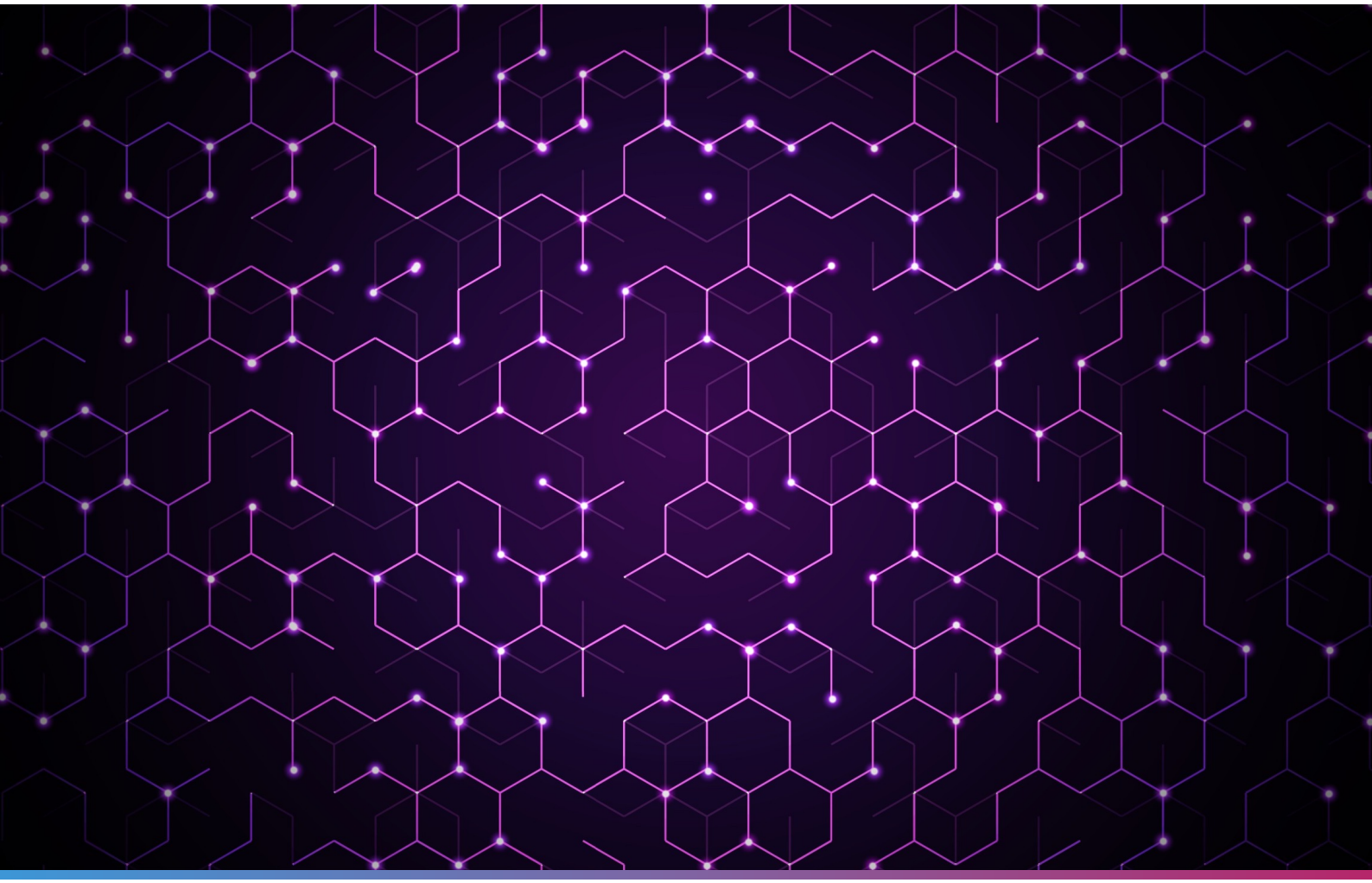
RadiGaMe



Handlungsoptionen zum Umgang mit Radikalisierung und Extremismus auf gamingnahen Plattformen und Messenger-Diensten

Befunde aus dem RadiGaMe-Forschungsverbund

Handlungsoptionen zum Umgang mit Radikalisierung und Extremismus auf gamingnahen Plattformen und Messenger-Diensten

Befunde aus dem RadiGaMe-Forschungsverbund

Stand November 2026

Lead-Autor*innen dieses Papiers: Robert Pelzer, Julian Junk, Linda Schlegel

Am Projekt beteiligte Forscher*innen: Maik Fielitz, Sebastian Golla, Simon Greipl, Matthias Heider, Judith Jaskowski, Julian Junk, Charlotte Korenke, Marius Kühne, Vivienne Ohlenforst, Michèle Leaman, Robert Pelzer, Hannes Jakim Peter, Diana Rieger, Linda Schlegel, Ulrike Schwertberger, Tobias Weidmann, Kevin Weyda, Lars Wiegold, Constantin Winkler

Zielsetzung der Handreichung

Der interdisziplinäre Forschungsverbund RadiGaMe – Radikalisierung auf Gaming-Plattformen und Messenger-Diensten wurde von Mai 2023 bis Oktober 2026 vom Bundesministerium für Forschung, Technologie und Raumfahrt (BMFTR) im Rahmen des Programms „Forschung für die zivile Sicherheit“ der Bundesregierung in der Fördermaßnahme „Zivile Sicherheit – Bedrohungen aus dem digitalen Raum“ gefördert. Koordiniert wurde der Verbund durch das Zentrum Technik und Gesellschaft der Technischen Universität Berlin; beteiligt waren zudem das Institut für Demokratie und Zivilgesellschaft Jena, das Landeskriminalamt Berlin, das Leibniz-Institut für Friedens- und Konfliktforschung PRIF, die Ludwig-Maximilians-Universität München, modus|zad – Zentrum für angewandte Deradikalisierungsforschung, Munich Innovation Labs und die Ruhr-Universität Bochum.

