early view via PubMed)*. https://pubmed.ncbi.nlm.nih.gov/40661269/

IDnow GmbH. (2025, September 4). *Vollautomatisierte Identitätsprüfung von der KJM für die Altersverifikation zugelassen*. https://www.idnow.io/de/pr-de/idcheckio-kjm-altersverifikation-zugelassen/

IMY (Swedish Authority for Privacy Protection). (2019). *Facial recognition in school renders Sweden's first GDPR fine* https://www.imy.se/en/about-us/arkiv/nyhetsarkiv/facial-recognition-in-school-renders-swedens-first-gdpr-fine/

Information Commissioner's Office. (2021). *Age Appropriate Design Code (Children's Code)*. https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/childrens-information/childrens-code-guidance-and-resources/age-appropriate-design-a-code-of-practice-for-online-services/

Internet Watch Foundation. (2024). *AI-generated child sexual abuse (IWF Annual Report 2023)*. https://www.iwf.org.uk/annual-report-2023/trends-and-data/ai-generated-child-sexual-abuse/

ITU. (2023). *Global offline population steadily declines to 2.6 billion people in 2023*. https://www.itu.int/itu-d/reports/statistics/2023/10/10/ff23-internet-use

Jugendschutz.net. (2025). *Bericht 2024. Jugendschutz im Internet. Risiken und Handlungsbedarf*. jahresbericht_2024.pdf

JUUUPORT e.V. (2025, September 4). *Über uns – JUUUPORT*. https://www.juuuport.de/ueber-uns

K12 Security Information eXchange (K12 SIX). (2023). *The K-12 Cyber Incident Map* [Dataset]. https://www.k12six.org/map

Kelly, R. (2024, August 28). *Survey: 86% of Students Already Use AI in Their Studies*. Campus Technology. https://campustechnology.com/articles/2024/08/28/survey-86-of-students-already-use-ai-in-their-studies.aspx

Khan Academy. (2023). *Free AI Quiz and Question Generator*. https://www.khanmigo.ai/question-generator

Khoo, Y. X. et al. (2024). "That's Kind of Sus(picious)": The Comprehensiveness of Mental Health Application Users' Privacy and Security Concerns. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (CHI '24). ACM. https://dl.acm.org/doi/10.1145/3613904.3642705

Kim, D. et al. (2025). Usefulness of Automatic Speech Recognition Assessment of Children With Speech Sound Disorders: Validation Study. *Journal of Medical Internet Research*. https://www.jmir.org/2025/1/e60520/

klicksafe / Medienanstalt Rheinland-Pfalz. (2025, September 4). *klicksafe: Die EU-Initiative für mehr Sicherheit im Netz*. https://www.klicksafe.de/

Kommission für Jugendmedienschutz (KJM). (2025, September 4). *Verfahren der Positivbewertung für AV-Systeme*. https://www.kjm-online.de/themen/technischer-jugendmedienschutz/unzulaessige-inhalte/

Kosoy, E., Jeong, S., Sinha, A., Kraljic, T. & Gopnik, A. (2024). *Children's Mental Models of Generative Visual and Text Based AI Models*. https://doi.org/10.48550/arXiv.2405.13081

Kouroupa, A., Laws, K. R., Irvine, K., Mengoni, S. E., Baird, A. & Sharma, S. (2022). The use of social robots with children and young people on the autism spectrum: A systematic review and meta-analysis. *PLOS ONE*, *17*(6), e0269800. https://doi.org/10.1371/journal.pone.0269800

Kulik, J. A. & Fletcher, J. D. (2016). Effectiveness of Intelligent Tutoring Systems: A Meta-Analytic Review. *Review of Educational Research*, *86*(1), 42–78. https://doi.org/10.3102/0034654315581420

Kultusministerkonferenz (KMK). (2025). *Handlungsempfehlung für die Bildungsverwaltung zum Umgang mit Künstlicher Intelligenz in schulischen Bildungsprozessen*. https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2024/2024_10_10-Handlungsempfehlung-KI.pdf

Leibniz-Institut für Medienforschung, Hans-Bredow-Institut (HBI). (2025a, September 4). *EU Kids Online*. https://leibniz-hbi.de/en/hbi-projects/eu-kids-online/

Leibniz-Institut für Medienforschung, Hans-Bredow-Institut (HBI). (2025b, September 4). *Sicherheit für Kinder in der digitalen Welt (SIKID) – HBI project*. https://leibniz-hbi.de/hbi-projects/sicherheit-fuer-kinder-in-der-digitalen-welt/

