



Europeiska
unionens råd

Bryssel den 19 december 2023
(OR. en)

17002/23

JAI 1704
ENFOPOL 554
CRIMORG 207
IXIM 247
DATAPROTECT 384
CYBER 326
COPEN 461
FREMP 382
TELECOM 403
COMPET 1303
MI 1154
CONSUM 504
DIGIT 303

FÖLJENOT

från: Europeiska kommissionens generalsekreterare, undertecknat av
Martine DEPREZ, direktör

inkom den: 19 december 2023

till: Thérèse BLANCHET, generalsekreterare för Europeiska unionens råd

Komm. dok. nr: COM(2023) 797 final

Ärende: RAPPORT FRÅN KOMMISSIONEN TILL EUROPAPARLAMENTET
OCH RÅDET
om genomförandet av förordning (EU) 2021/1232 av den 14 juli 2021
om ett tillfälligt undantag från vissa bestämmelser i direktiv 2002/58/EG
vad gäller användning av teknik hos tillhandahållare av
nummeroberoende interpersonella kommunikationstjänster för
behandling av personuppgifter och andra uppgifter i syfte att bekämpa
sexuella övergrepp mot barn på nätet

För delegationerna bifogas dokument – COM(2023) 797 final.

Bilaga: COM(2023) 797 final

Bryssel den 19.12.2023
COM(2023) 797 final

RAPPORT FRÅN KOMMISSIONEN TILL EUROPAPARLAMENTET OCH RÅDET

om genomförandet av förordning (EU) 2021/1232 av den 14 juli 2021 om ett tillfälligt undantag från vissa bestämmelser i direktiv 2002/58/EG vad gäller användning av teknik hos tillhandahållare av nummeroberoende interpersonella kommunikationstjänster för behandling av personuppgifter och andra uppgifter i syfte att bekämpa sexuella övergrepp mot barn på nätet

INNEHÅLL

FÖRTECKNING ÖVER TERMER OCH AKRONYMER	2
1. INLEDNING	4
2. GENOMFÖRANDEÅTGÄRDER	5
2.1. Tillhandahållarnas behandling av personuppgifter (artikel 3 g vii)	5
2.1.1. Typ och volym av de behandlade uppgifterna	5
2.1.2. Skäl för behandling enligt förordning (EU) 2016/679	6
2.1.3. Skäl för överföring av personuppgifter utanför unionen enligt kapitel V i dataskyddsförordningen, där så är tillämpligt	6
2.1.4. Antal fall av sexuella övergrepp mot barn på nätet som identifierats, uppdelat mellan material med sexuella övergrepp mot barn och kontaktsökning med barn	7
2.1.5. Prövning för användare och utfall	8
2.1.6. Antal och andel fel (falska positiva resultat) för de olika tekniker som användes	10
2.1.7. Åtgärder som vidtagits för att begränsa felfrekvensen och den felprocent som uppnåtts ...	13
2.1.8. Policyn för lagring samt de dataskyddsåtgärder som tillämpas enligt dataskyddsförordningen	14
2.1.9. Namn på organisationer som agerar i allmänhetens intresse mot sexuella övergrepp mot barn och med vilka uppgifter har delats enligt denna förordning	15
2.2. Statistik från medlemsstaterna (artikel 8)	15
2.2.1. Totalt antal rapporter om upptäckta sexuella övergrepp mot barn på nätet	16
2.2.2. Antalet barn som identifierats	23
2.2.3. Antalet dömda förövare	26
2.3. Utvecklingen av tekniska framsteg	28
2.3.1. Spårning av känt material med sexuella övergrepp mot barn	29
2.3.2. Spårning av nytt material med sexuella övergrepp mot barn	29
2.3.3. Spårning av grooming	30
2.3.4. Nya utmaningar i form av dialogrobotar och konst/bild-generatorer med artificiell intelligens	
31	
3. SLUTSATSER	32

FÖRTECKNING ÖVER TERMER OCH AKRONYMER

Term/akronym	Definition
AI	Artificiell intelligens
API	Gränssnitt för tillämpningsprogram
CG-CSAM	Computer-generated Child Sexual Abuse Material (datorgenererat material med sexuella övergrepp mot barn)
ChatGPT	ChatGPT (Chat Generative Pre-trained Transformer) är en form av generativ AI. Det är en stor språkmodellbaserad dialogrobot, utvecklad av OpenAI, som ger användarna möjlighet att finjustera och styra en konversation mot önskad längd, format, stil, detaljnivå och språk.
Klassificerare	En form av artificiell intelligens, en algoritm som sorterar data i angivna klasser eller kategorier
Klassificeraren Content Safety API	Googles klassificerare Content Safety API använder programmatisk åtkomst och artificiell intelligens för att klassificera och prioritera miljardtals bilder för granskning
CSA	Child Sexual Abuse (sexuella övergrepp mot barn)
CSAI Match	CSAI Match är en teknik som utvecklats av YouTube-tekniker för att upptäcka nya uppladdningar av tidigare identifierade sexuella övergrepp mot barn i videor
CSAM	Child Sexual Abuse Material (material med sexuella övergrepp mot barn), t.ex. bilder och videor som visar övergrepp
Direktivet om sexuella övergrepp mot barn	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
CSEA	Child Sexual Exploitation and Abuse (sexuell exploatering av och sexuella övergrepp mot barn)
Sexuella övergrepp mot barn på nätet	Den gemensamma term som används för de tre typer av sexuella övergrepp mot barn som definieras i direktivet om sexuella övergrepp mot barn, nämligen barnpornografi, pornografisk föreställning och kontaktsökning med barn ("gromning"), enligt definitionen i interimsförordningen (artikel 2.4)
EU	Europeiska unionen
Dataskyddsförordningen	Europaparlamentets och rådets förordning (EU) 2016/679 av den 27 april 2016 om skydd för fysiska personer med avseende på behandling av personuppgifter och om det fria flödet av sådana uppgifter och om

	upphävande av direktiv 95/46/EG (allmän dataskyddsförordning), EUT L 119, 4.5.2016, s. 1
Gromning	När gärningsmän bygger upp förtroende och ett förhållande med ett barn i ett försök att utsätta barnet för sexuell exploatering eller sexuella övergrepp. Formellt känt som kontaktsökning med barn, enligt definitionen i artikel 6 i direktivet om sexuella övergrepp mot barn.
Hash	En unik digital kod som skapas av en matematisk algoritm ("hashing") och som blir filens signatur, eller dess hashvärde
Interimsförordningen	Europaparlamentets och rådets förordning (EU) 2021/1232 av den 14 juli 2021 om ett tillfälligt undantag från vissa bestämmelser i direktiv 2002/58/EG vad gäller användning av teknik hos tillhandahållare av nummeroberoende interpersonella kommunikationstjänster för behandling av personuppgifter och andra uppgifter i syfte att bekämpa sexuella övergrepp mot barn på nätet, EUT L 274, 30.7.2021, s. 41
LLM	En stor språkmodell (Large Language Model) är en typ av artificiell intelligensmodell som har tränats genom djupinlärningsalgoritmer för att känna igen, generera, översätta och/eller sammanfatta stora mängder skrivet mänskligt språk och textdata.
MD5	Kryptografisk algoritm för autentiseringskod för meddelanden som används på internet
Meta SSN++	En AI-modell utvecklad av Meta som kan upptäcka nästan exakta kopior.
NCMEC	National Centre for Missing and Exploited Children (amerikansk privat, icke-vinstdrivande organisation) till vilken tillhandahållare av onlinetjänster enligt amerikansk lag är skyldiga att rapportera fall av potentiella sexuella övergrepp mot barn som upptäcks i deras nätverk
PDQ och TMK+PDQF	Verktyg som Facebook använder för att upptäcka skadligt innehåll. PDQ är en teknik för fotomatchning och TMK+PDQF är en teknik för videomatchning.
PhotoDNA	Det mest använda verktyget baserat på hashingteknik som finns tillgängligt kostnadsfritt och baserar sig på ett licensavtal som är skraddarsytt för att förhindra missbruk och användning i något annat syfte än att upptäcka sexuella övergrepp mot barn

