

**“DET ÄR COOLT MED TEKNIKEN,
MEN MAN MÄRKER ATT MÄNNISKORNA
SOM STYR DEN INTE HINNER MED”**

**Artificiell intelligens
och kampen mot
sexuella övergrepp
mot barn**

CHILDHOOD CHILD X

WORLD CHILDHOOD FOUNDATION
FOUNDED BY H.M. QUEEN SILVIA OF SWEDEN

INNEHÅLLSFÖRTECKNING

3 FÖRORD

4 INLEDNING

7 SÅ ANVÄNDER FÖRÖVARE AI FÖR ATT UTSÄTTA BARN FÖR SEXUELLA ÖVERGREPP

9 SÅ ANVÄNDS AI FÖR ATT FÖREBYGGA OCH STOPPA SEXUELLA ÖVERGREPP MOT BARN

13 SÅ STÄRKER AI STÖD TILL BARN OCH UNGA SOM UTSATTS FÖR SEXUELLA ÖVERGREPP

15 UNGAS REFLEKTIONER OCH REKOMMENDATIONER

19 BARNRÄTTSLIG SLUTSATS OCH VÄGEN FRAMÅT

21 REKOMMENDATIONER

23 KÄLLFÖRTECKNING

Utgivningsår: 2026

Ansvarig utgivare: Stiftelsen ChildX
och World Childhood Foundation

Produktion: Allison McMahan, Sophie Josephson,
Erika Olsson och Sara Hugosson

Grafisk formgivning: Scencia

Foto: Pexels, Cottonbro Studio

*För bearbetning av texterna har AI-verktyg
använts under mänsklig översyn.*

FÖRORD

AI är inte en framtidsfråga. Tekniken finns redan i allas vardag, inte minst barn och ungas. Den är en del av förövares verktyglåda, men också ett avgörande verktyg för alla som arbetar för att skydda barn mot sexuella övergrepp och exploatering.

Samma teknik som används för att skapa övergreppsmaterial, grooma barn i industriell skala och manipulera bilder används också för att upptäcka övergreppsmaterial, identifiera riskbeteenden och förbättra stödet för barn utsatta för sexuella övergrepp och exploatering.

Förövare har alltid utnyttjat nyteknologi för att nå barn. Digitaliseringen har förändrat brottsligheten i grunden. Förövare agerar nu anonymt, gränsöverskridande och i massiv skala. Samhällets förmåga att skydda barn måste ligga steget före. Den måste vara tekniskt avancerad, rättsligt robust och tydligt prioriterad.

Behovet av att agera är akut. Antalet anmälda sexualbrott mot barn i Sverige har ökat från 906 fall år 2015 till över 6 000 fall år 2024. De globala volymerna av övergreppsmaterial är enorma. Enbart under 2025 fick Internet Watch Foundation (IWF) in 312 030 rapporter om bekräftat övergreppsmaterial. 2025 rapporterade IWF dessutom en ökning på 26 362 procent av fotorealistiska AI-genererade filmer med sexuella övergrepp mot barn. Från 13 stycken år 2024 till 3 440. Siffrorna är bara toppen av ett isberg.

Den här rapporten bygger på en litteraturöversikt, samtal med barn och unga, överlevare, experter och yrkesverksamma i Sverige och internationellt. Den visar att AI redan påverkar barns säkerhet. Den avgörande frågan är nu hur samhället väljer att agera och vem som tar ansvar.

I februari antog Sverige en ny AI-strategi, med ambitionen att Sverige ska bli världsledande inom AI. Den ambitionen måste också omfatta skyddet av barn. Idag hänger lagstiftning, långsiktig finansiering och samverkan mellan myndigheter, civilsamhälle, forskning och techsektorn inte med. Det måste förändras.

Vår gemensamma övertygelse är att det är möjligt. Men det måste ske nu. Inte sen. Och det måste göras ansvarsfullt, rättssäkert och med barnets rättigheter i centrum. Skydd av barn kan inte vänta.

Britta Holmberg, global program- och påverkanschef och vice generalsekreterare World Childhood Foundation

Ida Östensson, generalsekreterare ChildX

INLEDNING

Artificiell intelligens (AI) förändrar landskapet för alla som arbetar för att stoppa sexuella övergrepp och exploatering av barn.

Samma teknik som kan användas för att skydda barn mot skadligt innehåll, minska spridningen av övergreppsmaterial och vägleda barn som utsatts för sexuella övergrepp till stöd, används också av förövare. De kan använda tekniken för att skapa AI-genererat övergreppsmaterial, anpassa grooming till enskilda barn och hitta sårbara grupper i syfte att utsätta dem för sexuella övergrepp och exploatering.

AI är idag en del av barns och ungas vardag. De använder chatbotar och AI-kompisar för skolarbete, relationsråd, sällskap och information om exempelvis psykisk hälsa (Internetstiftelsen, 2025). Tekniken är därmed inte en framtidsfråga, utan en del av den digitala verklighet som barn redan lever i. I rapporter från bland annat Ecpat (2023, 2026) uttrycker barn oro för att AI ska användas för att skapa och sprida övergreppsmaterial som föreställer dem själva. För de barn som utsätts kan konsekvenserna bli omfattande, med psykisk ohälsa, ryktesspridning och återtraumatisering. Skapandet och spridningen av AI-genererat övergreppsmaterial är ett växande problem.

Enligt artikel 34 i FN:s konvention om barnets rättigheter, Barnkonventionen, som sedan 2020 är en del av svensk lagstiftning, har varje barn rätt till skydd mot alla former av sexuellt utnyttjande och sexuella övergrepp (FN, 1989). Barnrättskommittén, som granskar staters efterlevnad av

konventionen, har i sin allmänna kommentar nummer 25 konstaterat att rätten till skydd mot sexuell exploatering och sexualiserat våld är teknikneutral (FN, 2021). Det betyder att barn ska skyddas oavsett vilken metod förövaren använder. Skyddet gäller också exploatering genom teknik som utvecklas nu och som kommer att fortsätta utvecklas i framtiden. Barn har även rätt att aktivt delta i alla frågor som rör dem (artikel 12) på lika villkor (artikel 2), där deras bästa ska sättas i främsta rummet (artikel 3) och deras liv, utveckling och värdighet ska respekteras (artikel 6) (FN, 1989).

För att samhället ska kunna ta ansvar och skydda barn mot sexuella övergrepp och exploatering behöver vi förstå både AI-teknikens dubbla natur och den digitala verklighet som barn och unga redan lever i. Mot denna bakgrund gav Childhood och ChildX Allison McMahan, PhD och legitimerad psykolog vid Karolinska Institutet, i uppdrag att genomföra en litteraturöversikt över forskning och tillgänglig rapportlitteratur från myndigheter, etablerade civilsamhällesorganisationer och andra verksamheter om hur AI används för att facilitera, upptäcka, förebygga och åtgärda sexuella övergrepp och exploatering av barn.