López-Zambrano, J. et al. (2021). *Early Prediction of Student Learning Performance Through Data Mining: A Systematic Review. Psicothema.* https://www.psicothema.com/pdf/4692.pdf

Makin, L. et al. (2025). Smart speakers are an acceptable and feasible speech practice tool for children with speech difficulties. *Disability and Rehabilitation: Assistive Technology*. https://pubmed.ncbi.nlm.nih.gov/40244154/

McKenzie, R. (2025). 493 Cases of Sextortion Against Children Linked to Notorious Scam Compounds. *WIRED*. https://www.wired.com/story/child-sextorition-scam-compounds-southeast-asia/

Medienpädagogischer Forschungsverbund Südwest. (2024). JIMStudie 2024. https://mpfs.de/app/uploads/2024/11/JIM_2024_PDF_barrierearm.pdf

Meta Transparency Center. (2025). *Community Standards Enforcement Report: Child Nudity and Sexual Exploitation*. https://transparency.meta.com/reports/community-standards-enforcement/child-nudity-and-sexual-exploitation/facebook/

Miao, F., Holmes, W. & UNESCO. (2023). *Guidance for generative AI in education and research*. https://unesdoc.unesco.org/ark:/48223/pf0000386693

Microsoft Research Blog. (2025, Juli 17). *Lack of children in public medical imaging data points to growing age bias in biomedical AI*. https://www.microsoft.com/en-us/research/articles/lack-of-children-in-public-medical-imaging-data-points-to-growing-age-bias-in-biomedical-ai/

Milmo, D. (2025). AI images of child sexual abuse getting 'significantly more realistic', says watchdog. *The Guardian*. https://www.theguardian.com/technology/2025/apr/23/ai-images-of-child-sexual-abuse-getting-significantly-more-realistic-says-watchdog

MIT RAISE. (2021). *Day of AI: Curriculum and program*. https://dayofai.org/

Moore, J., Grabb, D., Agnew, W., Klyman, K., Chancellor, S., Ong, D. C. & Haber, N. (2025). *Expressing stigma and inappropriate responses prevents LLMs from safely replacing mental health providers*. Proceedings of the 2025 ACM Conference on Fairness, Accountability, and Transparency (FAccT '25). https://doi.org/10.1145/3715275.3732039

Mozilla Foundation. (2023). *Are mental health apps better or worse at privacy in 2023?* https://www.mozillafoundation.org/en/privacynotincluded/articles/are-mental-health-apps-better-or-worse-at-privacy-in-2023/

Muralidharan, K., Singh, A. & Ganimian, A. J. (2019). *Disrupting Education? Experimental Evidence on Technology-Aided Instruction in India. American Economic Review*, *109*(4), 1426–1460. https://doi.org/10.1257/aer.20171112

Nagata, J. M. et al. (2024). *Screen time and mental health: A prospective analysis of the Adolescent Brain Cognitive Development(ABCD) Study. BMC Public Health*. https://doi.org/10.1186/s12889-024-20102-x

National Center for Missing & Exploited Children. (2024). *NCMEC Releases New Sextortion Data*. https://www.missingkids.org/blog/2024/ncmec-releases-new-sextortion-data

New York State Education Department. (2023). *NYSED Issues Determination on Biometric Identifying Technology in Schools*. https://www.nysed.gov/news/2023/state-education-department-issues-determination-biometric-identifying-technology-schools

Newman, M., Sun, K., Dalla Gasperina, I. B., Shin, G. Y., Pedraja, M. K., Kanchi, R., Song, M. B., Li, R., Lee, J. H. & Yip, J. (2024). *"I want it to talk like Darth Vader": Helping Children Construct Creative Self-Efficacy with Generative AI*. Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems (CHI '24). https://doi.org/10.1145/3613904.3642492

OECD. (2025). *How's Life for Children in the Digital Age?* https://doi.org/10.1787/0854b900-en

Ofcom. (2022). *Smart Speakers Research Report*. https://www.ofcom.org.uk/siteassets/resources/documents/research-and-data/multi-sector/general-communications-research/smart-speakers-research.pdf

Ofcom. (2024). *Children and parents: media use and attitudes report 2024 – interactive data* [Dataset]. https://www.ofcom.org.uk/media-use-and-attitudes/media-habits-children/children-and-parents-media-use-and-attitudes-report-2024-interactive-data

Ofcom. (2025). *Children and Parents: Media Use and Attitudes Report*. https://www.ofcom.org.uk/siteassets/resources/documents/research-and-data/media-literacy-research/children/childrens-media-use-and-attitudes-report-2025/childrens-media-literacy-report-2025.pdf?v=396621