1. INLEDNING

Enligt artikel 9 i interimsförordningen (nedan även kallad *förordningen*) ska kommissionen utarbeta en rapport om dess genomförande på grundval av rapporter från tillhandahållare av interpersonella kommunikationstjänster (*tillhandahållare*) om behandlingen av personuppgifter och om den statistik som medlemsstaterna lämnat. Enligt den ovannämnda bestämmelsen ska kommissionen i genomföranderapporten särskilt beakta

- a) de villkor för behandling av personuppgifter och andra uppgifter som behandlas enligt förordningen,
- b) proportionaliteten i det undantag som föreskrivs i förordningen, inbegripet en bedömning av den statistik som medlemsstaterna lämnat enligt artikel 8,
- c) utvecklingen av tekniska framsteg när det gäller verksamhet som omfattas av förordningen, och i vilken utsträckning denna utveckling förbättrar exaktheten och minskar antalet och andelen fel (falska positiva resultat).

Denna genomföranderapport enligt interimsförordningen bygger på de uppgifter som har lämnats av tillhandahållare och medlemsstater enligt artikel 3.1 g vii respektive artikel 8. Rapporteringen visade att det fanns betydande skillnader i tillgången till uppgifter och vilka typer av uppgifter som samlats in och därmed också i jämförbarheten mellan de uppgifter som samlats in av tillhandahållare och medlemsstater. Eftersom det i förordningen inte finns någon mall för rapportering lämnade tillhandahållarna olika typer av information som inte alltid var jämförbar. Kommissionens avdelningar gjorde därför en uppföljning för att se till att uppgifterna tolkas korrekt. Den övervägande delen av medlemsstaterna kunde inte lämna uppgifter i tid och vissa av dem har inte kunnat lämna några uppgifter alls fram till offentliggörandet av denna rapport. Det hade en betydande inverkan på rapportens tidsplan, fullständighet och användbarhet. Trots ansträngningarna för att se till att uppgifterna är enhetliga och jämförbara finns det fortfarande betydande skillnader, vilket framgår av tabellerna nedan. På alla punkter finns inte heller uppgifter från samtliga tillhandahållare eller medlemsstater.

Syftet med denna genomföranderapport är att utifrån tillgängliga uppgifter ge en faktabaserad översikt över läget när det gäller genomförandet av interimsförordningen. Rapporten innehåller inga tolkningar av förordningen och det tas inte ställning till på vilket sätt den har tolkats och tillämpats i praktiken.

2. GENOMFÖRANDEÅTGÄRDER

2.1. Tillhandahållarnas behandling av personuppgifter (artikel 3 g vii)

I artikel 3 g vii i interimsförordningen fastställs villkoren för tillhandahållare som agerar enligt undantaget i förordningen att senast den 3 februari 2022, och därefter senast den 31 januari varje år, offentliggöra och till den behöriga tillsynsmyndigheten och kommissionen lämna in en rapport om behandlingen av personuppgifter enligt förordningen. Google, LinkedIn, Meta, Microsoft och X (tidigare Twitter)¹ lämnade rapporter för 2021 och 2022.

2.1.1. Typ och volym av de behandlade uppgifterna

Tillhandahållarna rapporterade att de behandlade både innehålls- och trafikdata.

När det gäller innehållsdata som behandlas för att upptäcka sexuella övergrepp mot barn på nätet uppgav samtliga ovannämnda tillhandahållare bilder och videor. De förlitade sig främst på hashmatchningsteknikerna PhotoDNA och MD5 för att upptäcka matchningar av tidigare identifierat material med sexuella övergrepp mot barn. Googles CSAI Match-verktyg användes för att skapa digitala fingeravtryck av videor på plattformar och jämföra dem med filerna i Googles/YouTubes register över fingeravtryck (LinkedIn). I lika stor utsträckning användes enligt rapporterna automatiserad teknik (artificiell intelligens, maskininlärning) och mänsklig granskning (t.ex. av Google). Google och LinkedIn bekräftade att de även identifierat material med sexuella övergrepp mot barn som inte matchade tidigare identifierat material. Ingen av de fem tillhandahållarna som lämnade in uppgifter uppgav att de upptäckt kontaktsökning med barn via textdetektering inom ramen för det undantag som föreskrivs i denna förordning.

När det gäller insamlade trafikdata och de respektive volymerna av olika typer av innehålls- och trafikdata som behandlas fanns det stora variationer i tillhandahållarnas rapporter.

Trafikdata som samlas in av tillhandahållare och ingår i CyberTipline-rapporter till U.S. National Center for Missing and Exploited Children (NCMEC) omfattar (alla eller ett urval av) följande uppgifter:

- a) Uppgifter om användare/uppgiftslämnare/konto (Google, LinkedIn, Microsoft, X).
- b) Metadata relaterade till innehålls-/transaktionsdata (Google, LinkedIn, Microsoft).
- c) Uppgifter som rör ett potentiellt offer (Google).
- d) Uppgifter om missbruk (Google).

När det gäller volymen av de uppgifterna som behandlats enligt interimsförordningen rapporterade LinkedIn behandling av 8 miljoner bilder och videor med ursprung i EU mellan den 14 juli och den 31 december 2021 samt 21 miljoner bilder och 63 000 videor med ursprung i EU under 2022. Microsoft rapporterade en behandling enligt förordningen av 8,9 miljarder bilder och videor globalt mellan juli och december 2021 samt 12,3 miljarder innehållsposter globalt

¹ Twitter skickade in sitt bidrag före namnbytet, men omnämns som X i resten av denna rapport.

under 2022. Några EU-siffror fanns inte tillgängliga, vilket gör att det inte går att dra några slutsatser som är relevanta för denna rapport. De andra tillhandahållarna lämnade ingen information om de behandlade uppgifternas volym. Av de fem tillhandahållare som lämnade in rapporter var det därför bara en som lämnade uppgifter på den detaljnivå som krävdes.

För att illustrera det övergripande sammanhanget rapporterade NCMEC att organisationen under 2022 tog emot totalt 87,2 miljoner bilder och videor från hela världen och 5,1 miljoner bilder och videor från EU. Siffran för 2021 var 84,8 miljoner bilder och videor från hela världen och 1,8 miljoner bilder och videor från EU. Detta är enbart sådant material som har identifierats som potentiellt material med sexuella övergrepp mot barn av en tillhandahållare och kan därför inte ses som en indikation på de totala volymerna av uppgifter som behandlats enligt interimsförordningen.