Litteraturöversikten, med titeln "A Rapid Review of Artificial Intelligence in Facilitation, Detection, and Prevention of Child Sexual Abuse and Exploitation" ligger till grund för slutsatserna i den här rapporten (McMahan, 2025).

Nuvarande lagstiftning

De AI-verktyg som nämns i denna rapport omfattas av flera EU-regelverk. EU har en central roll i regleringen av digital teknik eftersom digitala tjänster och AI-system används över nationsgränserna. Barns rättigheter i digitala miljöer regleras bland annat genom Digital Services Act ((EU) 2022/2065) som tydliggör plattformars ansvar på EU:s inre marknad. DSA ställer bland annat krav på att plattformar skyndsamt rapporterar material som skildrar sexuellt utnyttjande av barn.

Även EU-förordningen AI Act, ((EU) 2024/1689), är relevant eftersom den reglerar AI-system, särskilt sådana som kan påverka barns rättigheter, säkerhet och integritet. I maj 2026 enades EU-parlamentet och EU-rådet om att förbjuda distributionen av AI-funktioner som utvecklas för att generera sexuellt övergreppsmaterial mot barn eller andra sexuellt kränkande manipulationer (Europaparlamentet, 2026).

Plattformar och techbolag som utvecklar och tillgängliggör AI-verktyg har ett avgörande ansvar för att skydda barn mot grooming, spridning av övergreppsmaterial och AI-genererat eller manipulerat sexuellt kränkande innehåll. Ansvaret omfattar både förebyggande säkerhetsåtgärder, tydliga rapporteringsvägar, aktiv moderering och att skyndsamt ta ned olagligt material. Tekniska lösningar, som AI-baserad moderering och automatiserad upptäckt, behöver samtidigt kombineras med mänsklig granskning i komplexa ärenden. För att skyddet ska fungera krävs också transparens, samverkan med myndigheter och civilsamhälle samt att barns rättigheter byggs in i tjänsterna från start.

Det pågår också en översyn av det EU-direktiv (2011/93/EU) som reglerar medlemsstaternas arbete mot sexuella övergrepp och exploatering av barn online. Syftet med översikten är bland annat att uttryckligen inkludera AI-genererat övergreppsmaterial.

Parallellt förhandlas en ny EU-förordning om sexuella övergrepp mot barn, med målet att stärka plattformars ansvar att förebygga, upptäcka och rapportera sexuella övergrepp mot barn i digitala miljöer. Behovet av långsiktig lagstiftning är brådskande. Det gäller särskilt eftersom det tillfälliga undantaget från ePrivacy-direktivet, som gav plattformar möjlighet att proaktivt skanna sina tjänster efter övergreppsmaterial och grooming, löpte ut den 3 april 2026.

EU-reglerna handlar i stor utsträckning om plattformars ansvar att upptäcka, rapportera och hantera olagligt innehåll. Samtidigt är medlemsländerna avgörande för att regelverken ska få effekt i praktiken. Det är också den nationella straffrätten som i stor utsträckning avgör vad som faktiskt är straffbart. I Sverige är därför barnpornografilagstiftningen central, eftersom den definierar vad som rättsligt räknas som dokumenterade sexuella övergrepp mot barn och vilket material som är olagligt (BRB, 2020:173, kap 16 § 10a).

Dagens svenska lagstiftning har begränsningar när det gäller att skydda barn mot sexuella övergrepp som faciliteras av AI. Dessa begränsningar gör det möjligt att kringgå reglerna. Ett exempel är när AI används för att manipulera övergreppsmaterial så att det faller utanför lagtextens nuvarande utformning, trots att materialet fortfarande tydligt utgör sexuell exploatering av barn.

Ordlista

Artificiell intelligens (AI): Maskiner eller datorprogram med förmåga att efterlikna mänsklig intelligens och kognitiva funktioner och som kan utveckla sina resonemang baserat på den information de samlar in.

Avatar: En digital, interaktiv karaktär som kan efterlikna en person i utseende, beteende och röst eller vara helt syntetiskt genererad.

Biometrisk analys: Analys av en persons unika kroppsliga egenskaper, som fingeravtryck, ansiktsform eller rörelsemönster.

Deepfake: Teknik som kan ersätta en persons ansikte, röst eller kropp med någon annans i bild eller video, ofta på ett mycket verklighetstroget sätt.

Generativ AI: En gren inom AI som har förmågan att generera nytt innehåll, som texter eller bilder, istället för att bara analysera befintlig data.

Hashing: En funktion som skapar ett unikt digitalt fingeravtryck av en bild, ett hashvärde. Används för att enkelt identifiera en bild som man lätt vill kunna känna igen.

Klassificerare: En funktion eller modell med förmågan att automatiskt kategorisera data i fördefinierade grupper, till exempel för att avgöra om ett fotografi föreställer en hund eller en katt.

Maskininlärning: En metod inom AI där datorer tränas att lära sig från data och hitta mönster för att kunna göra förutsägelser eller fatta beslut utan att vara programmerade för varje steg.

Metadata: Information som beskriver annan data. Det kan inkludera detaljer som när en fil skapades, vem som äger den eller hur stor den är.

Sexuella övergrepp mot barn: Sexualiserat våld mot personer under 18 år. Sexuellt våld innefattar till exempel sexuella kränkningar, ofredande och våldtäkt.

Språkmodell: En gren inom AI som fokuserar på att analysera, tolka och förstå text och tal.

Övergreppsmaterial: Dokumenterade sexuella övergrepp på barn. I den här rapporten används övergreppsmaterial som ett samlingsbegrepp för sexuella kränkningar av barn i bild och video, som deepfakes, AI-genererade övergreppsbilder och sexuellt kränkande bilder.

Metod

Med utgångspunkt i McMahans litteraturöversikt har Childhood och ChildX hämtat in perspektiv från ChildX expertgrupp av överlevare, samt från experter inom utveckling och praktisk applicering av AI. Litteraturöversikten låg också till grund för två fysiska fokusgrupper med barn och unga tillsammans med Fryshuset. Totalt deltog 15 barn och ungdomar i åldrarna 14–17 år.

Utifrån litteraturöversikten, experter och fokusgrupper tog Childhood och ChildX fram rekommendationer. Rekommendationerna presenterades och diskuterades i workshopformat vid en internationell konferens i Stockholm den 5 mars 2026. I workshopen deltog experter och yrkesverksamma från relevanta verksamheter inom näringslivet, civilsamhället, myndigheter, akademi och offentlig sektor. Totalt medverkade 49 personer. Efter konferensen har Childhood och ChildX bearbetat rekommendationerna som presenteras i denna rapport.

SÅ ANVÄNDER FÖRÖVARE AI FÖR ATT UTSÄTTA BARN FÖR SEXUELLA ÖVERGREPP

Litteraturöversikten visar att den snabba utvecklingen av AI skapar ett nytt landskap, där den tekniska kapaciteten ofta ligger steget före vad lagstiftare hinner reglera kring. Översikten lyfter flera situationer där AI används för att möjliggöra och skala upp sexuella övergrepp mot barn.