Office of the U.S. Surgeon General. (2023). *Social Media and Youth Mental Health: The U.S. Surgeon General's Advisory*. U.S. Department of Health and Human Services. https://www.hhs.gov/sites/default/files/sg-youth-mental-health-social-media-advisory.pdf

Open Knowledge Foundation Deutschland e.V.; mediale pfade.org. (2025). *About Jugend hackt*. https://jugendhackt.org/about/

Organisation for Economic Co-operation and Development. (2025). *The legal and policy landscape of age assurance online for child safety and well-being*. OECD Publishing. https://doi.org/10.1787/4a1878aa-en

Radesky, J. et al. (2022). Prevalence and Characteristics of Manipulative Design in Mobile Applications Used by Children. *JAMA Network Open, 5*(6), e2217641. https://doi.org/10.1001/jamanetworkopen.2022.17641

Radesky, J. et al. (Authors as listed by JAMA Network Open). (2024). Algorithmic Content Recommendations on a Video-Sharing Platform Used by Children. https://doi.org/10.1001/jamanetworkopen.2024.13855

Raffoul, A., Ward, Z. J., Santoso, M., Kavanaugh, J. R. & Austin, S. B. (2023). Social media platforms generate billions of dollars in revenue from U.S. youth: Findings from a simulated revenue model. *PLOS ONE*, *18*(12), e0295337. https://doi.org/10.1371/journal.pone.0295337

Raymond, N. (2025). Massachusetts college student to plead guilty to PowerSchool data breach. *Reuters*. https://www.reuters.com/legal/massachusetts-hacker-plead-guilty-powerschool-data-breach-2025-05-20/

Regulation (EU) 2016/679 General Data Protection Regulation (GDPR) — Recital 38 and Article 8 (2016). https://eur-lex.europa.eu/eli/reg/2016/679/oj

Risk Definition. (o. J.). Cambridge. https://dictionary.cambridge.org/dictionary/english/risk

Santos, R. M. et al. (2023). Accurate, timely, and portable: Course-agnostic early prediction of student performance from LMS logs. *Computers and Education: Artificial Intelligence, 5*, 100175. https://doi.org/10.1016/j.caeai.2023.100175

Scassellati, B., Boccanfuso, L., Huang, C.-M., Mademtzi, M., Qin, M., Salomons, N., Ventola, P. & Shic, F. (2018). Improving social skills in children with ASD using a long-term, in-home social robot. *Science Robotics*, *3*(21), eaat7544. https://doi.org/10.1126/scirobotics.aat7544

Sidoti, O., Park, E. & Gottfried, J. (2025,15. Januar). *About a quarter of U.S. teens have used ChatGPT for schoolwork – double the share in 2023*. Pew Research Center. https://www.pewresearch.org/short-reads/2025/01/15/about-a-quarter-of-us-teens-have-used-chatgpt-for-schoolwork-double-the-share-in-2023/

Smahel, D., MacHackova, H., Mascheroni, G., Dedkova, L., Staksrud, E., Ólafsson, K., Livingstone, S. & Hasebrink, U. (2025). *EU Kids Online 2020: survey results from 19 countries*. https://doi.org/10.21953/lse.47fdeqj01ofo

Southwell, R. et al. (2024). Automatic Speech Recognition Tuned for Child Speech in the Classroom *(ICASSP 2024)*. https://ieeexplore.ieee.org/document/10447428

Srinivas, N., Ricanek, K., Michalski, D., Bolme, D. S. & King, M. (2019). *Face Recognition Algorithm Bias: Performance Differences on Images of Children and Adults*. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops

(CVPRW). https://openaccess.thecvf.com/content_CVPRW_2019/papers/BEFA/Srinivas_Face_Recognition_Algorithm_Bias_Performance_Differences_on_Images_of_Children_CVPRW_2019_paper.pdf

Stoilova, M., Livingstone, S. & Bulger, M. (2023). Do parental control tools fulfil family expectations for child protection? A rapid evidence review of the contexts and outcomes of use. *Journal of Children and Media.* https://doi.org/10.1080/17482798.2023.2265512

Terre des Hommes IF (2024). *Children's Rights at the Centre of EU Digital Policy in 2024-2029*. https://terredeshommes.org/blog/2024/04/18/childrens-rights-at-the-centre-of-eu-digital-policy-in-2024-2029/#:~:text=,existing%20policies%20and%20legislation%2C%20notably