2.1.2. Skäl för behandling enligt förordning (EU) 2016/679

Tillhandahållarna uppgav följande specifika skäl enligt förordning (EU) 2016/679 (*dataskyddsförordningen*):

- Artikel 6.1 d i dataskyddsförordningen, dvs. behandlingen är nödvändig för att skydda intressen som är av grundläggande betydelse för barn och de som är offer för sexuella övergrepp mot barn på nätet (Google, Meta, X²).
- Artikel 6.1 e i dataskyddsförordningen, dvs. behandlingen är nödvändig för att utföra en uppgift av allmänt intresse (LinkedIn, Microsoft, Meta, X³).
- Artikel 6.1 f i dataskyddsförordningen, dvs. behandling som är nödvändig för ändamål som rör berättigade intressen hos
 - i. tillhandahållaren av att inom ramen för sina tjänster upptäcka, förhindra eller på annat sätt ta itu med sexuella övergrepp mot barn på nätet och att skydda andra användare, kunder, partner och allmänheten från denna form av olagligt innehåll (Google, Meta),
 - ii. offer för sexuella övergrepp mot barn och den organisation till vilken tillhandahållaren rapporterar sexuella övergrepp mot barn på nätet (t.ex. NCMEC) av att upptäcka, förebygga och avlägsna material med sexuella övergrepp mot barn på nätet från tjänsten (Google).

2.1.3. Skäl för överföring av personuppgifter utanför unionen enligt kapitel V i dataskyddsförordningen, där så är tillämpligt

Samtliga tillhandahållare uppgav att de använde standardiserade avtalsklausuler som godkänts av kommissionen i enlighet med artikel 46.2 c i dataskyddsförordningen. För överföringar av personuppgifter till NCMEC rapporterade LinkedIn också att man i den utsträckning det är tillämpligt åberopade undantag tillåtna enligt artikel 49.1 i dataskyddsförordningen.

² Meta och X specificerade inte uttryckligen den konkreta artikeln.

³ Meta och X specificerade inte uttryckligen den konkreta artikeln.

2.1.4. Antal fall av sexuella övergrepp mot barn på nätet som identifierats, uppdelat mellan material med sexuella övergrepp mot barn och kontaktsökning med barn

Tabell nr 1: Antal fall av sexuella övergrepp mot barn på nätet som upptäcktes under 2021

Tillhandahållare	Antal fall av material med sexuella övergrepp mot barn identifierade under 2021	Kommentarer
Google	33 Google Chat-konton	Detta avser antalet Google Chat-konton för EU-användare där sexuella övergrepp mot barn har upptäckts med hjälp av automatiserad teknik under perioden den 2 augusti till den 31 december 2021. Inga uppgifter lämnades om antalet identifierade innehållsposter.
LinkedIn	31 innehållsposter	31 innehållsposter bekräftades vara material med sexuella övergrepp efter manuell granskning och rapporterades till NCMEC, 6 var känt material och de återstående 25 var okänt material.
Meta	340 000 konton	Antal konton som upptäckts skicka minst en fil som utgör material med sexuella övergrepp mot barn i meddelandetrådar som omfattade en EU-användare mellan den 8 november 2021 och den 31 december 2021.
Microsoft	6 600 innehållsposter	6 600 innehållsposter (enstaka bild eller video) bekräftade som material med sexuella övergrepp mot barn från EU av över 20 000 innehållsposter som upptäckts globalt mellan juli 2021 och december 2021.
X (tidigare Twitter)	532 898 konton	Konton (oklart om endast från EU eller globalt) som stängdes av på grund av brott mot X:s policy om sexuell exploatering av barn mellan den 2 augusti 2021 och den 31 december 2021.

Av de uppgifter som lämnats av X framgår det inte om det enbart rör sig om tjänster som omfattas av tillämpningsområdet för interimsförordningen (nummeroberoende interpersonella kommunikationstjänster) eller om denna siffra även omfattar andra tjänster (såsom informationssamhällets tjänster). Detta gäller alla uppgifter som rör X i denna rapport.

Tabell nr 2: Antal fall av sexuella övergrepp mot barn på nätet som upptäcktes under 2022

Tillhandahållare	Antal fall av material med sexuella övergrepp mot barn identifierade under 2022	Kommentarer
Google	2 045 innehållsposter	Upptäckta innehållsposter som rapporterats till NCMEC i 752 Google-konton genom automatiserad teknik för EU-

		användare.
LinkedIn	2 innehållsposter	LinkedIn upptäckte 2 bilder och 0 videor som bekräftats vara material med sexuella övergrepp mot barn.
Meta	6,6 miljoner innehållsposter	Filer som utgör åtgärdat material med sexuella övergrepp mot barn upptäckta med hjälp av Metas mediamatchningsteknik i meddelandetrådar som omfattade en EU-användare.
Microsoft	12 800 innehållsposter	12 800 innehållsposter (enstaka bild eller video) bekräftade som material med sexuella övergrepp mot barn från EU av över 50 000 innehållsposter upptäckta globalt under 2022.
X (tidigare Twitter)	2 348 712 konton	Konton (oklart om endast från EU eller globalt) som stängdes av på grund av brott mot X:s policy om sexuell exploatering av barn.

2.1.5. Prövning för användare och utfall

Enligt artikel 3.1 g iv i interimsförordningen ska tillhandahållarna inrätta lämpliga förfaranden och prövningsmekanismer för att säkerställa att användarna kan lämna in klagomål till dem. Artikel 5 innehåller dessutom regler för rättsmedel.

Tillhandahållarna uppgav att de infört interna förfaranden och prövningsmekanismer för användare vars konton har begränsats på grund av delning av material med sexuella övergrepp mot barn och/eller där innehållet har tagits bort, så att de kan överklaga begränsningen/borttagningen och få sitt ärende granskat.

De rapporterade fall som omfattas av förordningen där en användare har lämnat in ett klagomål med hjälp av den interna prövningsmekanismen eller till en rättslig myndighet inom EU samt utfallen av sådana klagomål. Ingen tillhandahållare förutom Microsoft (som rapporterade 0 klagomål i båda kanalerna under 2021 och 2022) rapporterade separat statistik över intern och rättslig prövning. Därför innehåller tabellerna nedan både interna och rättsliga prövningsförfaranden.

Tabell nr 3: Antal fall där en användare har lämnat in klagomål med hjälp av den interna prövningsmekanismen eller till en rättslig myndighet samt utfallet av sådana klagomål under 2021

Tillhandahållare	Antal klagomål från användare	Antal återställda konton	Antal fall av återinfört innehåll	Kommentarer
Google	8	0	i.u.	Google Chat-konton avstängda på grund av sexuella övergrepp mot barn på nätet där användaren överklagade: 8. Inget återställdes.

LinkedIn	0	i.u.	i.u.	
Meta	4 900	i.u.	207	4 900 användare överklagade. Efter överklagandeprocessen fick 207 användare sitt innehåll återinfört och kontoåtgärderna upphävdes.
Microsoft	0	i.u.	i.u.	
X (tidigare Twitter)	ca 90 000	ca 3 000	i.u.	Ca 90 000 överklaganden. X återställde ca 3 000 av dessa konton.

Tabell nr 4: Antal fall där en användare har lämnat in klagomål med hjälp av den interna prövningsmekanismen eller till en rättslig myndighet samt utfallet av sådana klagomål under 2022.

Tillhandahållare	Antal klagomål från användare	Antal återställda konton	Antal fall av återinfört innehåll	Kommentarer
Google	378	0	i.u.	Google-konton inaktiverade p.g.a. sexuella övergrepp mot barn på nätet när användaren överklagade: 378. Inget återställdes.
LinkedIn	0	i.u.	i.u.	
Meta	29 000	i.u.	3 700	Användarna överklagade åtgärderna för cirka 29 000 delade filer. Efter överklagandeprocessen återställdes cirka 3 700 innehållsfiler och kontoåtgärderna upphävdes.
Microsoft	0	i.u.	i.u.	
X (tidigare Twitter)	Ca 430 000	Ca 4 000	i.u.	Ca 430 000 överklaganden. X återställde ca 4 000 av dessa konton.