AI-genererat och manipulerat material

Tekniken gör det möjligt för förövare att skapa övergreppsmaterial och sexuellt kränkande manipulerade bilder, till exempel genom att "klä av" ett brottsoffer för att utöva utpressning i kontakt med barn i ekonomiskt eller sexuellt syfte.

AI-genererat material som bygger på tidigare dokumenterat övergreppsmaterial kan bli mycket verklighetstroget. Förövare kan också simulera låg bildkvalitet i bilder eller filmer för att de genererade bilderna ska uppfattas som äkta. Det kan handla om att skapa nya övergreppsbilder av tidigare identifierade brottsoffer eller generera övergreppsmaterial utifrån vardagliga bilder på barn som publiceras öppet på internet. Den här formen av övergrepp kan skapa en upprepad traumatisering för barnet, vars bilder cirkulerar online.

Automatisering av grooming och utpressning

Genom avancerade chatbotar och digitala avatarer som simulerar konversationer kan förövare automatisera steg-för-steg-manualer i syfte att utsätta barn för grooming och sexuell utpressning. De generativa AI-verktygen gör det möjligt att föra trovärdiga och individuellt anpassade samtal som utgår från barnets eget sätt att prata. Denna acceleration leder till fler kontaktförsök och allt mer skraddarsydda och aggressiva metoder.

Algoritmer och rekommendationssystem

Översikten visar hur algoritmer skraddarsyr innehåll, bland annat på sociala medier. Det leder till att användare får förslag på liknande material, eller på innehåll som de spenderar mycket tid på att titta på. Unga användare är särskilt sårbara när rekommendationerna bygger på tidigare sökhistorik. Förslagen kan normalisera våldsamt innehåll och eskalera tittandet från lagligt pornografiskt innehåll till övergreppsmaterial, utan att användaren aktivt söker efter det. Denna normalisering kan bidra till att användare tar del av olagligt eller skadligt innehåll. Algoritmer kan även styra förövare direkt mot barn i sårbara grupper.

AI-genererat övergreppsmaterial utmanar nuvarande nationell och europeisk lagstiftning och juridiska definitioner av vad som ska räknas som ”riktigt” övergreppsmaterial. Detta försvårar när brottsutredare behöver samverka över nationella gränser eller dela bevismaterial.

Juridiska utmaningar och normalisering i digitala miljöer

AI-genererat övergreppsmaterial utmanar nuvarande nationell och europeisk lagstiftning och juridiska definitioner av vad som ska räknas som ”riktigt” övergreppsmaterial. Detta försvårar när brottsutredare behöver samverka över nationella gränser eller dela bevismaterial.

Litteraturoversikten lyfter flera studier som problematiserar det normaliserande språkbruk som förövare och andra aktörer använder. Det AI-genererade innehållet förminskas genom påståenden om att det inte har något verkligt brottsoffer eller inte är äkta. I forum för förövare argumenterar användare för att dessa bilder ska ses som konstnärliga och inte lika allvarliga som övergreppsmaterial som inte är AI-genererat. Ur ett barnrättsperspektiv kränker detta barnets integritet och kan därmed försvåra barnets möjligheter till rehabilitering och stöd.

Summering

- **Eftersläpande lagstiftning:** När den juridiska utvecklingen inte håller jämna steg med teknikutvecklingen bidrar det till att tekniken missbrukas för att utsätta barn för sexuella övergrepp.
- **Sexuell utpressning:** AI-genererat övergreppsmaterial kränker barns sexuella integritet och används aktivt av förövare för att utsätta barn för sexuell utpressning.
- **Konsekvenser för offren:** Missbruket av AI-teknik skalar upp groomingförsök, leder till en ökad produktion av övergreppsmaterial och återtraumatiserar brottsoffer.

SÅ ANVÄNDS AI FÖR ATT FÖREBYGGA OCH STOPPA SEXUELLA ÖVERGREPP MOT BARN

Flera sektorer använder redan AI i arbetet mot sexuella övergrepp, bland annat hälso- och sjukvården, brottsbekämpande myndigheter och digitala plattformar.

Inom dessa områden testas och används tekniken för att tidigt upptäcka brottsoffer, förövare och högriskbeteenden. Den används också för att snabbare hantera anmälda ärenden och analysera mönster i stora mängder data. AI kan dessutom minska personalens exponering för skadligt material och skapa stödstrukturer för beslutsfattande. Inom hälso- och sjukvården används chatbotar även i förebyggande utbildningsinsatser om kroppslig integritet, för att hjälpa barn att utveckla strategier som skyddar dem och minskar riskerna för att utsättas för sexuella övergrepp. Det visar att tekniken kan fungera som ett komplement i utbildningssituationer. Hittills fokuserar den tekniska utvecklingen främst på att upptäcka sexuella övergrepp. Det förebyggande arbetet är därmed ett område där AI:s användning fortfarande är mindre utvecklad.

Stoppa spridning av övergreppsmaterial mot barn

Företag, civilsamhällesorganisationer och brottsbekämpande myndigheter använder ofta så

kallade hashing för att automatiskt upptäcka och stoppa spridningen av övergreppsmaterial mot barn. Hasharna, som sammanställs till hashlistor består av tidigare identifierat övergreppsmaterial där varje bild och video har ett digitalt fingeravtryck, ett så kallat hashvärde. För att material ska läggas till en lista måste det först granskas och bedömas som illegalt av brottsbekämpande myndigheter och aktörer som samverkar med myndigheterna, till exempel hotlines. Med hashvärdet kan enskilt material och varianter av det upptäckas på plattformar och plockas ner när de delas.

För att hitta nytt eller manipulerat övergreppsmaterial, misstänkt groomingkommunikation eller annat skadligt innehåll används även så kallade klassificerare. Förenklat är det AI-verktyg som analyserar bild, text, video eller metadata och sorterar innehållet i förbestämda kategorier.

Utmaningen i att använda klassificerare för detta ändamål är att mellanmännisklig kommunikation innehåller nyanser och kontext som gör det svårt att identifiera när en konversation går från harmlös till att kategoriseras som skadlig. Analys av metadata, till exempel filnamn, är en skalbar metod som även visar goda resultat ur ett integritetsperspektiv.

Förebygga grooming och riskreducering online

Flera maskininlärningsmodeller visar goda resultat när det gäller att förebygga och upptäcka grooming. Modellerna bygger bland annat på typiska fraser som förövare använder och på hur kommunikationen utvecklas över tid.

Eftersom grooming är svårt att upptäcka möter modellerna samtidigt flera utmaningar. Förövare kan anpassa sin kommunikation till enskilda barn, använda kodord, byta plattform eller komma med antydningar.

På grund av detta har enbart textbaserade AI-modeller begränsningar i att upptäcka pågående grooming-beteende. Modeller som också analyserar mönster och beteenden i kommunikationen, till exempel vem som kontaktar vem och hur ofta, kan därför ge en bredare bild och bidra till att skilja groomingbeteenden från vanliga kontakter mellan användare.