Der Verbund hat Radikalisierungsprozesse auf Messenger-Diensten sowie Gaming- und gamingnahen Plattformen untersucht und rechtskonforme, ethisch reflektierte Ansätze für Früherkennung, Strafverfolgung und Prävention entwickelt. Eine besondere Stärke des Verbunds lag darin, beide Kommunikationsräume – Messenger-Dienste und gamingnahe Kommunikation – gemeinsam zu untersuchen. Dabei wurden plattformübergreifende Verbindungen ebenso berücksichtigt wie die Unterscheidung zwischen radikalen Positionen und einer für die liberale Demokratie und den gesellschaftlichen Zusammenhalt schädlichen Radikalisierung. Im Mittelpunkt standen Radikalisierungsprozesse, die mit Gewaltbefürwortung, strafbaren Handlungen oder demokratiegefährdenden Bestrebungen verbunden sind. Hierfür hat der Verbund Grundlagenforschung zu Radikalisierungsdynamiken mit anwendungsorientierter Entwicklung – von KI-gestützten Screening-Instrumenten und rechtlichen und ethischen Leitplanken bis zu Präventionsansätzen – verbunden und Sicherheitsbehörden sowie Präventionspraxis von Beginn an intensiv eingebunden.

Zentraler Kontext und Befunde

In öffentlich zugänglichen Kommunikationsräumen auf Messenger-Diensten sowie Gaming- und gamingnahen Plattformen finden sich in erheblichem Umfang menschenfeindliche, insbesondere antisemitische und rechtsextreme, islamistische und misogyne Inhalte, die teilweise ironisch gerahmt und uneinheitlich moderiert werden und denen selten widersprochen wird. Da diese Inhalte meist ohne Zugangshürden auffindbar sind, ist das Risiko hoch, ihnen auch unbeabsichtigt zu begegnen. Weder Messenger- noch Gaming-Räume radikalisieren per se – entscheidend sind die sozialen Dynamiken in diesen digitalen Gemeinschaften. Digitale Räume und auch Gaming-Inhalte und -Aktivitäten sind zugleich eine Ressource: Sie stiften Zugehörigkeit und eröffnen lebensweltnahe Zugänge für Prävention.

Die Empfehlungen richten sich an Politik, Verwaltung, Sicherheitsbehörden, Zivilgesellschaft, Plattformbetreiber und Wissenschaft auf europäischer, Bundes-, Landes- und kommunaler Ebene. Sie verbinden grundsätzliche Empfehlungen mit konkreten Anknüpfungspunkten an laufende Gesetzgebungs- und Förderprozesse (Stand: November 2026).

Überblick

A. Übergreifende Empfehlungen (3-6)

B. Gaming-Plattformen, Regulierung und Games-Branche (7-9)

C. Internetmonitoring und Nutzung KI-basierter Tools in Polizeibehörden (9-12)

D. Prävention und Distanzierung (12-14)

EMPFEHLUNGEN UND HANDLUNGSOPTIONEN

A. Übergreifende Empfehlungen und Handlungsoptionen

1 Interventionen auf die Radikalisierungsdynamiken digitaler Räume ausrichten

Aktivitäten in digitalen Räumen können Identität, Zugehörigkeit und Anerkennung stiften, sie können Gruppennormen erzeugen und für Handlungen außer- und innerhalb des Netzes mobilisieren. Sicherheits- und demokratiegefährdende Radikalisierung ist daher weniger Folge einzelner Inhalte als ein sozialer Prozess, in dem sich extremistische Orientierungen und Handlungsbereitschaft im Austausch mit anderen verstärken. Repressive Maßnahmen wie Löschung und Strafverfolgung und präventive Ansätze wie persönliche Ansprache und Distanzierungsangebote sollten deshalb aufeinander abgestimmt werden. Ziel muss sein, radikalisierungsfördernde Prozesse in einem Raum zu irritieren oder zu unterbinden, nicht nur auf einzelne Äußerungen oder Personen einzugehen. Interventionen sollten plattformübergreifend ansetzen und dabei die Wechselwirkungen zwischen offenen und geschlossenen Kommunikationsräumen berücksichtigen, da sich die dortigen Dynamiken gegenseitig verstärken können. Im Gaming-Bereich sollte außerdem das gesamte Spektrum digitaler Gaming-Räume – von Spielen über Gaming- und gamingnahe Plattformen bis hin zu Gaming-Inhalten auf Social Media – in die Maßnahmen einbezogen werden. Zugleich sollten diese Räume nicht pauschal problematisiert werden: Gaming-Räume etwa sind nach den Befunden nicht per se anfälliger als andere digitale Räume, und alarmistische Debatten erschweren den Zugang zu Communities.

2 Wissenschaftlich fundierte Kriterien für die Bewertung digitaler Risikoräume entwickeln

Für die Regulierung digitaler Plattformen sowie für präventive und repressive Interventionen wird ein empirisch fundierter, abgestufter Bewertungsrahmen benötigt, der unterschiedliche Risiken erfasst und differenziert bewertet – von der Normalisierung menschenfeindlicher und extremistischer Narrative über radikalisierungsfördernde Gruppendynamiken und Mobilisierungspotenziale bis hin zu Hinweisen auf Gewaltorientierung und operative Vorbereitung. Extremistische Kommunikation, strafrechtliche Relevanz und gefahrenabwehrrechtlicher Handlungsbedarf sollten nicht gleichgesetzt werden. Die Kriterien sollten plattformspezifisch sein, methodisch überprüft und regelmäßig aktualisiert werden; sie ersetzen weder die fachliche Bewertung im jeweiligen Sektor noch die Prüfung rechtlicher Eingriffsvoraussetzungen.

Konkrete Handlungsoptionen:

- *Bund (BMI, BMFTR):* Weiterentwicklung und Erprobung eines gemeinsamen Bewertungsrahmens als Transfervorhaben beauftragen – unter Beteiligung von Bundeskriminalamt, Landeskriminalämtern, Verfassungsschutz, Präventionspraxis und Wissenschaft – und seine Nutzung in den Gremien der Innenministerkonferenz abstimmen. Die Kriterien sollten wissenschaftlich entwickelt und offengelegt werden; ihre regelmäßige Aktualisierung sollte finanziert werden.
- *EU-Kommission:* Den Bewertungsrahmen als fachlichen Maßstab bei der Prüfung der Risikobewertungen nach Art. 34/35 Digital Services Act (DSA) heranziehen, insbesondere für Messenger-Dienste mit öffentlichen Kanälen sowie Gaming- und gamingnahe Dienste.