2.1.6. Antal och andel fel (falska positiva resultat) för de olika tekniker som användes

Enligt artikel 3.1 e i interimförordningen ska tillhandahållarna se till att den teknik som används är tillräckligt tillförlitlig genom att den i största möjliga utsträckning begränsar felfrekvensen vad gäller upptäckten av innehåll som utgör sexuella övergrepp mot barn på nätet.

Här uppgav tillhandahållarna att de inte använder varje typ av teknik separat för att upptäcka sexuella övergrepp mot barn på nätet. De använder i stället en skiktad metod för att spåra övergreppen genom att kombinera olika typer av detektionsteknik för att öka noggrannheten. För att minska fel eller falska positiva resultat kompletterar samtliga tillhandahållare sådan teknik med mänsklig granskning. Tillhandahållarna uppgav inte antalet och andelen fel (falska positiva resultat) för var och en av de olika typerna av teknik separat, utan rapporterade sammanlagda uppgifter för all teknik som användes.

De flesta tillhandahållare mäter antalet och andelen fel som upphävda verkställighetsbeslut, dvs. som den totala andelen återställningar/upphävanden efter överklagande (t.ex. vilken andel avstängda konton eller borttaget innehåll som tillhandahållaren återinförde efter att användaren överklagade). Tillhandahållarnas metod återspeglar inte nödvändigtvis definitionen av ”falska positiva resultat” i statistiken.

Den rapporterade andelen återställningar/upphävanden är följande:

Tabell nr 5: Andel återställningar/upphävanden efter överklagande

Tillhanda hållare	2021*		2022		Kommentarer
	% återställda konton av antalet överklaganden	% återställda av antalet avstängda konton	% återställda konton av antalet överklaganden	% återställda av antalet avstängda konton	
Google	0 % (0 av 8)	0 % (0 av 33)	0 % (0 av 378)	0 % (se kommentarer)	Antalet avstängda konton angavs inte för 2022. I stället angavs antalet innehållsfiler som upptäckts och rapporterats till NCMEC, vilket var 2 045. Inget konto återställdes efter överklagandet.
LinkedIn	0 %	0 %	0 %	0 %	Inga överklaganden. För perioden från den 13 juli 2021 till den 31 december 2021 rapporterade LinkedIn också att av de 75 filer som granskats och som kom från EU bekräftades 31 efter mänsklig granskning utgöra material med sexuella övergrepp mot barn. LinkedIn lämnade inga sådana uppgifter för 2022.
Meta	4,22 % (207 av 4 900)	0,06 % (207 av 340 000)	Se kommentarer	Se kommentarer	Det lämnades ingen information som gör att det går att exakt fastställa omfattningen, innehållet i överklagandena eller skälen till återställandena. För 2022 lämnades uppgifterna i form av innehåll, inte i form av konton: <ul style="list-style-type: none"> - Antal innehållsfiler som lett till avstängning ("åtgärdats"): 6,6 miljoner - Antal innehållsfiler som omfattades av överklaganden: 29 000 - Antal innehållsfiler som återställdes: 3 700 Detta innebär <ul style="list-style-type: none"> - % återställda innehållsfiler av antal överklaganden:

					12,8 % (3 700 av 29 000) - % återinförda innehållsfiler av antal borttagna innehållsfiler: 0,06 % (3 700 av 6,6 miljoner)
Microsoft	0 %	-	-	-	Otillräckliga uppgifter för att beräkna andelen återställningar/upphävanden efter överklaganden. Totalt 17 upphävanden av det ursprungliga beslutet om innehållsmoderering för 2022, inga uppgifter om totalt antal överklaganden.
X (tidigare Twitter)	1,43 % (100 av 7 000)	0,06 % (100 av 166 000)	2,17 % (500 av 23 000)	0,10 % (500 av 501 000)	Under andra halvåret 2021 stängdes ca 166 000 användare av via automatiserade mekanismer på grund av sexuella övergrepp mot barn. Av dessa användare lämnade ca 7 000 in överklaganden som resulterade i ca 100 upphävda beslut. Under 2022 stängdes 501 000 användare av via automatiserade mekanismer på grund av sexuella övergrepp mot barn. Av dessa användare lämnade ca 23 000 in överklaganden som resulterade i ca 500 upphävda beslut.

* Rapporteringsperioderna för 2021 skiljer sig åt för varje tillhandahållare

2.1.7. Åtgärder som vidtagits för att begränsa felfrekvensen och den felprocent som uppnåtts

Enligt artikel 3.1 e i interimsförordningen ska den teknik som används vara tillräckligt tillförlitlig och konsekvenserna av sporadiska fel som ändå uppstår ska åtgärdas utan dröjsmål. Dessutom krävs enligt artikel 3.1 g ii mänsklig tillsyn och, vid behov, mänskligt ingripande.

Samtliga tillhandahållare uppgav att de använder en skiktad metod för att upptäcka och bekämpa spridningen av sexuella övergrepp mot barn på nätet. Här ingår användning av hashmatchningsteknik (däribland PhotoDNA) för att upptäcka material med sexuella övergrepp mot barn i kombination med mänskliga granskningsprocesser för att bekräfta om en mediafil (bild och video) innehåller övergrepp, men också mänsklig tillsyn över behandlingen.

Tillhandahållarna uppgav att de tillämpade olika skyddsåtgärder för att begränsa och minska felfrekvensen i samband med detektering, rapportering och borttagning av sexuella övergrepp mot barn på nätet. Här ingår (ingen uttömmande förteckning)⁴

- i. övervakning och kvalitetsbedömning av prestandan hos de verktyg som används för att upptäcka sexuella övergrepp, både för att finjustera precisionen (att de endast upptäcker sexuella övergrepp mot barn på nätet) och för att få en tillbakablick (att de på sina plattformar inte missar sexuella övergrepp mot barn på nätet) (Google, X),
- ii. mänsklig granskning och tillsyn: ett urval av media som har identifierats som material med sexuella övergrepp mot barn av hashmatchningstekniker granskas av mänskliga granskare/utbildade analytiker (Google, LinkedIn, Meta, Microsoft),
- iii. flaggning och granskning av högvolymkluster (Meta),
- iv. användning av ytterligare manuella granskningsprocesser såsom löpande kontroller av hashkvalitet (LinkedIn, Microsoft),
- v. mänskliga granskare som genomgår specialiserade omfattande utbildningar under vägledning av rådgivare i hur man känner igen material med sexuella övergrepp mot barn för att se till att den mänskliga granskningen är korrekt (Google),
- vi. regelbundna kvalitetskontroller av mänskliga granskare och de bedömningar som gjorts (Google, X),
- vii. andra kvalitetskontrollprocesser för att minska antalet fel och omedelbara åtgärder, t.ex. oberoende hashverifiering (Google, LinkedIn), mänsklig granskning av varje fall av material med sexuella övergrepp mot barn som inte förekommit tidigare före rapportering (Google),
- viii. utveckling och regelbunden översyn av riktlinjer och verkställighetsstrategier av ämnesexperter utbildade inom sexuella övergrepp mot barn på nätet (Google),
- ix. arbete med kvaliteten i NCMEC:s CyberTipline-rapporter och falska positiva resultat, i de fall sådana förekommer (Google, LinkedIn, Meta, Microsoft, X).

⁴ De tillhandahållare som står inom parentes är de som specifikt uppgav de konkreta åtgärderna. Om en tillhandahållare inte finns med på förteckningen betyder inte det att de inte genomför denna åtgärd, utan bara att de inte har nämnt den i sin rapport.

2.1.8. Policyn för lagring samt de dataskyddsåtgärder som tillämpas enligt dataskyddsförordningen

Enligt artikel 3.1 h och i i interimsförordningen ska relevanta personuppgifter lagras på ett säkert sätt endast för vissa specifika syften och innehålla specifikationer för lagringsperioden. Dessutom måste de tillämpliga kraven i dataskyddsförordningen följas.