Verktyg för utredning och prioritering inom brottsbekämpande myndigheter

AI används idag för att effektivisera prioriteringen av brottmålsärenden. Det kan handla om att snabbt transkribera kommunikation så att utredaren kan söka efter ord i beslagtagen information. Det kan också handla om att lokalisera krypterade filer eller mappar som kan innehålla övergreppsmaterial. Klassificering av metadata, till exempel filnamn, visar hög träffsäkerhet men testas än så länge endast i kontrollerade miljöer. Sådan teknik kan hjälpa utredare att hantera stora datavolymer och identifiera AI-genererat övergreppsmaterial. Verktyg för biometrisk analys kan också användas för att identifiera gärningspersoner utifrån

deras unika fysiska egenskaper, till exempel knogar och nagelbäddar. Metoderna visar goda resultat i testmiljöer.

Barnliknande AI-avатарer används också för att utbilda utredare i utredande samtal och förhör med barn. Testerna befinner sig i ett tidigt skede men visar goda resultat. Metoden möjliggör utbildning utan att verkliga barn behöver involveras i utsatta situationer. Eftersom både avataren och utredarna kan ha svårt att uppfatta otydliga signaler från barn, ger en kombination av olika utbildningsmetoder bäst resultat.

Stora språkmodeller används även för att tolka teman i dokumentation från barn- och domstolsförhör. Enligt litteraturoversikten är modellernas tolkningar lika bra eller bättre än de som görs av människor. Ett viktigt resultat från dessa tester är att förhör i domstol är mer styrande än barnförhör, vilket kan få konsekvenser för barnets rätt till skydd och för beviskedjans integritet.

Hälso- och sjukvårdens användning av AI

Inom hälso- och sjukvården visar maskininläring och språkmodeller goda resultat när det gäller att flagga ärenden som en kliniker behöver granska närmare, enligt principen ”human in the loop”.

Verktygen kan uppmärksamma vanvård med högre känslighet än traditionella diagnoskoder och används bland annat för att identifiera barn som riskerar att utsättas för människohandel för sexuella ändamål. Samma teknik kan även användas inom socialtjänsten för att tidigare upptäcka riskfaktorer för utsatthet. För att AI ska ge goda resultat krävs dock tillförlitlig och

representativ data. Det är dessutom avgörande att de ärenden som lyfts fram alltid granskas av en människa med rätt kompetens.

Åldersestimering och ansiktsigenkänning

AI-verktyg som uppskattar åldern på personer som förekommer i övergreppsmaterial kan effektivisera hanteringen av stora mängder material, vilket också minskar utredares och granskares exponering för materialet. Bildbaserade modeller för åldersestimering visar bäst resultat när de ställs inför två val, till exempel att avgöra om personen på bilden är över eller under 18 år.

Modellernas träffsäkerhet påverkas av den data de tränas på. Därför är det viktigt att datan som modellen tränats på är av tillräckligt hög kvalitet, vilket bland annat innebär att den är balanserad, alltså representativ, och inkluderar barn av olika kön, ålder, utveckling och etnicitet.

Litteraturöversikten lyfter även andra verktyg som kan stärka säkerheten för barn och unga i digitala miljöer. Det handlar till exempel om appar som är särskilt framtagna för att skydda barn, åldersklassificering av tjänster och olika former av innehållsfilter. Samtidigt visar litteraturöversikten att tekniska lösningar behöver kombineras med innehållsmoderering, blockeringsfunktioner, hänvisningar till stödverksamheter och utbyte av bästa praxis för att bli effektiva.

Plattformer

Plattformars ansvarstagande identifieras genomgående som avgörande för att brottsutredningar ska kunna genomföras snabbt och effektivt. Det handlar

Modellernas träffsäkerhet påverkas av den data de tränas på. Därför är det viktigt att datan som modellen tränats på är av tillräckligt hög kvalitet, vilket bland annat innebär att den är balanserad, alltså representativ, och inkluderar barn av olika kön, ålder, utveckling och etnicitet.

också om lagstadgade granskningsprocesser, samarbete och gemensamma insatser mellan olika aktörer.

Litteraturöversikten visar att dagens regelverk inte är tillräckligt starka och att flera internationella organisationer vill att barns säkerhet ska vara utgångspunkt när ny teknik tas fram. De efterfrågar därför lagstadgade krav på granskning och gemensamma satsningar mellan olika sektorer för att skyndsamt ta ner övergreppsmaterial.

Summering

- **Möjligheter i praktiken:** Flera AI-verktyg kan idag användas för att stärka barnrättsperspektivet i praktiken. Den tekniska utvecklingen kan bidra till att skapa nya arbetssätt, skydda barn från skadligt innehåll, minska spridningen av övergreppsmaterial samt ge stöd och vägledning till barn som utsatts.
- **Effektiv brottsbekämpning:** Brottsbekämpande myndigheter testar nu AI-teknik för att förebygga och upptäcka brott. Med hjälp av maskininlärning och stora språkmodeller kan man effektivisera arbetsprocesser, genomföra bättre barnförhör och utbilda utredare och hälso- och sjukvårdspersonal.
- **Utmaningar med data:** En stor utmaning är att säkerställa att verktygen tränas på balanserade data som inkluderar olika etniciteter, kön och åldrar. Annars riskerar AI-verktygen ge snedvridna eller inkorrekta svar. Verktygens träffsäkerhet tenderar också att minska när de används utanför kontrollerade träningsmiljöer.
- **Reglering för samtida utmaningar:** AI-genererat sexuellt övergreppsmaterial utmanar nuvarande nationell och europeisk lagstiftning, samt de juridiska definitionerna av vad som ska räknas som "riktigt" övergrepps-material. Juridiken måste värna om både brottsoffers integritet och rättigheter, och samtidigt främja innovativa metoder för att förebygga, upptäcka och rapportera sexuella övergrepp mot barn.

SÅ STÄRKER AI STÖD TILL BARN OCH UNGA SOM UTSATTS FÖR SEXUELLA ÖVERGREPP

AI har redan en roll inom verksamheter som arbetar med stöd till brottsutsatta. Även om användningen av verktygen ökar, är det vetenskapliga stödet för att dessa fungerar än så länge begränsat i jämförelse med verktyg som används för att upptäcka och förebygga brott.

Tester i kontrollerade terapeutiska miljöer visar dock positiva resultat. Bland annat kan AI stötta rehabilitering och bidra till att stärka motståndskraften hos brottsoffer för sexuella övergrepp. Verktygen visar också att de kan förbättra yrkesverksammas självförtroende och samtals teknik i kontakten med barn som är utsatta.

Prediktiva och kliniska tillämpningar

Maskininlärning som tränas på stora mängder patientdata används för att förutse suicidtankar och suicidalt beteende hos unga efter sexuella övergrepp. Tekniken kan identifiera riskfaktorer i olika faser efter utsattheten, även om modellerna har svårt att skilja mellan faktiska suicidförsök och planering. Maskininlärning används även för att förutse posttraumatiskt stressyndrom, PTSD, och utveckling av depression bland barn efter sexuella övergrepp.