Thorn. (2025). *New data reveals the devastating toll of sextortion on kids*. https://www.thorn.org/press-releases/new-data-reveals-the-devastating-toll-of-sextortion-on-kids/

Turnitin. (2019). *Gradescope revolutionizes grading experience for university physics courses*. https://www.turnitin.com/customer-stories/gradescope-revolutionizes-grading-experience-for-university-physics-courses

UN Committee on the Rights of the Child. (2021). *General comment No. 25 (2021) on children's rights in relation to the digital environment*. United Nations. https://www.ohchr.org/en/documents/general-comments-and-recommendations/general-comment-no-25-2021-childrens-rights-relation

UNICEF. (2019). *UNICEF poll: More than a third of young people in 30 countries report being a victim of online bullying*. https://www.unicef.org/press-releases/unicef-poll-more-third-young-people-30-countries-report-being-victim-online-bullying

UNICEF. (2021). *Policy guidance on AI for children*. UNICEF-Global-Insight-policy-guidance-AI-children-2.0-2021.pdf

UNICEF Innocenti. (2024). *Generative AI: Risks and opportunities for children*. https://www.unicef.org/innocenti/generative-ai-risks-and-opportunities-children

UNICEF Office of Research – Innocenti. (2025). *Childhood in a Digital World (Flagship report and statistical analysis)*. https://www.unicef.org/innocenti/innocenti/media/11296/file/UNICEF-Innocenti-Childhood-in-a-Digital%20World-report-2025.pdf

United Nations. (2021). *General comment No. 25 (2021) on children's rights in relation to the digital environment*. https://www.ohchr.org/en/documents/general-comments-and-recommendations/general-comment-no-25-2021-childrens-rights-relation

University of Edinburgh. (2024, März 21). *Children may overestimate smart speakers' abilities*. https://www.ed.ac.uk/research-innovation/latest-research-news/children-may-overestimate-smart-speakers-abilities

U.S. Department of Justice. (2023, Juli 19). *Amazon Agrees to Injunctive Relief and $25 Million Civil Penalty for Alleged Violations of Children's Privacy Law Relating to Alexa*. https://www.justice.gov/archives/opa/pr/amazon-agrees-injunctive-relief-and-25-million-civil-penalty-alleged-violations-childrens

Vasconcellos, R. P. et al. (2025). *Electronic Screen Use and Children's Socioemotional Problems: A Systematic Review and Meta-Analysis of Longitudinal Studies. Psychological Bulletin*. https://www.apa.org/pubs/journals/releases/bul-bul0000468.pdf

Wang, S. et al. (2024). *Artificial intelligence in education: A systematic literature review. Expert Systems with Applications, 252*, 124167. https://doi.org/10.1016/j.eswa.2024.124167

Weale, S. (2021, Oktober 18). *ICO to step in after schools use facial recognition to speed up lunch queue*. https://www.theguardian.com/education/2021/oct/18/privacy-fears-as-schools-use-facial-recognition-to-speed-up-lunch-queue-ayrshire-technology-payments-uk

WHO. (2024). *One in six school-aged children experiences cyberbullying, finds new WHO/Europe study*. https://www.who.int/europe/news/item/27-03-2024-one-in-six-school-aged-children-experiences-cyberbullying--finds-new-who-europe-study

Wood, S. G., Moxley, J. H., Tighe, E. L. & Wagner, R. K. (2018). *Does Use of Text-to-Speech and Related Read-Aloud Tools Improve Reading Comprehension for Students With Reading Disabilities? A Meta-Analysis. Journal of Learning Disabilities*, *51*(1), 73–84. https://doi.org/10.1177/0022219416688170

WSJ Staff. (2025). *TikTok Might Be Breaching Digital Services Act Rules on Advertising, EU Says*. https://www.wsj.com/tech/eu-says-tiktok-might-break-digital-services-act-ad-rules-18e3a2df

Yang, M. (2021, Dezember 30). *Amazon's Alexa device tells 10-year-old to touch a penny to a live plug socket*. https://www.theguardian.com/technology/2021/dec/29/amazons-alexa-child-penny-live-plug

Ye, Z. et al. (2023). Meta-analysis of the relationship between bullying and depressive symptoms in children and adolescents. *BMC Psychiatry, 23*, 215. https://doi.org/10.1186/s12888-023-04681-4

Zhou, R. et al. (2024). Meta-Analysis of Internet Gaming Disorder Prevalence: Assessing the Impacts of DSM5 and ICD11 Diagnostic Criteria. *International Journal of Mental Health and Addiction*. https://pmc.ncbi.nlm.nih.gov/articles/PMC11203952/