Tillhandahållarna uppgav att de hade robusta policyer för lagring och hade vidtagit åtgärder för skydd av personuppgifter. Policyer för datalagring skiljer sig åt beroende på vilken typ av uppgifter det handlar om. De uppger att lagringsperioden i varje enskilt fall är begränsad till den tid som anses lämplig utifrån datatypen och syftet med behandlingen och att uppgifterna raderas i slutet av lagringsperioden. I tillhandahållarnas integritetspolicyer/-förklaringar och service-/användaravtal finns mer detaljerad information om metoder för datalagring.

De flesta tillhandahållarna (LinkedIn, Meta, Microsoft) tillämpar en lagringpolicy på 90 dagar för media som bekräftats innehålla material med sexuella övergrepp mot barn och som upptäckts inom förordningens tillämpningsområde. Under denna period lagras det innehåll som bekräftats vara material med sexuella övergrepp mot barn i en separat och säker lagring som hanteras av specialiserade team (t.ex. Microsofts team för brottsbekämpning och nationell säkerhet). I dessa lagringssystem raderas automatiskt det lagrade materialet med sexuella övergrepp mot barn efter 90 dagar om inte någon förlängning av lagringsperioden har begärts i samband med rättsliga processer, vilket är normalt när brottsbekämpande myndigheter följer upp NCMEC:s rapporter.

Google uppgav att material som upptäcks inom förordningens tillämpningsområde inte lagras längre än absolut nödvändigt för de relevanta syftena enligt förordningen och i varje fall inte längre än 12 månader från det att materialet har identifierats och rapporterats, men denna period kan förlängas om det finns en giltig rättslig begäran att bevara det.

X (tidigare Twitter) uppgav att de sparar profilinformation och innehåll under den tid som användarkontot är i bruk och personuppgifter som samlas in i samband med användningen av tjänsten i högst 18 månader. När ett konto inaktiveras av användaren sparar X i allmänhet uppgifterna i ytterligare 30 dagar innan kontot raderas. Användaruppgifter relaterade till klagomål och överträdelse av policyn, däribland kontoinformation för den som har brutit mot policyn (t.ex. identifierare som används för att skapa kontot som e-postadress eller telefonnummer) sparas på obestämd tid. Detta för att förhindra att personer som upprepade gånger bryter mot X:s policy skapar nya konton, men också för att se till att personer som bryter mot X:s policy inte bara kan vänta ut avstängningsperioden och sedan bryta mot policyn igen⁵.

Bland de skyddsåtgärder för personuppgifter som tillhandahållarna tillämpar ingår standardåtgärder för branschen (alla eller ett urval), såsom (ingen uttömmande förteckning)⁶

⁵ *X Privacy Policy, 4. How Long We Keep Information*, tillgänglig på: <https://twitter.com/en/privacy>.

⁶ De tillhandahållare som står inom parentes är de som specifikt uppgav de konkreta åtgärderna. Om en tillhandahållare inte finns med på förteckningen betyder inte det att de inte genomför denna åtgärd, utan bara att de inte har nämnt den i sin rapport.

- i. användning av teknik för avidentifiering eller pseudonymisering och anonymisering av data (t.ex. maskering, hashing, differentiell integritet) (Google, LinkedIn, Meta, Microsoft),
- ii. utlämning av endast hashvärden till tredje part i syfte att upptäcka material med sexuella övergrepp mot barn (Google, LinkedIn),
- iii. användning av kryptering enligt branschstandard (algoritmer och protokoll) för uppgifter som överförs mellan privatägd infrastruktur och offentliga nätverk (Meta),
- iv. strategier för datastyrning/omfattande integritetsprogram (X, tidigare Twitter, Google) och strikta interna restriktioner för åtkomst till data (Meta) (tillämpas exempelvis på personal, entreprenörer, agenter som behöver informationen för att kunna behandla den), användning av åtkomstkontrollistor mellan olika granskningsverktyg för material med sexuella övergrepp mot barn och hashförbud (Meta) samt strikta avtalsenliga sekretesskrav för alla som har åtkomst,
- v. granskning av strategier för anonymisering och datastyrning, dvs. genomförande av integritetsgranskningar för att identifiera, komma åt och minska potentiella risker för integriteten vid insamling, behandling, lagring och delning av personuppgifter samt granskning av praxis för skydd av personuppgifter (Microsoft),
- vi. beredskapsplaner för hantering av säkerhetsincidenter för att övervaka, upptäcka och hantera eventuella säkerhetsproblem och incidenter i infrastrukturen (Google, Meta).

2.1.9. Namn på organisationer som agerar i allmänhetens intresse mot sexuella övergrepp mot barn och med vilka uppgifter har delats enligt denna förordning.

Under båda rapporteringsperioderna (juli/augusti 2021 till december 2021 och januari 2022 till december 2022) uppgav samtliga tillhandahållare att de delade uppgifter som behandlas enligt denna förordning med NCMEC. I enlighet med artikel 7.1 i interimsförordningen uppgav samtliga tillhandahållare att de även hade meddelat kommissionen att de rapporterade sexuella övergrepp mot barn på nätet till NCMEC enligt denna förordning⁷.

2.2. Statistik från medlemsstaterna (artikel 8)

Medlemsstaterna är skyldiga att enligt artikel 8.1 i interimsförordningen lämna statistik om följande:

- (a) Det totala antalet rapporter om upptäckta sexuella övergrepp mot barn på nätet som har lämnats in av tillhandahållare och organisationer som agerar i allmänhetens intresse mot sexuella övergrepp mot barn till de behöriga nationella brottsbekämpande myndigheterna, uppdelat mellan det absoluta antalet fall och fall som rapporterats flera gånger om sådan information finns tillgänglig samt typen av tillhandahållare på vars tjänst det sexuella övergreppet mot barn på nätet upptäcktes.
- (b) Antalet barn som identifierats genom åtgärder enligt artikel 3, uppdelat efter kön.

⁷ Informationen om de organisationer som agerar i allmänhetens intresse och till vilka tillhandahållarna rapporterar sexuella övergrepp mot barn på nätet enligt denna förordning har offentliggjorts på https://home-affairs.ec.europa.eu/policies/internal-security/child-sexual-abuse/legal-framework-protect-children_sv i linje med kommissionens skyldigheter enligt artikel 8.2 i interimsförordningen.

(c) Antalet dömda förövare.

De flesta medlemsstater lämnade åtminstone en del information, men i många av medlemsstaterna hade ännu inga relevanta system för insamling och rapportering av uppgifter införts. Det innebär att i de fall statistik lämnades avsåg den olika rapporteringsperioder och skilde sig avsevärt åt när det gäller detaljnivå. Några medlemsstater lämnade in årlig statistik från och med det datum då förordningen trädde i kraft. De flesta av dem redovisade per kalenderår eftersom de troligtvis inte hade tekniska möjligheter att särskilja den begärda statistiken per år från och med det datum då förordningen trädde i kraft. Ett fåtal medlemsstater lämnade inga uppgifter alls.

Det bör också noteras att de statistiska uppgifterna i vissa fall hämtas från s.k. aktuella databaser, eller journalförings- och ärendehanteringssystem, dvs. inga egentliga statistiksystem. De lämnade siffrorna baseras ibland på dynamiska uppgifter, vilket innebär att de inte är slutgiltiga utan kan komma att ändras. Beroende på tidpunkten för inhämtning av uppgifter (t.ex. i Slovenien och Danmark) sker förändringar när fler utredningar och domstolsförfaranden avslutas.