Genom så kallad deepfake-teknik får överlevare dessutom möjlighet att konfrontera en digital, simulerad version, en avatar av förövaren i kontrollerade och trygga miljöer. Efter dessa konfrontationer visar de minskade symtom på PTSD samt minskad skuld och skam. Resultaten visar att AI kan användas i terapiesammanhang för att underlätta brottsoffers rehabilitering och bygga upp deras motståndskraft.

Fortbildning för yrkesverksamma och utredande samtal

AI-drivna chatbotar och avatarer används för att träna personal inom tandvård, sjukvård, socialtjänst och civilsamhälle i samtals teknik och i att hålla svåra samtal med barn. Verktygen erbjuder hög visuell, auditiv och känslomässig realism. De stärker personalens självförtroende när de samlar in känslig information från barn som är utsatta för vanvård. Fortbildningen ger också personalen mer kunskap om att ställa öppna och icke-ledande frågor. Samma teknik används även för att förbereda socialtjänsten inför utredande samtal, där botarna dessutom bidrar med direkt återkoppling och förslag på förbättringar av utredarnas samtals teknik och bemötande.

“ **Tester i kontrollerade terapeutiska miljöer visar positiva resultat. Bland annat kan AI stötta rehabilitering och bidra till att stärka motståndskraften hos brottsoffer för sexuella övergrepp.**

Direkt stöd till brottsutsatta

Flera ideella verksamheter använder idag AI-verktyg, som chatbotar, för att möta stödsökande. Det kan handla om stöd vid anmälan, hänvisningar till terapi eller hjälp med att planera för den egna säkerheten. Chatbotar kan även träna unga i skyddande strategier mot grooming samt ge information om sexuell och reproduktiv hälsa.

Denna användning visar på det stora behovet av tillgängligt och tillförlitligt lågröskelstöd. Det finns även verksamheter som praktiskt hjälper barn och unga att ta ner sexuellt kränkande övergreppsmaterial med hjälp av så kallade hashlistor.

Summering

- **Kompetensutveckling för yrkesverksamma:** Det finns goda förutsättningar för att använda AI för att utbilda och stärka kompetensen, samtalstekniken och självförtroendet hos de yrkesverksamma som möter barn som blivit utsatta för övergrepp.
- **Identifiering av riskfaktorer:** Genom kliniska tillämpningar av maskininlärning går det att identifiera kritiska riskfaktorer – såsom depression och suicidtankar – i olika faser efter att ett barn har utsatts för sexuella övergrepp.
- **Tillgängligt direktstöd:** AI-verktyg kan användas för att ge barn enklare information om sexuella och reproduktiva rättigheter, tips på var man kan få stöd, samt få hjälp med nedtagning av sexuellt kränkande material.
- **Terapeutisk rehabilitering:** Inom kontrollerade terapeutiska forskningsmiljöer har AI-verktyg visat goda resultat som komplement i vården för att stödja brottsoffers rehabilitering och bygga upp deras motståndskraft.

UNGAS REFLEKTIONER OCH REKOMMENDATIONER

Med utgångspunkt i litteraturoversiktens resultat samlade Childhood och ChildX unga i två fokusgrupper för att diskutera hur AI kan missbrukas för att begå sexuella övergrepp mot barn, hur tekniken kan förebygga utsatthet och hur AI-verktyg kan användas för att upptäcka övergrepp. Deltagarna bidrog också med egna perspektiv och förbättringsförslag. Nedan följer en sammanställning av de viktigaste teman som diskussionerna bidrog till.

AI i vardagen

“Vi använder AI typ överallt, i mobilen. Inom skolor för att få hjälp. Och i reklam. Och den lär sig mer av människor, ju mer man använder AI, desto mer lär den sig om sättet vi pratar eller sättet vi skriver på, alltså hur vi tänker. Det beror på vad man pratar med AI om så tar den information, alltså den lagrar information på allt vi säger till den, så att den ska improve.”

Ungdom, fokusgrupp 2

I fokusgrupperna blir det tydligt att AI redan är en integrerad del av ungas vardag. Flera deltagare beskriver att de använder chatbotar för att ställa enklare frågor om skoluppgifter eller sitt mående, och att de generellt har positiva erfarenheter av detta.

Deltagarna lyfter att AI lär sig av hur de pratar och skriver, och att tekniken ständigt utvecklas. Samtidigt betonar de att det är viktigt att vara uppmärksam på att verktygen ofta vill bekräfta eller ”pleasa” användaren. De påpekar också att en chatbot inte alltid kan skilja mellan en användare som mår genuint dåligt och en användare som främst behöver uppmuntran.

Sexuellt kränkande manipulerade bilder

Flera deltagare lyfter hur AI bidrar till att sexuellt kränkande och manipulerade bilder normaliseras i ungas relationer. Det kan handla om allt från att skicka manipulerade bilder som ett skämt till sin partner, till att sprida dem på sociala medier:

“Människor blir inte ens chockade längre över det. De har blivit vana med det. De har blivit vana med att den här personen sprider falska bilder på en.”

Ungdom, fokusgrupp 1

Åldersverifiering och ansiktsigenkänning

Litteraturoversikten lyfter flera säkerhetsåtgärder som kan minska ungas exponering för sexualiserat

material. Det handlar till exempel om åldersverifiering av användare, begränsningar som gör att barn bara kan kontakta andra i sin egen ålder och moderering av chattar. Även deltagarna i fokusgrupperna lyfter dessa åtgärder som viktiga. Samtidigt nyanserar deltagarna bilden. De påpekar att vuxna förövare lätt kan kringgå systemet genom att använda bilder på syskon eller andra unga personer vid verifieringen. Som en möjlig lösning föreslår de att vuxna som köper mobiltelefoner ska behöva genomgå en användarverifiering som är förinstallerad i mobilen före köpet. Det skulle kunna vara en säkerhetsåtgärd som minskar barns utsatthet.

“Ja, man borde göra så när man köper telefoner att man verifierar sig då när man köper den, för man kan ju se vem som har köpt mobilen.”

Ungdom, fokusgrupp 2

När det gäller åldersverifiering betonar deltagarna att funktionerna inte får hindra kommunikationen mellan unga. De understryker också att systemet inte ska gå att lura med någon annans foto, utan att identiteten i stället bör verifieras mot en id-handling. Deltagarna lyfter även moderering av chattar som en viktig säkerhetsåtgärd, eftersom det kan avskräcka förövare från att kontakta barn.

“Ja, de (moderatorerna) borde vara mer aktiva när det gäller barn under 18, att de tittar mer på deras chattar då och då, att när man skriver till någon att båda ändå vet i bakhuvudet att okej, vad än jag skriver här kan vuxna från den här appen se. Så då skapar det lite så att de kanske de som försöker leta efter barn kanske tänker till: "amen om jag skriver nånting så kan det finnas folk som ser det vi skriver så de kanske backar och inte vet du vad väljer att göra sånt.”