Zhu, X. et al. (2025). *Digital Detection Meets Crisis Intervention: A National County-level study of GoGuardian Beacon implementation and sustainment on Youth Suicide Rates* [Preprint]. medRxiv. https://doi.org/10.1101/2025.04.22.25326214

9. Anhang



Clemens Wasner ist einer der renommiertesten KI-Experten Österreichs und eine führende Stimme in der Debatte um die gesellschaftlichen Auswirkungen Künstlicher Intelligenz. Er ist Gründer und CEO von EnliteAI, einem KI-Unternehmen, das er 2025 erfolgreich an den globalen Geospatial-Technologies-Anbieter Cyclomedia (größter Geodatenanbieter nach Google) verkauft hat. Wasner ist außerdem Mitbegründer und Vorstandsmitglied von AI Austria, dem größten KI-Netzwerk des Landes, in dem Bildung und KI-Kompetenz wiederkehrende Schwerpunkte sind. Er berät regelmäßig politische Entscheidungsträger:innen, darunter das Bildungsministerium, und moderiert den vielbeachteten Podcast „AI Talk“. Mit über 18.000 LinkedIn-Followern spielt er eine zentrale Rolle bei der Gestaltung des KI-Diskurses in Österreich.

Institution: AI Austria
Stakeholder-Klassifizierung: Tech-Industrie/Zivilgesellschaft
Beschreibung der Institution: AI Austria ist das größte KI-Netzwerk in Österreich und bringt Industrie, Forschung und zivilgesellschaftliche Akteure zusammen. Es fördert die verantwortungsvolle, gemeinnützige KI-Entwicklung und den öffentlichen Dialog über die gesellschaftliche Wirkung von KI.

Frank Woithe ist leitender Experte bei jugendschutz.net, dem gemeinsamen Kompetenzzentrum der deutschen Bundesländer zum Online-Schutz von Jugendlichen. Seine Arbeit konzentriert sich auf die technischen und funktionalen Aspekte digitaler Dienste, einschließlich Altersdifferenzierung, Sicherheitseinstellungen und deren realer Wirksamkeit. Er und sein Team testen Plattformen systematisch aus der Perspektive von fiktiven Minderjährigen, um tatsächliche Expositionsrisiken zu bewerten. Ursprünglich als Lehrer ausgebildet, arbeitet Woithe seit über 15 Jahren im Jugendschutz und konzentriert sich nun stark auf die Auswirkungen von KI auf die Mediennutzung von Minderjährigen.

Institution: jugendschutz.net
Stakeholder-Klassifizierung: staatliche Institution
Beschreibung der Institution: jugendschutz.net fungiert als das gemeinsame Kompetenzzentrum von Bund, Ländern und Landesmedienanstalten für den Schutz von Kindern und Jugendlichen im Internet. Die Stelle recherchiert Gefahren und Risiken in jugendaffinen Diensten und drängt Anbieter und Betreiber, ihre Angebote so zu gestalten, dass Kinder und Jugendliche sie unbeschwert nutzen können. Sie nimmt über ihre Hotline Hinweise auf Verstöße gegen den Jugendmedienschutz entgegen und sorgt dafür, dass diese schnell beseitigt werden.

Martin Fischer ist Referent für Medien- und Digitalpolitik beim Deutschen Kinderhilfswerk mit Schwerpunkt auf Kinderrechten im digitalen Raum. Er arbeitet in der „Koordinierungsstelle Kinderrechte in der digitalen Welt“, wo KI zu einem zentralen Thema geworden ist. Seine Rolle verbindet Interessenvertretung, politische Koordination und Dialog zwischen Zivilgesellschaft, Regulierungsbehörden und internationalen Institutionen. Fischer positioniert sich bewusst zwischen der Hervorhebung von Risiken und der Anerkennung von Chancen der KI für Kinder und Jugendliche.



Institution: Deutsches Kinderhilfswerk e.V.
Stakeholder-Klassifizierung: NGO
Beschreibung der Institution: Das Deutsche Kinderhilfswerk ist eine gemeinnützige Organisation, die sich seit über 50 Jahren für Kinderrechte in Deutschland einsetzt. Sie verbindet politische Interessenvertretung, Projektfinanzierung und direkte Unterstützung, um Chancengleichheit für alle Kinder zu fördern.