I flera medlemsstater inrättar de behöriga myndigheterna nya avdelningar för att utreda brott som rör sexuella övergrepp mot barn på nätet samt central rapportering för dessa brott (Lettland, Tjeckien). I framtiden bör detta leda till mer exakt statistik.

Tyskland uppgav att det inte kunde lämna någon statistik enligt artikel 8.1 i interimsförordningen, eftersom man inte ansåg sig ha någon rättslig grund för frivillig spårning⁸. På sin webbplats rapporterar dock den tyska federala kriminalpolisen (BKA) att den under 2022 tog emot 89 844 rapporter från NCMEC samtidigt som NCMEC uppgav att 138 193 rapporter skickades till de tyska myndigheterna.⁹ Tre medlemsstater lämnade inga uppgifter eller motiveringar till varför de inte rapporterade enligt denna bestämmelse (Malta, Portugal och Rumänien).

2.2.1. Totalt antal rapporter om upptäckta sexuella övergrepp mot barn på nätet

De flesta medlemsstater lämnade viss statistik om det totala antalet rapporter om sexuella övergrepp mot barn på nätet enligt artikel 8.1 a i interimsförordningen. Eftersom medlemsstaterna lämnade uppgifter för olika rapporteringsperioder gick det inte att beräkna hur många rapporter om upptäckta sexuella övergrepp mot barn på nätet som sammanlagt tagits emot på EU-nivå under en viss period, exempelvis tiden efter genomförandet av förordningen.

⁸ Rapport från Tyskland enligt artikel 8 i förordning (EU) 2021/1232, mottagen den 18 oktober 2022. NCMEC publicerar alla uppgifter om rapporter som tagits emot och som rör EU:s medlemsstater, däribland Tyskland, i sina CyberTipline-rapporter uppdelat per land. Se följande länk: NCMEC, [2021 CyberTipline Reports by Country](#), hämtad i juli 2023; NCMEC, [2022 CyberTipline Reports by Country](#), hämtad i juli 2023.

⁹ https://www.bka.de/SharedDocs/Kurzmeldung/DE/Kurzmeldung/230623_Mindestspeicherfristen_IP-Adressen.html

Medlemsstaterna uppgav i de flesta fall det totala antalet rapporter som mottagits av de nationella brottsbekämpande myndigheterna från tillhandahållare eller andra organisationer som agerar i allmänhetens intresse mot sexuella övergrepp mot barn. Med tanke på att de flesta USA-baserade tillhandahållare rapporterar till NCMEC uppgav också majoriteten av medlemsstaterna att de mottog de flesta eller samtliga rapporter från NCMEC. Medlemsstaterna uppgav inte antalet åtgärdbara rapporter, dvs. rapporter som lämpar sig för utredning, men vissa uppgav antalet inledda ärenden, vilket är betydligt lägre. Skillnaden mellan mottagna rapporter och utredda fall hade flera orsaker, t.ex. att rapporten innehöll material med sexuella övergrepp mot barn, men att man inte hade tillräckligt med information för att inleda en utredning, sammanslagning av rapporter när mer än en rapport gäller en och samma misstänkt eller att materialet, trots att det visade exploaterande situationer, inte bedömdes som brottsligt enligt nationell lag. Dessutom skilde de flesta medlemsstaterna inte mellan det absoluta antalet fall och de fall som rapporterats flera gånger. I de rapporter som lämnades av NCMEC hade NCMEC redan i förväg kategoriserat rapporterna från tillhandahållarna som ”åtgärdbara” eller ”informativa”. En åtgärdbar rapport definieras av NCMEC som en rapport med tillräcklig information för att starta en utredning. Här ingår vanligtvis användaruppgifter, bilder och en möjlig plats. Rapporten kategoriseras som ”informativ” när den innehåller otillräckligt med information eller när bilden anses vara viral och har rapporterats många gånger. Under 2022 betecknade NCMEC 49 % av rapporterna som ”åtgärdbara”, medan 51 % betecknades som ”informativa”.

Endast ett fåtal medlemsstater angav typen av tillhandahållare på vars tjänst sexuella övergrepp mot barn på nätet hade upptäckts (t.ex. Belgien, Tjeckien, Estland, Frankrike och Polen) och endast en medlemsstat gjorde en detaljerad fördelning (Belgien).

Slovenien uppgav att man inte kunde lämna siffror för brott som utretts enbart på grund av rapporter från tillhandahållare och organisationer, utan bara kunde lämna siffror för alla utredningar om sexuella övergrepp mot barn på nätet, oavsett vilken informationskälla som ledde till att utredningen inleddes.

Tabell nr 6: Totalt antal rapporter om upptäckta sexuella övergrepp mot barn på nätet per medlemsstat

Land	Rapporteringsperiod	Totalt antal rapporter om sexuella övergrepp mot barn på nätet	Källa till rapporterna	Kommentarer
Österrike	2021 till 2022	16 311	NCMEC	
Belgien	1 augusti 2021 till 31 juli 2022	26 226	Rapporter från tillhandahållare (sociala medier) och Childfocus hotline	
Bulgarien	2021 till 2022	9 120	Tillhandahållare och INHOPE hotline genom "Safenet" och andra	Av dessa handlade 9 112 varningar om webbsidor som innehöll material med sexuella övergrepp mot barn och som drevs av bulgariska tillhandahållare.
Kroatien	1 januari 2021 till 31 oktober 2022	9 044	NCMEC	
Cypern	1 juli 2021 till 31 december 2022	3 570	NCMEC	
Tjeckien	1 januari 2022 till 31 juli 2022	13 279	NCMEC	
Danmark	2 augusti 2021 till 20 januari 2023	10 744	NCMEC	
Estland	-	-	NCMEC, Child Helpline 116 111	Estland uppgav att statistik för polis och gränsbevakning, inklusive NCMEC, inte är offentlig. För 2021 rapporterade Estland 360 sexualbrott mot ett barn utan fysisk kontakt. Dessutom begicks 86 % av alla sexualbrott utan fysisk kontakt i internetmiljön eller med hjälp av informationstekniska verktyg.
Finland	2022	25 000	NCMEC och Rädda Barnen	

Frankrike	1 augusti 2021 till 1 augusti 2022	120 000	NCMEC	
Tyskland	-	-	-	Uppgifter ej tillgängliga/ej rapporterade.
Grekland	2021 till 2022	142	NCMEC, grekisk hotline för olagligt internetinnehåll – Safeline, nationella telekommunikations- och postkommissionen, nationella linjen SOS 1056 – The smile of the child, grekiska ombudsmannen	
Ungern	2022	0	Ingen av de rapporter som lämnades av tillhandahållarna skickades enligt interimsförordningen	
Irland	2021 till 2022	15 355	NCMEC	
Italien	2022	4 607	Ej specificerat	
Lettland	1 augusti 2022 till 6 mars 2023	Cirka 115 till 220 rapporter per månad	Från icke-lettiska tillhandahållare och organisationer som agerar i allmänhetens intresse mot sexuella övergrepp mot barn (främst NCMEC) och lettiska tillhandahållare och organisationer (främst Latvian Safer Internet Center)	
Litauen	1 januari 2021 till 30 juni 2022	4 142	Ej specificerat	
Luxemburg	2021 till 2022	2 491	Ej specificerat	
Malta	-	-	-	Inga uppgifter har lämnats/rapporterats.
Nederländerna	2021	36 537	Tillhandahållare och organisationer som agerar i allmänhetens intresse mot	