3 Ein wissenschaftlich verankertes, risikobasiertes Längsschnittmonitoring institutionell verankern

Empfohlen wird der Aufbau einer wissenschaftlich verankerten Monitoring- und Analysestelle, die anhand eines gemeinsam entwickelten Risikobewertungsrahmens (vgl. Nr. 2) Kommunikationsräume auf Social-Media-Plattformen, Messenger-Diensten sowie Gaming- und gamingnahen Plattformen – einschließlich neu entstehender Plattformen und digitaler Formate – längsschnittlich und phänomenübergreifend untersucht, um Räume mit akteur*innenübergreifendem Interventionsbedarf zu identifizieren und zu priorisieren sowie konkrete Ansatzpunkte für die Abstimmung geeigneter Interventionen (hierzu Nr. 4) herauszuarbeiten. Hierzu sollten Kommunikationsdaten fortlaufend standardisiert erhoben und nach abgestimmten Risikokriterien (vgl. Nr. 2) bewertet werden. Dabei werden auf zwei Ebenen Ziele verfolgt: Auf Plattformebene sollen Risikostrukturen und die Verbreitung relevanter Inhalte erfasst und vergleichend eingeordnet werden. Auf Ebene konkreter Kommunikationsräume sollen Entwicklungen sowie die plattformübergreifende Vernetzung von Communities und zentralen Akteur*innen kontinuierlich beobachtet werden, um neue oder veränderte Interventionsbedarfe zeitnah zu erkennen und zu priorisieren. Dabei sollten auch Verlagerungseffekte von Moderations- und Sperrmaßnahmen – etwa die Abwanderung von Akteur*innen und Communities in weniger regulierte Räume – erfasst werden, da sie die Wirkung einzelner Interventionen relativieren können. Die Ergebnisse sollen eine gemeinsame Evidenzgrundlage für wissenschaftliche Einordnung, Plattformregulierung, Strafverfolgung, Gefahrenabwehr und Prävention – etwa auch für Landesmedienanstalten und die BzKJ einschließlich KidD – bereitstellen. Der spezifische Mehrwert gegenüber sicherheitsbehördlichem Phänomenmonitoring sowie wissenschaftlichem Monitoring von Radikalisierungsdynamiken (MOTRA) liegt in einem sektorübergreifend angelegten, interventionsorientierten Monitoring, das die unterschiedlichen Praxisbedarfe von Strafverfolgung, Prävention und Plattformregulierung systematisch zusammenführt, um eine gemeinsame Risikobewertung und abgestimmte Interventionen in konkreten digitalen Räumen zu ermöglichen.

Konkrete Handlungsoptionen:

- *Bund (BMFTR, BMI):* Wie oben beschrieben eine Monitoringstelle dauerhaft statt projektförmig finanzieren (bei Bedarf nach einer mehrjährigen Aufbauphase), an einer unabhängigen wissenschaftlichen Einrichtung ansiedeln und mit einem Beirat aus Sicherheitsbehörden, Prävention, Jugendmedienschutz und Plattformaufsicht ausstatten.
- *Bundesregierung (EU-Ebene):* Sich bei der Weiterentwicklung des DSA für Transparenz- und Datenzugangspflichten auch kleinerer Dienste einsetzen, da viele relevante Messenger-Dienste sowie gamingnahe Kommunikations-, Streaming- und Mod-Plattformen unterhalb der Schwelle für sehr große Plattformen liegen.
- *EU-Kommission und zuständige Koordinatoren für digitale Dienste:* Betonen, klarstellen und durchsetzen, dass Messenger-Dienste mit öffentlichen Kanälen und großen offenen Gruppen insoweit als Online-Plattformen den Pflichten des DSA unterliegen – einschließlich Melde-, Transparenz- und, bei Erreichen der Schwellenwerte, Datenzugangspflichten.
- *Bund, Länder, Wissenschaft*: Die Entwicklung forschungsethischer und datenschutzrechtlicher Kriterien und Leitlinien für die Forschung in zugangsbeschränkten digitalen Kommunikationsräumen vorantreiben. Dazu gehören die Erforderlichkeit des Vorgehens, die Abwägung von Forschungs- und

Schutzinteressen sowie Schutzmaßnahmen für Beforschte und Forschende. Auf dieser Grundlage sollen Forschende die ethische Vertretbarkeit ihres Vorgehens im Einzelfall prüfen und nachvollziehbar darlegen können.

4 Akteur*innenübergreifende Interventionen in digitalen Risikoräumen etablieren und Zuständigkeiten klären

Auf Grundlage der im Längsschnittmonitoring identifizierten Risikodynamiken und Interventionsansatzpunkte (vgl. Nr. 3) sollten Sicherheitsbehörden, Präventionsträger, zuständige Regulierungs- und Jugendmedienschutzstellen sowie möglichst auch Plattformbetreibende eine gemeinsame Befassung mit priorisierten digitalen Risikoräumen etablieren. Hierfür sind ein wechselseitiges Verständnis der jeweiligen Aufträge und Handlungslogiken sowie klare Zuständigkeiten und Schnittstellen erforderlich. Die Akteur*innen sollten ihre Interventionen unter Wahrung ihrer unterschiedlichen Ziele, Aufgaben, rechtlichen Voraussetzungen und professionellen Grenzen aufeinander abstimmen. Je nach Risikokonstellation sollten geeignete Kombinationen aus Strafverfolgung, Löschersuchen, plattformseitiger Sanktionierung von Regelverstößen, Jugendmedienschutz, Online-Ansprache und Distanzierungsangeboten – auch unter Einbeziehung gamebasierter Formate – entwickelt und erprobt werden. Angesichts der begrenzten Evidenzbasis für indizierte Prävention in Messenger- und Gaming-Räumen sollten verschiedene Präventions- und Interventionsformen in ihrem akteur*innenübergreifenden Zusammenspiel erprobt und evaluiert werden. Strafverfolgung kann einen Zugang zu Beratung eröffnen, indem namhaft gemachten Personen gezielt Beratungsangebote unterbreitet werden. Dies muss rechtlich zulässig erfolgen und sozialarbeiterische Vertrauensbeziehungen müssen in der Zusammenarbeit mit Sicherheitsakteur*innen geschützt bleiben. Auch sichtbare polizeiliche Präsenz in digitalen Räumen, wie sie in anderen EU-Staaten bereits umgesetzt wird, kann geprüft werden. Akteur*innenübergreifende Interventionen setzen vor allem sekundär- und tertiärpräventiv an. Sie ersetzen aber nicht universelle Prävention, Demokratieförderung und politische Bildung, die eigenständig und verlässlich gefördert werden müssen (vgl. Abschnitt D).