Rebecca Wald ist eine deutsche Forscherin mit Sitz in Amsterdam, die sich auf die alltäglichen Interaktionen von Kindern und Jugendlichen mit digitalen Technologien spezialisiert hat. Sie promovierte an der Universität Amsterdam über intelligente Lautsprecher in Fa-



milien mit kleinen Kindern und zeigte damit frühe und weit verbreitete Adaptionsmuster auf. Ihre Forschung zeigte, dass nicht Datenschutzbedenken, sondern die digitalen Fähigkeiten der Eltern die Technologienutzung von Kindern und Jugendlichen am stärksten beeinflussen. Heute arbeitet sie als Fachexpertin zum Thema Online-Kindheit am Niederländischen Jugendinstitut und berät sowohl Praxis als auch Politik.

Institution: Nederlands Jeugdinstituut
Stakeholder-Klassifizierung: Forschung
Beschreibung der Institution: Das Nederlands Jeugdinstituut ist das unabhängige nationale Wissenszentrum der Niederlande für Kinder und Erziehung. Es unterstützt Fachkräfte, politische Entscheidungsträger:innen, Eltern sowie Kinder und Jugendliche mit wissenschaftlich fundierten Erkenntnissen, um die Chancengleichheit für Kinder und Jugendliche zu fördern.



Larissa Leonhard ist Postdoktorandin am Institut für Kommunikations- und Medienforschung der LMU München. Der Schwerpunkt ihrer Arbeit liegt auf digitaler Kompetenz in algorithmisch geprägten Umgebungen, wobei ihr besonderes Augenmerk Kindern und Jugendlichen gilt. Im Jahr 2023 führte sie eine groß angelegte Studie mit gemischten Methoden unter 14- bis 17-Jährigen zum Thema KI und Algorithmen im Alltag durch.

Institution: Ludwig-Maximilians-Universität München
Stakeholder-Klassifizierung: Forschung
Beschreibung der Institution: Die LMU München ist eine der führenden Forschungsuniversitäten Europas mit einer starken Tradition in den Sozialwissenschaften und der Medienforschung. Ihre Kommunikationsforschung trägt wesentlich zu internationalen Debatten über digitale Medien und Gesellschaft bei.



Giovanna Mascheroni ist Soziologin für digitale Medien und Professorin am Institut für Kommunikationswissenschaften der Università Cattolica del Sacro Cuore. Ihre Arbeit konzentriert sich auf die sozialen Auswirkungen digitaler Medien, der Datenidentifizierung und der Künstlichen Intelligenz auf Kinder, Jugendliche und Familien. Sie trat 2007 als italienische Teamleiterin dem „EU Kids Online"-Netzwerk bei und ist seit 2023 dessen stellvertretende Koordinatorin. Sie spielte eine führende Rolle im von Horizon 2020 finanzierten ySKILLS-Konsortium – als Leiterin der eingehenden Forschung zur Rolle digitaler Kompetenzen für das Wohlbefinden schutzbedürftiger oder gefährdeter Kinder – und ist Partnerforscherin am ARC Centre of Excellence for the Digital Child. Sie hat eng mit der Europäischen Kommission, der Gemeinsamen Forschungsstelle (JRC), dem Europarat (CoE) und der OECD über die Auswirkungen von KI auf Kinder zusammengearbeitet und ist Mitglied der Beratungsgruppe für den BIK Knowledge Hub (zuvor war sie Mitglied der BIK Map Advisory Group). Mit fast 10.000 wissenschaftlichen Zitierungen prägt ihre Arbeit globale akademische und politische Debatten.

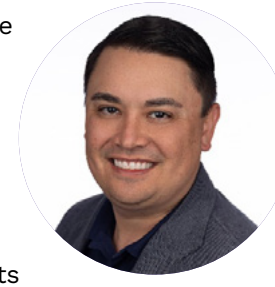
Institution: Università Cattolica del Sacro Cuore/EU Kids Online
Stakeholder-Klassifizierung: Forschung
Beschreibung der Institution: Die Università Cattolica del Sacro Cuore ist eine führende europäische Universität mit starker interdisziplinärer Forschung in den Sozialwissenschaften. EU Kids Online ist ein multinationales Forschungsnetzwerk, das das Wissen über die Online-Chancen und -Risiken von Kindern und Jugendlichen fördert.