Ungdom, fokusgrupp 2

Alltså, chat-botten borde veta sina gränser

Ungdom, fokusgrupp 2

Deltagarna betonar dock att modereringen behöver skötas av människor, eftersom de inte anser att AI-moderering är tillräckligt tillförlitlig. De vill också att experter som utvecklar AI-moderering förklarar hur de tänker lösa problemet med att AI inte alltid går att lita på som moderator.

“Alltså jag vill bara lägga till att ni borde så här fråga experter på området om vad de tycker om vad de ska göra åt att modereringen inte funkar. För att man ser jätteofta att det funkar inte att moderera genom AI, det funkar inte att relya på AI så mycket utan det känns som att ofta företag bara relyar på en AI och de tittar inte igenom liksom: "okej, har det modererat bra? Har de bannat rätt person? Eller så här har man genererat rätt ålder till rätt barn för att inte en fyraåring hamnar med nittonåringar?”

Ungdom, fokusgrupp 2

Samtidigt ser deltagarna möjligheter att använda AI som stöd för plattformarna. AI skulle till exempel kunna bidra till att känna igen vuxna förövare utifrån deras kommunikationsmönster när de kontaktar barn. Barnen lyfter också en annan säkerhetsåtgärd: att användaren verifierar sin identitet gentemot företaget, samtidigt som användarnamnet fortfarande kan vara anonymt för andra användare. En sådan lösning skulle också kunna göra det lättare för företag att anmäla och blockera förövare.

Direkt stöd till brottsoffer

Användningen av AI för direkt stöd till barn och unga är det tema som engagerar flest deltagare. En deltagare lyfter anmälningsplikten som ett hinder för att kontakta traditionella stödverksamheter. Det kan vara en anledning till att unga i stället vänder sig till en chatbot. Till skillnad från en kurator inom exempelvis elevhälsan eller socialtjänsten har en chatbot ingen skyldighet att göra en orosanmälan. Detta beskrivs som särskilt viktigt i situationer där barnet främst vill få stöd för att må bättre i stunden, utan att vuxna ingriper eller gör situationen större än vad barnet själv önskar. Samtidigt lyfter ungdomarna flera hinder med att söka stöd hos en chatbot. Deltagarna påpekar att en chatbot kan hålla med användaren, även när användaren uttrycker en vilja att skada sig själv eller andra. Ett barn hänvisar också till uppgifter på sociala medier om barn som tagit sitt liv efter kontakt med en chatbot som uppmuntrat till suicid.

“För alltså Chat GPT är sånt, det är så här specifik kodat för att hålla med dig om det. Så om du tänker så här, ja men jag har inget liv, jag har ingen livslust längre. Jag ser ingen mening med livet, så hetsar de upp det och bara så här, ja men det hade du rätt i. Kolla på alla krig som händer i världen och sånt.”

Ungdom, fokusgrupp 1

Deltagarna lyfter även att olika chatbotar kan ha olika funktioner beroende på barnets behov. Vissa botar kan vara användbara för att få länkar eller information om hur man söker vidare stöd när man mår dåligt. Vid allvarliga problem betonar unga dock att man inte ska förlita sig på en bot, utan ta hjälp av en vuxen som man litar på och som kan agera.

“ Lagarna är gamla och internet går så här jättefort. De måste göra det straffbart att ens försöka lura barn med AI.

Ungdom, fokusgrupp 2

Deltagarna i en av fokusgrupperna påpekar också att avsändaren påverkar viljan att söka stöd. Om en chatbot kommer från en verksamhet med experter och ett tydligt uppdrag att hjälpa unga ökar tilliten till just den boten. Verktuget kan då fungera som ett första steg, fram till dess att barnet känner sig tryggt nog att ta emot hjälp från en människa.

“Men jag vet att det finns folk som använder ChatGPT som terapeuter som de sa. Men jag tycker att det är ganska värdelöst att gå dit för den säger inte farliga saker, för oftast om jag (...) säger till den till exempel att jag inte gillar skolan så börjar den skicka hjälplinje-nummer. Det är därför jag tycker att använda ChatGPT är ganska värdelöst för den skickar bara länkar till olika typer av nummer som man kan ringa stödlinjer för saker som egentligen inte är så deep. Så jag tycker att man borde bara använda ChatGPT för skoluppgifter och små saker för det hjälper inte.”

Ungdom, fokusgrupp 2

Juridiska utmaningar

Med lika delar frustration och uppgivenhet beskriver deltagarna vuxenvärldens oförmåga att skydda dem mot sexuella övergrepp och exploatering. En deltagare säger att det borde finnas någon lag som uttryckligen förbjuder AI-genererade nakenbilder, särskilt när det finns så många lagar. En annan deltagare uttrycker stor oförståelse för att samhället inte lyckas hålla jämna steg med den snabba teknikutvecklingen.

“Lagarna är gamla och internet går så här jättefort. De måste göra det straffbart att ens försöka lura barn med AI.”

Ungdom, fokusgrupp 2

Deltagarna i fokusgrupperna uttrycker även en frustration över att personer som redan har anmälts för att ha försökt begå eller har begått brott mot barn, tillåts att skapa nya användarkonton och kontakta fler barn, trots att de borde vara avstängda från plattformarna. De menar att AI skulle kunna användas för att hjälpa polisen att skanna plattformar och identifiera personer som tidigare har anmälts eller stängts av.

Plattformar

Plattformarnas ansvar är ett centralt tema i fokusgrupperna. Deltagarna lyfter att plattformar behöver skydda barn mot både grooming, spridning av AI-genererat övergreppsmaterial och sexuellt kränkande manipulerade bilder. Många beskriver en tydlig dubbelmoral och menar att plattformarnas egna användarregler inte verkar gälla företagen själva. De ifrågasätter hur algoritmer kan tillåtas föreslå skadligt material till barn, när det så tydligt strider mot plattformarnas egna riktlinjer.

Framtiden

Det råder delade meningar i fokusgrupperna om hur AI påverkar framtiden. Vissa tror att utvecklingen kan bli bättre när de själva blir vuxna, eftersom de har vuxit upp med tekniken och förstår den bättre. Andra uttrycker oro över vart utvecklingen är på väg. En deltagare sätter ord på denna dubbelhet: det är coolt med tekniken men man märker att människorna som styr den inte hinner med.

Summering

- **Aktiv och positiv användning:** Barn och unga är idag aktiva användare av AI inom flera olika områden. De ser positivt på att tekniken används som hjälpmedel i skolarbetet, och för att hitta information om till exempel var man kan söka stöd.
- **Teknikens dualitet:** Samtidigt uppmärksammar barnen det problematiska i att vuxenvärlden inte lyckas hänga med i den snabba teknikutvecklingen, eftersom det påverkar deras säkerhet.
- **Konsekvenser i barnens vardag:** Konsekvenserna som unga tvingas hantera i vardagen inkluderar risken att exponeras för övergreppsmaterial, utsätts för spridning av sexuellt kränkande, manipulerat material samt bristande moderering av olämpliga inlägg.