Konkrete Handlungsoptionen:

- *Bund und Länder (Innen- und Justizressorts):* Zuständigkeiten für akteur*innenübergreifende Interventionen in digitalen Räumen klären, wenn eine eindeutige Zuordnung zu einzelnen Bundesländern nicht möglich ist: Polizeibehörden auf Bundesebene sollten in Kooperation mit spezialisierten staatsanwaltschaftlichen Zentralstellen (wie der ZAC-NRW oder der hessischen ZIT) akteur*innenübergreifende Interventionen unter Einbeziehung von Präventionsakteur*innen, der Bundesnetzagentur sowie ggf. von weiteren Akteur*innen initiieren. Hierfür sollte im Bundeskriminalamt beim polizeilichen Staatsschutz oder bei der Zentralen Meldestelle für strafbare Inhalte im Internet eine Koordinierungsstelle eingerichtet werden. Diese sollte insbesondere bei unklarer örtlicher Zuständigkeit die Federführung klären und Expert*innen aus Wissenschaft und zivilgesellschaftlicher Präventionspraxis systematisch einbinden.
- *Bund und Länder (Innen- und Sozialressorts, Landespräventionsräte, Landes-Demokratiezentren und Sicherheitsbehörden):* Schnittstellen zwischen akteur*innenübergreifenden Interventionen in digitalen Risikoräumen und dem multiprofessionellen Fallmanagement herstellen, damit dabei identifizierte straffällig gewordene oder sicherheitsgefährdende Nutzende bei Bedarf in eine

personenbezogene Fallbearbeitung übergeleitet werden können. Damit verbunden sind:

  - Klare Standards für die Kooperation zwischen polizeilichem Staatsschutz und Distanzierungs- und Ausstiegsberatung entwickeln und in verbindlichen Kooperationsvereinbarungen verankern, insbesondere zu Übergabewegen, Interventionsschwellen und Datenschutzregeln.
  - Rechtsgrundlagen für Datenübermittlungen im multiprofessionellen Fallmanagement prüfen und klarstellen, auch mit Blick auf den Kinderschutz.

- *Polizeibehörden:* Erkennbare polizeiliche Online-Präsenz hinsichtlich Wirkung und Nebenfolgen evaluieren – in Messenger-Diensten etwa über erkennbare Behördenprofile und ggf. Ansprachen in öffentlichen Kanälen und Gruppen, auf Gaming-Plattformen über erkennbare Profile auf Community-Servern und bei Community-Events.
- *Wissenschaft und Fördergeber (BMFTR, BMI, BMBFSFJ):* Akteur*innenübergreifende Interventionen von Beginn an wissenschaftlich begleiten und evaluieren (Prozess- und Wirkungsevaluation), um eine Evidenzbasis dafür aufzubauen, welche Kombinationen von Maßnahmen in welchen Risikokonstellationen wirken; Erkenntnisse der Radikalisierungs- und Präventionsforschung systematisch in die Planung einbeziehen.

5 Transsektoralen Wissenstransfer ausbauen und dauerhaft verankern

Damit Wissen nicht in einzelnen Sektoren verbleibt oder mit dem Ende befristeter Projekte verloren geht, sollte der transsektorale Wissenstransfer dauerhaft verankert werden. Regelmäßige, strukturierte und vertrauensvolle Austauschformate sollten Forschung, Sicherheitsbehörden, Präventionspraxis, Plattformbetreiber und Regulierungsbehörden verbinden. Der Austausch sollte wissenschaftliche Erkenntnisse und Praxiswissen zusammenführen, aktuelle Entwicklungen aufnehmen und das Verständnis der jeweiligen Aufgaben und Grenzen fördern. Dabei geht es um einen wechselseitigen Transfer, der auch praktische Erfahrungen und offene Fragen in die Forschung zurückführt. Hierfür kann an die im RadiGaMe-Verbund entwickelten Formate wie das CrossPlay-Netzwerk angeknüpft werden. Diese sollten ausgebaut und über die Projektlaufzeit hinaus institutionell und finanziell abgesichert werden.

Konkrete Handlungsoptionen:

- *BMFTR*: Förderlinien der zivilen Sicherheitsforschung um verbindliche Transfer- und Verstetigungsphasen ergänzen, ggf. mit kurzer Evaluation und nicht mit eigenen kompletten Neuanträgen in eigenständigen Förderlinien, damit erfolgreiche Netzwerke nicht mit Projektende abbrechen.
- *Länder und Plattformbetreiber:* Hochschulen der Polizei, Justizakademien, Landesjugendämter und Landesmedienanstalten als regionale Transferknoten nutzen und bundesweit rollierende Weiterbildungs- und Workshopformate entwickeln, in die auch Messenger-Dienste, Gaming-Plattformen, Studios und Community-Management regelmäßig eingebunden werden.

B. Gaming-Plattformen, Regulierung und Games-Branche

6 Gaming- und gamingnahe Plattformen konsequent in die Plattformaufsicht einbeziehen

Plattformregulierung und -aufsicht konzentrieren sich bislang auf große soziale Netzwerke. Kommunikations-, Vertriebs-, Streaming- und Mod-Plattformen sowie Spiele mit Community-Funktionen rücken jedoch zunehmend in den Fokus der Plattformaufsicht. Bestehende Instrumente sollten auf sie konsequent angewendet werden. Die Einstufung einer ersten Gaming-Plattform als sehr große Online-Plattform, erste Aufsichtsverfahren der Bundesnetzagentur gegen eine große Vertriebsplattform für Spiele und der Vorschlag für einen EU KIDS Act zeigen, dass Gaming in der Plattformregulierung an Bedeutung gewinnt; diese Entwicklung sollte systematisch weiterverfolgt werden.

Konkrete Handlungsoptionen:

- *EU-Kommission:* Im Rahmen der Aufsicht über Plattformen darauf hinwirken, dass extremistische und menschenfeindliche Inhalte und Interaktionen bei Risikobewertungen, Risikominderungsmaßnahmen und Audits ausdrücklich berücksichtigt werden. Zudem die Nutzer*innenzahlangaben weiterer gamingnaher Dienste im Hinblick auf eine Einstufung als sehr große Plattform kritisch prüfen.
- *Bundesnetzagentur und BzKJ* (Minderjährigenschutz nach Art. 28 DSA): Gaming- und gamingnahe Dienste zu einem Aufsichtsschwerpunkt machen und sich dazu idealerweise mit den Koordinator*innen für digitale Dienste der Sitzstaaten abstimmen. Die Bundesnetzagentur ist dabei für Dienste zuständig, deren Niederlassung oder gesetzliche Vertretung in der EU in Deutschland liegt; dazu zählen bereits große Gaming-Vertriebsplattformen.
- *Plattformen:* Niedrigschwellige Meldewege nach Art. 16 DSA auch direkt im Spiel und für Voice-Kommunikation anbieten, mit eigener Kategorie für extremistische Inhalte und Rückmeldung an Meldende.
- *Bund (federführend BMDS):* Die Plattformaufsicht bei Bundesnetzagentur, BzKJ und KidD im Haushalt 2027 und darüber hinaus bedarfsgerecht mit Personal und Mitteln ausstatten.