			sexuella övergrepp mot barn	
Polen	3 augusti 2021 till 3 augusti 2023	227	Tillhandahållare och organisationer som agerar i allmänhetens intresse mot sexuella övergrepp mot barn	För perioden från den 3 augusti 2022 till den 3 augusti 2023 fick Polen in 1 rapport om grooming av barn och 105 rapporter om material med sexuella övergrepp mot barn.
Portugal	-	-	-	Inga uppgifter har lämnats/rapporterats.
Rumänien	-	-	-	Inga uppgifter har lämnats/rapporterats.
Slovakien	1 augusti 2021 till 31 juli 2022	7 206	Tillhandahållare och organisationer som agerar i allmänhetens intresse mot sexuella övergrepp mot barn	
Slovenien	1 januari 2021 till 14 juli 2023	452	Siffran anger brott i samband med aktiviteter på internet. För närvarande kan inte Slovenien utifrån den befintliga statistiken skilja mellan statistiska uppgifter om brott som utretts på grundval av rapporter som lämnats av tillhandahållare och organisationer och statistiska uppgifter om andra rapporter.	
Spanien	2022	31 474	Organisationer som agerar i allmänhetens intresse mot sexuella övergrepp mot barn	
Sverige	Augusti 2021 till 31 december 2022	32 830	Främst NCMEC	

Med tanke på att NCMEC är den främsta källan till rapporter är det intressant att jämföra medlemsstaternas siffror för hur många rapporter som de har mottagit från NCMEC med NCMEC:s siffror för hur många rapporter som har skickats till medlemsstaterna. NCMEC mottog under 2021 totalt 29 397 681 rapporter från branschen globalt, varav 99,7 % (eller 29 309 106) avsåg en eller flera bilder eller videor med sexuella övergrepp mot barn, 0,15 % (eller 44 155) rörde grooming och 0,05 % (eller 16 032) rörde sexhandel med barn. Under 2022 mottog NCMEC totalt 32 059 029 rapporter, varav 99,5 % (eller 31 901 234) avsåg bilder eller videor med sexuella övergrepp mot barn, 0,25 % (80 524) avsåg grooming och 0,06 % (eller 18 336) sexhandel med barn. För EU är fördelningen följande:

Tabell nr 7: NCMEC:s rapporter om misstänkta sexuella övergrepp mot barn på nätet lämnade till EU:s medlemsstater under 2021 och 2022

Land	Totalt antal rapporter 2021 ¹⁰	% av EU totalt 2021	Totalt antal rapporter 2022 ¹¹	% av EU totalt 2022	% av EU:s befolkning
Österrike	7 580	1,36 %	18 501	1,23 %	2,00 %
Belgien	15 762	2,84 %	50 255	3,34 %	2,60 %
Bulgarien	13 584	2,44 %	31 937	2,12 %	1,53 %
Kroatien	4 744	0,85 %	11 693	0,78 %	0,86 %
Cypern	2 657	0,48 %	7 361	0,49 %	0,20 %
Tjeckien	15 004	2,70 %	61 994	4,12 %	2,36 %
Danmark	5 891	1,06 %	30 215	2,01 %	1,31 %
Estland	2 729	0,49 %	6 408	0,43 %	0,30 %
Finland	6 079	1,09 %	10 904	0,73 %	1,24 %
Frankrike	98 233	17,67 %	227 465	15,13 %	15,16 %
Tyskland	79 701	14,34 %	138 193	9,19 %	18,59 %
Grekland	14 616	2,63 %	43 345	2,88 %	2,37 %
Ungern	31 710	5,70 %	109 434	7,28 %	2,16 %
Irland	7 327	1,32 %	19 770	1,31 %	1,13 %
Italien	37 480	6,74 %	96 512	6,42 %	13,32 %
Lettland	1 537	0,28 %	3 688	0,25 %	0,42 %
Litauen	3 509	0,63 %	16 603	1,10 %	0,63 %
Luxemburg	2 005	0,36 %	2 004	0,13 %	0,14 %
Malta	750	0,13 %	4 713	0,31 %	0,12 %
Nederländerna	36 790	6,62 %	57 012	3,79 %	3,96 %
Polen	37 758	6,79 %	235 310	15,65 %	8,41 %
Portugal	34 415	6,19 %	42 674	2,84 %	2,31 %
Rumänien	32 765	5,89 %	96 287	6,40 %	4,25 %
Slovakien	7 275	1,31 %	39 748	2,64 %	1,21 %

¹⁰ NCMEC, [2021 CyberTipline Reports by Country](#), hämtat i november 2023.

¹¹ NCMEC, [2022 CyberTipline Reports by Country](#), hämtat i november 2023.

Slovenien	3 162	0,57 %	14 795	0,98 %	0,47 %
Spanien	33 136	5,96 %	77 727	5,17 %	10,60 %
Sverige	19 635	3,53 %	48 883	3,25 %	2,33 %
Totalt	555 834		1 503 431		

De avsevärda skillnaderna mellan antalet rapporter 2021 och 2022 med en kraftig ökning under 2022 beror främst på att den frivilliga spårningen låg på en lägre nivå mellan januari och augusti 2021 innan interimsförordningen började tillämpas.

I NCMEC:s statistik per EU-medlemsstat skiljer man inte mellan källan till rapporten, i synnerhet anges inte om den kommer från en nummeroberoende interpersonell kommunikationstjänst. NCMEC lämnar bara statistik över det totala antalet rapporter som rör EU och som härstammar från nummeroberoende interpersonella kommunikationstjänster. Under 2021 kom 283 265 rapporter som rörde medlemsstaterna från chatt-, meddelande- eller e-posttjänster, dvs. 51 % av det totala antalet rapporter som rörde EU. Ytterligare 164 645 (30 % av det totala antalet) rapporter kom från sociala medier eller spelplattformar, där det också förekommer integrerade meddelande- eller chatttjänster. Under 2021 handlade 3 565 av de EU-relaterade rapporterna om grooming. Under 2022 kom 1 015 231 rapporter som gällde medlemsstaterna från chatt-, meddelande- eller e-posttjänster, dvs. 68 % av det totala antalet rapporter som rörde EU. Ytterligare 325 847 (22 % av det totala antalet) rapporter kom från sociala medier eller spelplattformar, där det också förekommer integrerade meddelande- eller chatttjänster. Under 2022 handlade 7 561 av de EU-relaterade rapporterna om grooming. Skillnaderna i antalet rapporter från tillhandahållare av nummeroberoende interpersonella kommunikationstjänster under 2021 och 2022 beror återigen på att den frivilliga spårningen var lägre mellan januari och augusti 2021 innan interimsförordningen började tillämpas.

Andelen rapporter per medlemsstat motsvarar i stort sett medlemsstatens befolkningsandel i förhållande till hela EU:s befolkning, vilket skulle kunna tyda på en jämförbar förekomst av sexuella övergrepp mot barn på nätet i medlemsstaterna. Betydande avvikelser förekom för Spanien och Italien, vars procentandelar under båda åren framstår som låga jämfört med procentandelen av EU:s befolkning, samtidigt som andelen rapporter för andra medlemsstater varierade kraftigt (t.ex. Tyskland, Polen, Nederländerna och Slovakien). Dessa avvikelser återspeglas inte som sådana i rapporterna om antalet fall och det är därför återigen svårt att dra slutsatser om sambandet mellan rapporter och utredningar.