BARNRÄTTSLIG SLUTSATS OCH VÄGEN FRAMÅT

Den här rapporten visar något som inte längre kan ignoreras, AI förändrar förutsättningarna för sexuella övergrepp mot barn och för arbetet med att förebygga och stoppa dem. Frågan nu är inte hur samhället ska förhålla sig till denna utveckling. Frågan är om samhället agerar tillräckligt snabbt, ansvarsfullt och med barnets rättigheter i centrum.

Genom hela rapporten framträder AI som en teknik med dubbla konsekvenser. Samma verktyg som förövare använder för att skapa övergreppsmaterial, automatisera grooming och manipulera bilder, används också för att upptäcka brott, identifiera offer och stärka yrkesverksamma.

Att avstå från att använda AI i arbetet mot sexuella övergrepp mot barn är inte en neutral position. Det är ett aktivt val med verkliga konsekvenser för riktiga barn.

Barnen ser det vi missar

Ett av rapportens viktigaste bidrag är barnens egna röster, kloka och frustrerade, de talar av erfarenhet. De ser hur vuxenvärlden, lagstiftningen och plattformarna inte hänger med. De har konkreta förslag: bättre åldersverifiering, mänsklig moderering, uppdaterade lagar och chatbotar från betrodda avsändare med ett tydligt uppdrag att hjälpa. För att hantera de risker och brister som barnen så tydligt beskriver måste vi vuxna överbrygga tre centrala glapp som identifieras i den här rapporten.

“ Det är coolt med tekniken, men man märker att människorna som styr den inte hinner med.

Ungdom, fokusgrupp 2

Tre centrala glapp

Teknikutveckling och lagstiftning: AI-genererat övergreppsmaterial utmanar nuvarande juridiska definitioner och den gråzon som finns utnyttjas av förövare. Juridiska och organisatoriska hinder sätter stopp för datadelning som krävs för att effektivisera arbetet med att hitta förövare samt identifiera barn i utsatthet. Lagstiftningen måste vara teknikneutral och skydda barn oavsett vilket verktyg förövaren använder. Redan i utvecklingsfasen av ny teknik måste utgångspunkten vara barns rättigheter.

Tillgängliga AI-verktyg och faktisk användning: Verktyg för att analysera bilder, kartlägga beteenden, uppskatta ålder och identifiera brottsoffer finns redan och används i andra länder. För att de inte ska bli missvisande krävs tillgång till data som är balanserad och representativ. Men brottsbekämpande myndigheter i Sverige saknar idag de rättsliga, tekniska och finansiella förutsättningarna som krävs för att använda dem i praktiken.

Plattformarnas ansvar och faktiska agerande:

Barnen i fokusgrupperna beskriver en tydlig dubbelmoral: att plattformar, trots att det strider mot deras riktlinjer, kan tillåta algoritmer som rekommenderar olämpligt eller skadligt innehåll till barn. Det räcker inte att uppmana plattformarna till transparens, skyndsam hantering och samverkan med myndigheter och civilsamhälle, det måste vara lagstadgade krav. Det tillfälliga undantaget från ePrivacy-direktivet som löpte ut i april 2026 gör detta ännu mer akut, utan lagstadgad grund riskerar plattformarnas möjlighet att proaktivt skanna efter övergreppsmaterial att försämrans dramatiskt.

Vad vi vet och vad vi ännu inte vet

Rapporten är tydlig med att evidensläget varierar. Inom brottsbekämpning och upptäckt av övergreppsmaterial är resultaten starka. Inom stöd och terapeutisk rehabilitering är forskningen mer begränsad, och många verktyg befinner sig fortfarande i testmiljöer. Det är viktigt att vara medveten om denna distinktion, inte för att bromsa användningen, utan för att säkerställa att verktygen utvärderas löpande utifrån barnets bästa och att en människa alltid finns i beslutsledet. Mer forskning behövs. Men behovet av mer forskning får inte bli en ursäkt för att inte agera här och nu.

Vägen framåt

AI kommer att fortsätta utvecklas. Förövare kommer att fortsätta att utveckla och utnyttja ny teknik. Det som avgör om skyddet för barn blir bättre eller sämre är inte tekniken i sig, det är hur samhället väljer att använda den. Det måste ske på ett ansvarsfullt sätt. Fler måste våga investera långsiktigt, samverkan mellan olika sektorer måste stärkas och ett barnrättsligt perspektiv måste integreras i hela ansvarskedjan.

“ Sverige har ambitionen att bli världsledande inom AI. Den ambitionen får inte utesluta skyddet av barn.

Sverige har ambitionen att bli världsledande inom AI. Den ambitionen får inte utesluta skyddet av barn. Deras rätt till skydd, integritet, stöd och delaktighet gäller även i digitala miljöer. Rätten är teknikneutral och den gäller redan idag. Det betyder att lagstiftning, finansiering och samverkan mellan myndigheter, civilsamhälle, forskning och techsektor måste hänga med. Inte sen. Inte om några år. Nu.

REKOMMENDATIONER

Vårt gemensamma mål är att rekommendationerna ska vara ett konkret underlag för att intensifiera användningen av AI i arbetet med att upptäcka, förebygga och stoppa sexuella övergrepp och exploatering av barn och för att motverka att AI används som ett verktyg av förövare.

1. Modernisera lagstiftningen för att skydda barn mot sexuella kränkningar

Målgrupp: Politiska beslutsfattare

1.1 Säkerställ att nuvarande lagstiftning är teknikneutral och skyddar barn mot övergrepp som faciliteras av AI: Tapp till lagstiftningens luckor så att användning av AI för att skapa sexuellt kränkande material mot barn blir straffbart, och så att manipulering av övergreppsmaterial med hjälp av AI i syfte att kringgå lagen inte längre är möjligt. Lagstiftningen måste säkerställa ett likvärdigt skydd för alla barn, oavsett ålder, kön, etnicitet eller pubertetsutveckling.

1.2 Förbjud skadliga tjänster: Vidta åtgärder mot specifika funktioner med hög skaderisk för barn, såsom tjänster som möjliggör sexuell kränkande manipulationer ('nudify-funktioner').

1.3 Harmonisera definitioner och processer internationellt: Driv processer på EU-nivå och mellan aktörer för mer likriktade juridiska definitioner av övergrepp och övergreppsmaterial, samt mer synkade processer för nedtagning av material.

2. Investera i teknik och samverkan som skyddar barn

Målgrupp: Beslutsfattare, myndigheter, finansärer och forskare

2.1 Säkra förutsättningar för datadelning: Se över juridiska och organisatoriska hinder för att tillåta säkert informationsutbyte mellan sektorer.

2.2 Etablera kontrollerade testmiljöer: Skapa kontrollerade testmiljöer, till exempel tekniska och regulatoriska sandlådor, där nya verktyg och AI-modeller kan utvärderas innan de når marknaden.