7 Safety by Design und Minderjährigenschutz in digitalen Gaming-Räumen verankern

Viele Nutzende von Gaming-Plattformen sind minderjährig; identitätsbasierter Hass und extremistische Inhalte finden sich gerade auch auf Plattformen mit sehr junger Nutzer*innenschaft. Schutz sollte vor allem über sichere Voreinstellungen, Moderation und Design erreicht werden, nicht allein über Zugangsbeschränkungen.

Konkrete Handlungsoptionen:

- *Bundesregierung (BMBFSFJ, BMDS) in den Verhandlungen zum EU KIDS Act*: Die vorgeschlagenen Safety-by-Design-Pflichten für Spiele unterstützen, insbesondere sichere Voreinstellungen, Kontakt- und Interaktionsschutz sowie Schutz vor Weiterleitung auf riskante Dienste. Darauf hinwirken, dass extremistische Ansprache und problematische Gruppendynamiken ausdrücklich als Interaktionsrisiko erfasst werden. Der vorgesehene Verhaltenskodex für Spieleanbieter sollte Mindeststandards für Moderation gegen Hass und Hetze enthalten. Der vorgesehene Schutz vor Weiterleitung auf riskante Dienste sollte erhalten und gestärkt werden, da sich Dynamiken zwischen offenen und

geschlossenen Räumen gegenseitig verstärken können (vgl. Nr. 1). Hierfür sollten die laufende Konsultation zum Kommissionsvorschlag (bis 26. November 2026) und die Beratungen im Europäischen Parlament genutzt werden.

- *Bundesregierung (BMBFSFJ):* Gaming und Extremismus in die Umsetzung der im Oktober 2026 vom Bundeskabinett beschlossenen Eckpunkte zum Kinder- und Jugendschutz in der digitalen Welt aufnehmen. Altersgrenzen und datensparsame Altersprüfung sollten nur in Verbindung mit Design- und Moderationspflichten greifen, um Ausweichbewegungen auf weniger regulierte Dienste zu vermeiden. Die dort geplante kindgerechte Meldestelle bei den Sicherheitsbehörden sollte ausdrücklich auch Vorfälle in Spielen und auf Gaming-Plattformen erfassen.
- *BzKJ, Länder:* Die bei der Alterskennzeichnung von Spielen bereits berücksichtigten Interaktionsrisiken in Chat- und Community-Funktionen im Hinblick auf extremistische Ansprache konkretisieren und transparent ausweisen. Für Gaming-Plattformen, die nicht der Alterskennzeichnung unterliegen, wirksame Vorsorgemaßnahmen nach § 24a JuSchG bzw. Art. 28 DSA einfordern; das Monitoring auf Gaming-Räume ausweiten. Berücksichtigung dabei sollten auch die Leitlinien der EU-Kommission zum Schutz Minderjähriger nach Art. 28 DSA finden.

8 Games-Branche und entsprechende Communities als Partner gewinnen

Studios, Plattformen, Community-Manager*innen, Moderator*innen, Streamer*innen und E-Sport-Vereine prägen die Kommunikation, Themen und auch Normen in Gaming-Räumen. Moderation ist bislang oft intransparent und uneinheitlich, teils reproduzieren Moderierende selbst menschenfeindliche Narrative; Gegenrede ist selten. Zugleich engagieren sich viele Communities für Vielfalt und demokratische Werte – daran lässt sich anknüpfen. Forschung zeigt, dass organisierte Gegenrede und das Eingreifen von Mitspielenden (Bystander-Interventionen) die Verbreitung von Hassrede eindämmen können; dieses Potenzial wird in Gaming-Communities bislang kaum genutzt.

Konkrete Handlungsoptionen:

- *Bund (BMFTR, Games-Förderung):* Bei geförderten Spielen mit Online- und Community-Funktionen immer ein Moderations- und Schutzkonzept als Förderkriterium vorsehen und Spiele mit demokratiepädagogischem Anspruch gezielt fördern, da sie im Wettbewerb mit großen Studios oft kaum bestehen können.
- *Plattformen und Studios:* Moderierende kontinuierlich schulen und deren Handeln überprüfen; weiterhin gemeinsame Standards entwickeln und aktualisieren; Moderationsentscheidungen transparent machen; positives Sozialverhalten wie Fair Play belohnen; Frühwarnmechanismen für plötzliche Anstiege von Hass mit Schnittstellen zu Präventionsakteur*innen schaffen; professionelle Präventionsakteur*innen bei Maßnahmen auf Plattformen und in Spielen unterstützen. Moderierende zudem psychosozial unterstützen, da die dauerhafte Konfrontation mit Hass- und Gewaltinhalten belastet und die Qualität von Moderationsentscheidungen beeinträchtigen kann.
- *Streaming-Plattformen:* Monetarisierung (Spenden, Abonnements) bei wiederholten Verstößen sperren und die Umgehung von Sperren durch Gastauftritte gesperrter Akteur*innen unterbinden; Kooperationen mit Creator*innen, die Zivilcourage vorleben, fördern.
- *Bund, Branche, Zivilgesellschaft:* Die Einführung eines Labels für gut moderierte, sichere Gaming-Communities prüfen.

- *Länder, Sportverbände, E-Sport-Verbände*: Die seit 2026 geltende Gemeinnützigkeit des E-Sports nutzen: E-Sport-Vereine wie andere Sportvereine in Präventions- und Demokratiearbeit einbeziehen; Schulungen für Trainer*innen und Moderierende anbieten, Verhaltenskodizes entwickeln, Ansprechpersonen benennen und Zugang zu Präventionsprogrammen ermöglichen.
- *Zivilgesellschaft, Plattformen, Fördergeber:* Trainings und Werkzeuge für Gegenrede und Bystander-Interventionen in Gaming-Communities entwickeln und fördern, etwa für Community-Manager*innen, Streamer*innen und E-Sport-Teams.