Jennifer Kaberi ist eine der prominentesten Stimmen Afrikas für Kinder und digitale Rechte und lebt in Nairobi, Kenia. Sie ist Gründerin von Mtoto News, einer Plattform, die direkt mit Kindern und Jugendlichen in digitalen Räumen arbeitet, um ihre Erfahrungen mit Technologie und KI zu verstehen. Ihre Arbeit verbindet Bildung, Forschung und politische Interessenvertretung, einschließlich der Zusammenarbeit mit der London School of Economics. Jennifer Kaberi



setzt sich aktiv für die Teilnahme von Kindern an nationalen und regionalen KI-Politikprozessen ein und hat über 12.000 LinkedIn-Follower. Kaberi wurde bei den Expert:innen-Interviews von Joseph Atulo und Liz Busisa vertreten und/oder begleitet.

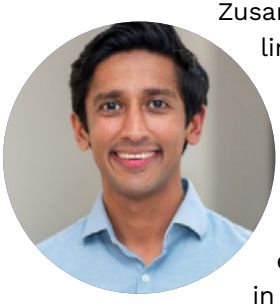
Institution: Mtoto News
Stakeholder-Klassifizierung: NGO
Beschreibung der Institution: Mtoto News arbeitet mit Kindern zusammen, um ihre Stimmen in digitalen und KI-bezogenen Debatten zu stärken. Die Organisation konzentriert sich auf Bildung, sicheren Einsatz von KI und politische Interessenvertretung in ganz Afrika.



Robbie Torney ist Leiter des Bereichs KI und digitale Bewertungen bei Common Sense Media, wo er die Forschung zur KI-Sicherheit und die Methodik zur Risikobewertung der Organisation leitet. Unter seiner Führung hat Common Sense Media systematische Risikobewertungen für große KI-Plattformen wie ChatGPT, Gemini, Meta AI und Grok sowie für aufstrebende Kategorien wie KI-Begleit-Chatbots und KI-Spielzeug entwickelt und durchgeführt. Seine Arbeit umfasst Forschung, politische Interessenvertretung und Zusammenarbeit mit der Industrie. Er hat vor dem US-Kongress und dem kalifornischen Parlament zur KI-Sicherheit für Jugendliche ausgesagt und arbeitet direkt mit Technologieunternehmen zusammen, um Branchenstandards mitzugestalten. Seine Forschung konzentriert sich auf die Bewertung von KI-Systemen anhand von Rahmenwerken, die Entwicklungsangemessenheit und Kindersicherheit durch Design in den Mittelpunkt stellen. Bevor er zu Common Sense Media kam, war er über ein Jahrzehnt lang in leitenden Positionen im Bildungswesen in Kalifornien tätig und brachte so praktisches Verständnis dafür, wie Technologie Kinder und Familien beeinflusst, in Fragen der KI-Politik und der verantwortungsvollen Entwicklung ein.

Institution: Common Sense Media
Stakeholder-Klassifizierung: NGO
Beschreibung der Institution: Common Sense Media ist eine führende gemeinnützige US-Organisation, die unabhängige Bewertungen, Lehrpläne zur digitalen Staatsbürgerschaft und politische Interessenvertretung anbietet. Sie unterstützt Eltern, Pädagog:innen und politische Entscheidungsträger:innen dabei, eine sichere und sinnvolle Mediennutzung für Kinder und Jugendliche zu fördern.

Viraj Doshi ist Platform Safety Lead bei Snap Inc. und verantwortlich für Sicherheit und digitales Wohlbefinden auf Snapchat. Sein Aufgabenbereich umfasst Sicherheitsprodukte und -richtlinien, die Zusammenarbeit mit externen Partnern im Bereich Online-Schäden sowie proaktive Initiativen zum Wohlbefinden. Außerdem ist er Mitglied des Vorstands der Tech Coalition. Vor seiner Tätigkeit bei Snap leitete er bei Meta die Bereiche Jugendwohlbefinden und Content-Moderation und war als Berater für Politik und Interessenvertretung tätig. Er bringt eine globale Perspektive mit, die durch seine Arbeit in den USA, Indien und Lateinamerika geprägt ist.



Institution: Snap Inc. (Snapchat)
Stakeholder-Klassifizierung: Tech-Industrie
Beschreibung der Institution: Snap Inc. ist ein globales Technologieunternehmen, das vor allem für Snapchat bekannt ist, eine multimediale Messaging-Plattform, die weltweit von Hunderten Millionen junger Menschen genutzt wird. Das Unternehmen hat stark in Werkzeuge und Berichtsmechanismen für Jugendsicherheit investiert.

Youmna Hashem ist Senior Analyst bei RAND Europe und spezialisiert auf verantwortungsvolle KI-Einführung und aufkommende Technologien. Ihre Arbeit liegt an der Schnittstelle von öffentlicher Politik, Ethik und menschenzentrierter Innovation. Zuvor arbeitete sie am Alan Turing Institute und bei Synapse Analytics mit Schwer-



punkt auf KI-Governance in Europa und der MENA-Region. Ihre Forschung legt den Schwerpunkt auf kontextspezifische, evidenzbasierte Ansätze zur KI-Regulierung und -Nutzung.