2.3 Tillsätt särskilda utvecklingsmedel för AI-projekt inom brottsbekämpning: Identifiera de områden där AI kan ge störst effekt inom brottsbekämpning. Anpassa tekniker som redan används internationellt till svenska förhållanden. Varje projekt bör utgå från en konsekvensanalys med utgångspunkt i barnets rätt till skydd, stöd och delaktighet och måste följas av systematisk utvärdering för att säkerställa faktisk effekt i form av identifierade barn och lagförda förövare.

2.4 Systematisera barnrättsprinciper och villkora stöd: Ge i uppdrag att systematiskt inkludera barnrättsprinciper i tech-utveckling, med oberoende granskningar och standardiserade mätetal för att säkerställa att AI-verktyg inte kan missbrukas för sexuella övergrepp mot barn. Villkora statliga stöd och offentliga upphandlingar mot att AI-systemen lever upp till att vara trygga för barn.

2.5 Skapa förutsättningar för en hållbar infrastruktur: Gå från kortsiktiga pilotprojekt till långsiktig finansiering samt säkerställ stabil drift, underhåll och kontinuerlig träning av AI-stöd i kärnverksamheten.

2.6 Säkerställ att användare enkelt kan anmäla online: Skapa enkla, säkra och tillgängliga sätt att anmäla sexuella kränkningar och övergrepp online.

3. Kräv säkerhet från start och ansvarsfull moderering

Målgrupp: Techsektorn och plattformar

3.1 Integrera säkerhet från start: Säkerställ att säkerhetsdesign utgår från barns rättigheter.

3.2 Tydliggör ansvaret i hela kedjan:

- **Infrastrukturleverantörer och datacenter:** Implementera kontrollmekanismer som säkerställer att lagringstjänster för träning av storskaliga modeller inte innehåller illegalt material, såsom övergreppsmaterial mot barn.
- **Ägare av onlineplattformar och utvecklare av AI-verktyg:** Arbeta proaktivt med att hindra att övergreppsmaterial laddas upp på plattformar, sök kontinuerligt efter olagligt innehåll, blockera försök till grooming och samverka över sektorsgränser. Skapa enkla vägar för rapportering och stöd för barn.
- **Utvecklare av generativa AI-modeller och verktyg:** Verka för att systemen inte kan generera skadligt material, vare sig genom förekomst i träningsdata eller till följd av otillräckliga tekniska skyddsåtgärder.

3.3 Garantera mänsklig översyn i modereringen:

Säkerställ att mänskliga moderatorer fattar beslut i komplexa ärenden och tillhandahåll transparenta vägar för att överklaga automatiserade beslut.

4. Rusta yrkesverksamma och säkra digitala stödstrukturer

Målgrupp: Yrkesverksamma och civilsamhälle

4.1 Genomför ett nationellt kompetenslyft: Erbjud kontinuerlig utbildning till personal som möter barn om barns användning av AI, AI-relaterade hot, juridiska ramverk och nya användningsområden för tekniken.

4.2 Systematisera frågor om digital utsatthet:

Integrera frågor om utsatthet för våld, sexuella övergrepp online och AI-relaterade skaderisker i utredande och stödjande samtal med barn.

4.3 Utveckla säkra AI-stödlinjer: Utveckla och utvärdera lågtröskelstöd, såsom AI-chattjänster som komplement när mänskligt stöd inte är tillgängligt och som brygga till professionellt stöd.

4.4 Introducera AI-simulering för

färdighetsträning: Implementera och utvärdera AI-avatare och AI-verktyg för att träna yrkesverksamma i förhørs- och samtalsmetodik utan att involvera barn i utsatta situationer.

KÄLLFÖRTECKNING

Andersson, J., & Wranger, T. (2025). *Barnen och internet 2025*. Internetstiftelsen. Hämtad 2026-06-08 från: <https://svenskarnaochinternet.se/app/uploads/2025/11/internetstiftelsen-barnen-och-internet-2025.pdf>

Brottsbalken (1962:700). Hämtad 2026-06-08 från: https://www.riksdagen.se/sv/dokument-och-lagar/dokument/svensk-forfattningssamling/brottsbalk-1962700_sfs-1962-700/

Ecpat Sverige. (2026). "AI eller inte – det är en invasion av någons privatliv, eller deras kropp": En rapport om barns syn på AI och nudify-verktyg. Hämtad 2026-06-08 från: https://ecpat.se/wpcontent/uploads/2026/05/AI-RAPPORT_2.pdf

Ecpat Sverige (2023). Jag tittade bara, jag gjorde inget elakt. En rapport om barns utsatthet på exposekonton. Hämtad 2026-06-08 från: https://ecpat.se/wp-content/uploads/2020/12/Ecpat_Exposerapport_2023_uppslag_low.pdf

FN:s kommitté för barnets rättigheter. (2021). *Allmän kommentar nr 25 om barnets rättigheter i relation till den digitala miljön (CRC/C/GC/25)*. OHCHR. Hämtad 2026-06-08 från: <https://www.ohchr.org/en/documents/general-comments-and-recommendations/general-comment-no-25-2021-childrens-rights-relation>

Förenta nationernas konvention om barnets rättigheter, 20 november 1989, United Nations Treaty Series, vol. 1577, s. 3. Hämtad 2026-06-08 från: https://treaties.un.org/pages/viewdetails.aspx?src=treaty&mtdsg_no=iv-11&chapter=4&clang=_en

Europaparlamentets och rådets förordning (EU) 2022/2065 av den 19 oktober 2022 om en inre marknad för digitala tjänster och om ändring av direktiv 2000/31/EG (Digital Services Act). EUT L 277, 27.10.2022, s. 1–102.

Europaparlamentets och rådets direktiv 2011/93/EU av den 13 december 2011 om bekämpande av sexuella övergrepp mot barn, sexuell exploatering av barn och barnpornografi samt om ersättande av rådets rambeslut 2004/68/RIF. EUT L 335, 17.12.2011, s. 1–14.

Europaparlamentets och rådets förordning (EU) 2024/1689 av den 13 juni 2024 om harmoniserade regler för artificiell intelligens och om ändring av förordningarna (EG) nr 300/2008, (EU) nr 167/2013, (EU) nr 168/2013, (EU) 2018/858, (EU) 2018/1139 och (EU) 2019/2144 samt direktiven 2014/90/EU, (EU) 2016/797 och (EU) 2020/1828 (EU AI Act). EUT L, 2024/1689, 12.7.2024.

Europaparlamentet (2026). *AI Act: deal on simplification measures, ban on nudifier apps*. Hämtad 2026-06-08 från: <https://www.europarl.europa.eu/news/en/pressroom/20260427IPR42011/ai-act-deal-on-simplification-measures-ban-on-nudifier-apps>.

McMahan, A. (2025). A rapid review of artificial intelligence in the facilitation, detection, and prevention of child sexual abuse and exploitation (Content area expert reviewer: L. Adomaitis). Karolinska Institutet & Stockholm Health Care Services, Region Stockholm.

CHILDHOOD.SE

CHILDX.SE