C. Internetmonitoring und Nutzung KI-basierter Tools in Polizeibehörden

Für alle folgenden Empfehlungen gilt: Der Einsatz KI-gestützter Werkzeuge in Polizeibehörden setzt klare Rechtsgrundlagen, abgestufte Eingriffsschwellen, Transparenz und eine menschliche Letztentscheidung voraus (vgl. Nr. 12). Klassifikatoren dienen der Vorselektion; Bewertungen und Maßnahmen bleiben Aufgabe fachlich qualifizierter Auswertender.

9 Polizeiliches Internetmonitoring risikobasiert weiterentwickeln und Eingriffsschwellen abstufen

Das polizeiliche Internetmonitoring sollte risikobasiert weiterentwickelt werden, um die Früherkennung sicherheitsgefährdender Radikalisierungsdynamiken im digitalen Raum zu verbessern. Dabei ist zu prüfen, ob insbesondere Gaming- und gamingnahe Kommunikationsräume entsprechend ihrer Risikolage hinreichend erfasst werden. Hierzu sollten bestehende fachliche Kriterien unter Einbeziehung wissenschaftlicher und rechtlicher Erkenntnisse zu einem einheitlichen, abgestuften Kriterienkatalog weiterentwickelt und zwischen BKA und Landeskriminalämtern abgestimmt werden. Dieser sollte die Einbeziehung von Plattformen, die Auswahl und Beobachtung konkreter Datenquellen sowie die anlassbezogene Vertiefung durch weiterführende Analysen und personenbezogene Recherchen steuern. Die fachlichen Kriterien sollten dabei von vornherein mit den jeweils rechtlich erforderlichen tatsächlichen Anhaltspunkten verknüpft werden und die unterschiedlichen Eingriffsschwellen operationalisieren.

Konkrete Handlungsoptionen:

- *Kommission Staatsschutz*: Bestehende fachliche Kriterien zu einem einheitlichen, risikobasierten Kriterienkatalog weiterentwickeln und regelmäßig evaluieren. Der Katalog sollte die Plattformauswahl steuern und abgestufte Eingriffsschwellen operationalisieren, indem er die jeweils erforderlichen tatsächlichen Anhaltspunkte konkret beschreibt – für die Aufnahme und Fortführung der Quellenbeobachtung sowie deren anlassbezogene Vertiefung etwa durch Netzwerkanalysen und personenbezogene Recherchen.
- *Innenministerien von Bund und Ländern, BMJV*: Unter Einbeziehung rechtswissenschaftlicher Expertise den gesetzlichen Regelungsbedarf für unterschiedliche Formen des Internetmonitorings einschließlich automatisierter und KI-gestützter Analysen prüfen. Für nachgewiesene Regelungslücken sollten spezifische, am Eingriffsgewicht orientierte Befugnisse vorbereitet werden, die zulässige Erhebungs- und Analyseverfahren sowie deren Voraussetzungen und Grenzen bestimmen. Anforderungen an Datenqualität, Schutz- und Kontrollvorkehrungen sowie parlamentarische Berichts- und Evaluationspflichten sollten die Regelungen flankieren.

10 LLM-basierte Klassifikatoren zur Unterstützung von Online-Datenauswertungen einsetzen

Für umfangreiche digitale Kommunikationsbestände sollten auf großen Sprachmodellen (Large Language Models, LLMs) basierende Instrumente zur Vorselektion und Analyseunterstützung erprobt werden. Dabei sind strategisches und operatives Internetmonitoring sowie Auswertungen zur Prüfung strafrechtlich relevanter Inhalte nach ihren jeweiligen Zwecken und rechtlichen Voraussetzungen zu unterscheiden.

Im RadiGaMe-Verbund wurden anhand der Codebücher zu Warnverhalten und strafbaren Äußerungen LLM-basierte Klassifikatoren für Text-, Bild-, Video- und Audioinhalte entwickelt und in der Auswertungsplattform INspectre erprobt. Diese ermöglichen es, Kommunikationsdaten nach fachlichen Relevanzkriterien vorzubewerten und mittels Tags für eine manuelle Prüfung durch Auswertende zu markieren. Die Klassifikatoren können die Sichtung von Datenquellen im Rahmen eines Live-Monitorings ebenso unterstützen wie das gezielte Filtern retrograder Kommunikationsdaten nach relevanten Inhalten.

Für den produktiven Einsatz von LLM-Klassifikatoren in polizeilichen Auswertevorgängen müssen Mindestanforderungen festgelegt werden. Diese betreffen die Zuverlässigkeit (Güte) der Klassifikatoren ebenso wie eine Klärung rechtlicher Grundlagen des KI-Einsatzes sowie festgelegte Verfahren zur fachlichen Prüfung und Bearbeitung von Hinweisen auf mögliche Gefahren oder strafrechtliche Relevanz.

Konkrete Handlungsoptionen:

- *BKA und Landeskriminalämter:* LLM-basierte Klassifikatoren in unterschiedlichen Analyse- und Auswertungsvorgängen erproben, wie beim operativen oder strategischen Monitoring, bei der Identifikation von strafbaren Inhalten in von Meldestellen zugelieferten Datensätzen oder im Rahmen von Aktionstagen zur Verfolgung von Hassdelikten.
- *Polizeiliche IT-Verantwortliche von Bund und Ländern:* In Abstimmung mit P20 eine resiliente KI-Betriebsarchitektur entwickeln, die dezentrale, on-premise betriebene Open-Weight-Modelle mit gemeinsam bereitgestellten Cloud-Rechenkapazitäten verbindet. Einheitliche Schnittstellen sowie räumlich getrennte Ausweichkapazitäten und erprobte Umschaltverfahren sollten bei Ausfällen den Weiterbetrieb ermöglichen.