Institution: RAND Europe
Stakeholder-Klassifizierung: Forschung
Beschreibung der Institution: RAND Europe ist eine unabhängige, gemeinnützige Forschungsorganisation, die evidenzbasierte Politikgestaltung unterstützt. Sie betreibt interdisziplinäre Forschung in den Bereichen Technologie, Gesundheit, Sicherheit und Sozialpolitik.

Sonia Livingstone ist Sozialpsychologin und eine der weltweit einflussreichsten Wissenschaftler:innen zum Thema Leben von Kindern in digitalen Umgebungen. Seit über 35 Jahren erforscht sie, wie Kinder, Jugendliche und Familien mit Medien interagieren, mit zunehmendem Fokus auf KI-vermittelte Systeme. Sie ist Professorin für Sozialpsychologie an der London School of Economics und leitet ein Forschungszentrum, das von der 5Rights Foundation finanziert wird. Ihre Arbeit basiert auf einem Kinderrechtsansatz, geprägt durch die langjährige Zusammenarbeit mit UNICEF und ihren Beitrag zur Allgemeinen Bemerkung Nr. 25 zu Kinderrechten in der digitalen Umgebung. Livingstone verbindet groß angelegte quantitative Studien mit tiefgehender qualitativer Feldarbeit unter Einbeziehung von Kindern, Jugendlichen, Eltern, Lehrer:innen und Schulen. Sie hat Regierungen, Regulierungsbehörden und internationale Organisationen weltweit beraten. Ihre Forschung untersucht kritisch Macht, Design und Verantwortlichkeit in digitalen Systemen. Sie hat umfangreich zu Online-Risiken, -Chancen und -Governance veröffentlicht. Mit über 91.500 akademischen Zitationen gehört sie zu den meistzitierten Forscher:innen weltweit auf diesem Feld. Ihre Arbeit hat grundlegend geprägt, wie die digitalen Rechte von Kindern in Politik und Praxis verstanden werden.



Institution: London School of Economics
Stakeholder-Klassifizierung: Forschung
Beschreibung der Institution: Die London School of Economics ist eine führende globale Universität, die sich auf Sozialwissenschaften spezialisiert hat. Ihre Medien- und Kommunikationsforschung hat großen Einfluss auf internationale Debatten über Digitalpolitik.



Marta Kuljon ist Politikassistentin bei der Europäischen Kommission, GD Justiz und Verbraucher, und arbeitet im Team des Koordinators der Kommission für Kinderrechte, wobei ihr Schwerpunkt auf der Einbeziehung der Kinderrechte in alle Politikbereiche der EU liegt. Sie leitet die EU-Plattform für Kinderbeteiligung und sorgt dafür, dass die Stimmen von Kindern in die Entscheidungsfindung auf EU-Ebene einfließen. Sie koordiniert Konsultationen mit Kindern und Jugendlichen und unterstützt Initiativen zur Bekämpfung von Cybermobbing und digitalen Risiken. Mit über zehn Jahren Erfahrung, die akademische Expertise in kinderbezogenen Bereichen und humanitäre Arbeit umfasst, bringt sie wertvolle praktische Erkenntnisse in die EU-Politikgestaltung ein.

Institution: Europäische Kommission – GD Justiz und Verbraucher
Stakeholder-Klassifizierung: (über)staatliche Institution
Beschreibung der Institution: Die GD Justiz und Verbraucher ist verantwortlich für die Wahrung der Grundrechte und des Verbraucherschutzes innerhalb der Europäischen Union. Sie spielt eine zentrale Rolle bei der Gestaltung der EU-Politik zu Kinderrechten und digitaler Sicherheit.

„Es ist ein bisschen gruselig und verwirrend. Aber ich glaube, Neugier ist eine gute Haltung.“

Jugendlicher (18), Großbritannien

SOS-Kinderdörfer weltweit
Hermann-Gmeiner-Fonds Deutschland e. V.
Ridlerstraße 55
80339 München
Tel.: +49 89 17914-140
info@sos-kinderdoerfer.de
sos-kinderdoerfer.de

Terre des Hommes Deutschland e. V.
Ruppenkampstraße 11 A
49084 Osnabrück
Tel.: +49 541 7101-0
info@tdh.de
tdh.de