11 LLM-basierte Klassifikatoren behördenübergreifend validieren und kontinuierlich weiterentwickeln

Für LLM-basierte Klassifikatoren und Analyse-Pipelines sollte eine behördenübergreifende, von fachlicher Phänomenexpertise gesteuerte Validierungsstruktur aufgebaut werden. Grundlage ist eine fachlich annotierte, versionierte Referenzdatenbasis zu polizeilich relevanten Inhalten wie etwa Äußerungsdelikten, Gefahrenhinweisen oder spezifischen Phänomenbezügen über Plattformen, Phänomenbereiche und Modalitäten hinweg, die laufend um neue sprachliche, visuelle und subkulturelle Ausdrucksformen ergänzt wird. Die Klassifikatoren sollten regelmäßig anhand unabhängiger Testdaten hinsichtlich ihrer Klassifikationsgüte und möglicher systematischer Verzerrungen validiert werden; Änderungen an Modellen, Prompts und Kriterien sollten dokumentiert werden. Die Ergebnisse sollten ausweisen, für welche Aufgaben und Datenarten ein Klassifikator hinreichend zuverlässig ist und wo Einsatzgrenzen liegen.

Konkrete Handlungsoptionen:

- *BKA, ZITiS und Landeskriminalämter:* Eine gemeinsame fachliche Kompetenzstelle für Monitoring, Validierung und Weiterentwicklung von im polizeilichen Wirkbetrieb verwendeten KI-basierten Klassifikatoren einrichten. Die Kompetenzstelle muss IT-Fachlichkeit zwingend mit Phänomenexpertise, kriminologischer und sozialwissenschaftlicher Fachlichkeit und ethischer Expertise verknüpfen.

12 Den Einsatz von KI im polizeilichen Internetmonitoring durch Legality und Ethics by Design absichern

Für die rechtliche und ethische Ausgestaltung des Einsatzes von KI im polizeilichen Internetmonitoring sowie in polizeilichen Datenauswertungen allgemein sollte ein wissenschaftlich gestütztes Gesamtkonzept entwickelt und technisch-organisatorisch umgesetzt werden. Für unterschiedliche Auswertungszwecke, etwa strategische und operative Analysen, sollten jeweils fachlich geeignete und ethisch vertretbare Güteanforderungen an eine KI-gestützte Vorselektion von Inhalten festgelegt werden. Da KI zugleich eine erhebliche Ausweitung des Beobachtungs- und Auswertungsumfangs ermöglicht, muss neben der Qualität auch die Begrenzung ihres Einsatzes sichergestellt werden. Hierzu sollten risikobasierte Kriterien für die Auswahl und Beobachtung digitaler Kommunikationsräume und die Vertiefung ihrer Auswertung festgelegt werden. Die Kriterien sollten mit den jeweils rechtlich erforderlichen tatsächlichen Anhaltspunkten verknüpft sein und so abgestufte Eingriffsschwellen für Art, Umfang und Dauer der Beobachtung von digitalen Räumen sowie für den Übergang zum Monitoring einzelner Personen konkretisieren.

Für ein Auswerteprojekt bspw. in der Plattform INspectre sollten Zweck, Rechtsgrundlage, Datenumfang, Erhebungszeitraum und Speicherdauer und die verwendeten LLM-Klassifikatoren dokumentiert und die Verhältnismäßigkeit nachvollziehbar dargelegt werden. Jede Datenquelle sollte vor ihrer Aufnahme manuell geprüft werden; tatsächliche Anhaltspunkte sollten bestätigt werden. Gesondert zu rechtfertigen ist die Einbeziehung von personenbezogenen Profilen. Verlinkungen allein sollten keine automatische Aufnahme auslösen. Für als strafrechtlich relevant oder mit Gefahrenhinweis getaggte Inhalte sollten Prüf- und Bearbeitungsabläufe vorgesehen werden, in denen Auswertende ihre Bewertung und daraus folgende Entscheidungen dokumentieren können.

Konkrete Handlungsoptionen:

- *Bundesgesetzgeber, Sicherheitsbehörden von Bund und Ländern / Innen- und Justizministerien in Bund und Ländern:* Entwicklung eines wissenschaftlich gestützten Gesamtkonzeptes für die KI-Unterstützung bei polizeilichen Datenauswertungen.
- *Industrie, BMI und BMFTR:* Die Entwicklung und Erprobung von Legality- und Ethics-by-Design-Funktionen für Monitoring-Tools wie INspectre vorantreiben. BMI und BMFTR sollten entsprechende Entwicklungsvorhaben fördern; Industriepartner*innen sollten die Funktionen gemeinsam mit polizeilichen Auswertungsstellen und rechtswissenschaftlicher Expertise umsetzen.
- *Sicherheitsbehörden von Bund und Ländern / BMI:* Sicherstellen, dass für die Anwendungsfälle und Analysefunktionen KI-gestützter Monitoring-Tools geklärt ist, ob sie als Hochrisiko-Systeme nach der KI-Verordnung einzustufen sind. Für entsprechend eingestufte Anwendungen sicherstellen, dass die erforderlichen Dokumentationsverfahren, menschliche Aufsicht und Grundrechte-

Folgenabschätzungen rechtzeitig vor Geltung der einschlägigen Pflichten am 2. Dezember 2027 umgesetzt sind.

13 Plattform- und Phänomenwissen intern verfügbar machen und kontinuierlich aktualisieren

Wissen über subkulturelle Codes, Kommunikationsformen und technische Eigenschaften von Plattformen sollte für den behördeninternen Transfer aufbereitet und im Austausch mit Wissenschaft und Zivilgesellschaft aktuell gehalten werden, um neue Entwicklungen – etwa neue Räume auf Gaming-Plattformen – frühzeitig aufzugreifen. Länder und Bund sollten entsprechende Module in Aus- und Fortbildung an Polizei- und Verwaltungshochschulen, der Hochschule des Bundes und Justizakademien verankern. Der Austausch mit Wissenschaft und Zivilgesellschaft sollte institutionalisiert werden, da diese wesentlich zur Erschließung, Einordnung und Vermittlung aktuellen Plattform- und Phänomenwissens beitragen (siehe auch Nr. 5).

D. Prävention und Distanzierung

14 Strukturelle Voraussetzungen für Präventionsarbeit im Gaming schaffen

Präventionsfachkräfte nennen zahlreiche Hürden, Gaming-Räume zu nutzen: fehlendes subkulturelles Wissen und wenig Zeit, es sich anzueignen; kaum Fortbildungsangebote; fehlende und oft nicht förderfähige Technik (Gaming-PC, Konsolen, Internet-Zugänge, Headsets, digitale Spiele, Streaming-Ausstattung); Unsicherheit über Plattformen und Kommunikationsnormen; praktische Hürden wie mitspielen zu müssen, um in In-Game-Chats präsent zu sein, oder flüchtige Sprachkommunikation; Sorge vor Backlash; ethische und datenschutzrechtliche Fragen beim Eintritt in (als) privat empfundene Räume; sowie die Internationalität der Gaming-Szene, die der national organisierten Prävention widerspricht. Ohne passende Strukturen bleiben gute Ansätze wirkungslos.

Konkrete Handlungsoptionen:

- *Fördergeber (Bund: BMBFSFJ, BMI, BAMF; Länder; Kommunen*): Technische Ausstattung, Lizenzen und Arbeitszeiten am Abend, am Wochenende und in den Ferien ausdrücklich zuwendungsfähig machen; englischsprachige und grenzüberschreitende Formate zulassen; Erprobung mit offener Fehlerkultur ermöglichen.
- *BMBFSFJ, BMFTR, BMI:* Digitale und Gaming-bezogene Prävention in den Förderprogrammen des Bundes als Schwerpunkt verstetigen; Unterstützungs- und Finanzierungsstrukturen für entsprechende Projekte und Organisationen ausbauen.
- *Bund, Länder, Fachverbände:* Fachliche Standards für Präsenz und Ansprache in (halb-)öffentlichen Online-Räumen entwickeln – zu Datenschutz, Interventionsschwellen, Selbstschutz der Fachkräfte und Umgang mit Minderjährigen.

15 Beziehungsorientierte Online-Distanzierungsarbeit ausbauen

Online-Distanzierungsarbeit sollte die Bedeutung von Zugehörigkeit in digitalen Gemeinschaften berücksichtigen und persönliche Ansprache sowie Beziehungsaufbau stärken, insbesondere dort, wo Personen für Gespräche erreichbar erscheinen. Möglichkeiten und Grenzen einer Kontaktaufnahme, gerade bei unbekannten Nutzer*innen, sind sorgfältig zu prüfen; die Befunde zeigen hier erhebliche fachliche und ethische Vorbehalte. Der Ausbau sollte daher mit professionellen Standards verbunden und evaluiert werden, ohne die Wirkung bereits als gesichert vorauszusetzen.

16 Gaming als Ressource für Prävention und Distanzierung nutzen

Gemeinsames Spielen kann lebensweltnahe Zugänge eröffnen, Kommunikation erleichtern und Beratungsbeziehungen stützen. Spielerfahrungen bieten Anknüpfungspunkte für Gespräche über Erfolgserlebnisse und erlernte Kompetenzen wie Teamwork, Perspektivübernahme, den Umgang mit Frustration sowie erfahrene Anerkennung. gamebasierte Methoden – etwa gemeinsamer Avatarbau in der Biografiearbeit oder professionell begleitetes Spielen – sollten an den Interessen der Personen ausgerichtet, erprobt und weiterentwickelt werden und Raum für kritische Reflexion lassen. Neben Serious Games sollten unter gewissen Rahmenbedingungen auch kommerzielle Spiele, Livestreams und Offline-Gaming-Formate genutzt werden.

Konkrete Handlungsoptionen:

- *Kommunen:* Bildungs- und präventionsorientierte Gaming-Angebote in Jugendzentren, Bibliotheken und Partnerschaften für Demokratie als niedrigschwellige Begegnungsorte ausbauen.
- *Länder:* Gaming und gamingnahe Kommunikation in Landesprogramme der Medienbildung und Demokratieförderung aufnehmen.

17 Fachkräfte für digitale Lebenswelten qualifizieren

Qualifizierungsangebote sollten Kenntnisse über Plattformen, Kommunikationskulturen und soziale Beziehungen in digitalen Räumen vermitteln, Methodenkompetenz für Beratung und Fallarbeit stärken und Berührungsängste sowie stereotype Vorstellungen aufgreifen. Zentral ist eine interessierte, fragende Haltung, die Gespräche über digitale Praktiken erleichtert – auch Fachkräfte ohne eigene (digitale) Gaming-Erfahrung können so ins Gespräch kommen. Qualifizierung sollte eine offene Fehler- und Lernkultur fördern.

Konkrete Handlungsoptionen:

- *Länder (Wissenschafts- und Kultusressorts, KMK):* Digitale Spielkulturen als fakultative Angebote in Studium und Ausbildung der Sozialen Arbeit, Pädagogik und des Lehramts sowie in die Lehrkräftefortbildung aufnehmen.
- *Bund, Länder:* Ein dauerhaft finanziertes Fortbildungsangebot für Fachkräfte aus Jugendhilfe, Schule und Distanzierungsarbeit schaffen. Wo dies bereits geschieht, entsprechende Angebote ausdrücklich zuwendungsfähig machen und ausbauen.

18 Eltern bei der Begleitung digitaler Gaming-Lebenswelten unterstützen

Eltern sollten unterstützt werden, die digitalen Lebenswelten ihrer Kinder aufmerksam zu begleiten, ohne Gaming pauschal zu problematisieren: nicht nur zu fragen, was, sondern auch mit wem gespielt wird und welche Erfahrungen Kinder dabei machen. Eine offene

Haltung und gemeinsames Spielen erleichtern Gespräche über Zugehörigkeit, Gruppennormen und problematische Erlebnisse.

Konkrete Handlungsoptionen:

- *Bund (BMBFSFJ), Länder, Kommunen:* Gaming in die von der Expert*innenkommission empfohlene frühe Elternberatung aufnehmen; Familienbildung, Schulen und Jugendhilfe sollten kurze, leicht zugängliche Informationen und Peer-Formate wie gemeinsame Spielabende anbieten. Qualitätsgesicherte, leicht zugängliche Informationsplattformen mit Materialien für Eltern und Fachkräfte schaffen.

19 Digitale Präventionsangebote dauerhaft fördern

Digitale Präventions- und Distanzierungsangebote brauchen verlässliche Förderbedingungen. Neben neuen Formaten muss die kontinuierliche Weiterentwicklung bestehender Ansätze möglich sein, damit wichtige Strukturen und Praxiswissen nicht mit Projektlaufzeiten verloren gehen; wiederkehrender Innovationsdruck darf nicht zulasten bewährter Angebote gehen. Förderung sollte kontinuierliche Beziehungsarbeit sowie den Aufbau von Kompetenzen und die Weiterentwicklung gamebasierter Ansätze ermöglichen.

Konkrete Handlungsoptionen:

- *Bund, Länder:* Mehrjährige Förderzeiträume (mindestens drei bis fünf Jahre) mit Verstetigungsperspektive vorsehen und Wirkungsevaluation verbindlich mitfinanzieren.