Med tanke på de varierande rapporteringsperioderna går det inte att göra någon direkt jämförelse, men det finns ändå betydande skillnader mellan NCMEC:s statistik och de siffror som medlemsstaterna har lämnat. Dessutom kan inte heller NCMEC:s siffror för medlemsstaterna fullt ut jämföras med de siffror från branschen som rapporterades i föregående avsnitt. Vissa av skillnaderna kan bero på att rapporter om sexuella övergrepp mot barn på nätet kommer från andra källor än interpersonell kommunikation. Det skulle kräva närmare analys, eftersom det med tanke på den förteckning över tillhandahållare som rapporterar till NCMEC även kan finnas andra tillhandahållare som vidtar frivilliga spårningsåtgärder som rör EU än de som hittills har

lämnat in rapporter till kommissionen¹². Men det faktum att det för de flesta medlemsstater tycks finnas en betydande skillnad mellan det antal rapporter som NCMEC anger att de har skickat till medlemsstaten och det antal rapporter som medlemsstaten anger att den har mottagit, tyder på att medlemsstaternas insamling och rapportering av uppgifter inte är fullständig.

Vid varje NCMEC-rapport som nämns ovan raderades bilder och videoklipp av sexuella övergrepp mot barn och hindrades från att spridas. Detta är särskilt viktigt för de nuvarande brottsoffren och för barn som tidigare har utsatts för sexuella övergrepp. Studier har visat att den fortsatta spridningen av bilder och videoklipp som visar övergreppen begränsar offrens förmåga att hantera de psykiska effekterna av övergreppen och skapar en sekundär form av viktimering.

2.2.2. Antalet barn som identifierats

De flesta medlemsstater lämnade fullständig eller partiell statistik över antalet barn som identifierats, uppdelad efter kön, i enlighet med artikel 8.1 b i interimsförordningen. Men flera medlemsstater lämnade inga uppgifter eller skäl till att de inte lämnade in rapporter enligt denna bestämmelse.

Flera medlemsstater lämnade ingen eller enbart partiell statistik för rapporteringsperioden, men angav skälen till detta. De skäl som angavs var att

- det inte går att räkna antalet barn som fallit offer för sexuella övergrepp på nätet (Frankrike),
- uppgifterna inte är tillgängliga eftersom de inte samlats in inom ramen för den nationella statistikinsamlingen eller att de nationella myndigheterna inte registrerade sådan statistik (Danmark, Litauen),
- uppgifterna inte är uppdelade efter kön i den nationella statistikinsamlingen (Belgien, Cypern, Tjeckien, Grekland, Irland, Italien, Litauen, Nederländerna),
- informationen inte är tillgänglig på den begärda detaljnivån i de befintliga informationssystemen (Finland),
- informationen inte har samlats in (Tyskland).

Vissa av de medlemsstater som uppgav att de inte kunde lämna statistik bekräftade att deras nationella myndigheter har uppmanats att ändra sitt registreringsförfarande för frivilliga rapporter och utredningar samt insamling av statistik (Danmark) och/eller håller på att införa nya informationssystem som ska göra det möjligt att rapportera på den detaljnivå som krävs (Finland).

I en medlemsstat skiljer uppgifterna nedan inte mellan barn som fallit offer för sexuella övergrepp på och utanför nätet (Ungern). I vissa fall omfattar statistiken även barn som själva har producerat och laddat upp material (egengenererat material, mestadels videor) (Tjeckien, Estland).

¹² NCMEC:s uppgifter finns tillgängliga [här](#).

Eftersom medlemsstaterna i de flesta fall lämnade uppgifter för olika rapporteringsperioder gick det inte att beräkna det totala antalet barn som i EU identifierats som offer för sexuella övergrepp mot barn på nätet per år och/eller under samma rapporteringsperiod.

Tabell nr 8: Antalet barn som identifierats, uppdelat efter kön

Land	Rapporteringsperiod	Flickor	Pojkar	Totalt	Kommentarer
Österrike	2021 till 2022	11	6	17	
Belgien	2021 till 2022	-	-	63	Könsuppdelade uppgifter inte tillgängliga.
Bulgarien	2022	50	12	62	
Kroatien	1 januari 2021 till 31 oktober 2022	20	0	20	
Cypern	2022	-	-	102	Könsuppdelade uppgifter inte tillgängliga.
Tjeckien	2022	-	-	30	Könsuppdelade uppgifter inte tillgängliga.
Danmark	-	-	-	-	Inga uppgifter tillgängliga.
Estland	2021	6	12	18	
Finland	-	-	-	-	Inga uppgifter tillgängliga.
Frankrike	-	-	-	-	Inga uppgifter tillgängliga.
Tyskland	-	-	-	-	Inga uppgifter tillgängliga.
Grekland	2021 till 2022	-	-	4	Könsuppdelade uppgifter inte tillgängliga.
Ungern	2021 till 2022	379	47	426	Omöjligt att skilja mellan offer för sexuella övergrepp på och utanför nätet. Endast barn under 16 år ingår.
Irland	2021 till 2022	-	-	101	Uppgifterna för 2021 (50 offer) kan inte delas upp efter kön. Uppgifter uppdelade

					efter kön för 2022: 25 flickor och 26 pojkar identifierade.
Italien	2022	-	-	385	Könsuppdelade uppgifter inte tillgängliga.
Lettland	1 augusti 2022 till 6 mars 2023	1	-	1	
Litauen	-	-	-	-	Inga uppgifter tillgängliga.
Luxemburg	2021 till 2022	0	0	0	
Malta	-	-	-	-	Inga uppgifter har lämnats/rapporterats.
Nederländerna	2021	-	-	222	Könsuppdelade uppgifter inte tillgängliga.
Polen	2022	2 368	487	3 014	År 2022 kom från Polens nationella polisinformationssystem uppgifter om 3 014 offer för sexuella övergrepp mot barn (2 368 flickor, 487 pojkar och 159 där kön inte angetts).
Portugal	-	-	-	-	Inga uppgifter har lämnats/rapporterats.
Rumänien	-	-	-	-	Inga uppgifter har lämnats/rapporterats.
Slovakien	Augusti 2021 till juli 2022	13	8	21	
Slovenien	1 januari 2021 till 14 juli 2023	220	85	305	
Spanien	2022	80	39	119	
Sverige	2022	8	4	12	
TOTALT FÖR ALLA MEDLEMSSTATE R	1 januari 2021 till 6 mars 2023	3 156	700	4 922	

Flera ytterligare förbehåll gäller för ovanstående uppgifter. I den befintliga statistiken kan inte medlemsstaterna alltid skilja mellan uppgifter om offer som identifierades utifrån en rapport från en tillhandahållare och uppgifter där till exempel offret själv eller någon annan som kände offret eller upptäckte brottet gjorde en anmälan (vilket nämndes av Slovenien). Sverige uppgav att även

barn som identifierats genom chattloggar ingår i rapporteringen, trots att foton eller videor av övergreppen aldrig har hittats eller konstaterats komma från det specifika offret.

Uppgifterna i tabell nr 8 motsvarar inte nödvändigtvis rapporteringsskyldigheterna i interimsförordningen, som endast avser offer som räddats tack vare de rapporter som lämnats av tillhandahållare och organisationer som agerar i allmänhetens intresse mot sexuella övergrepp mot barn. I de uppgifter som lämnats ingår i vissa fall offer som identifierats på grund av en rad andra orsaker och tillvägagångssätt.

Uppgifterna ger därför ingen heltäckande bild över antalet barn som identifierats som offer för sexuella övergrepp mot barn på nätet i EU.

Även om ett offer har identifierats innebär det inte nödvändigtvis att identifieringen har lett till en fällande dom. I vissa fall identifierades offret, men i utredningen gick det sedan inte att fastställa en misstänkt eller leda utredningen till en fällande dom (Sverige).

Av uppgifterna kan man dock dra slutsatsen att ett betydande antal offer har identifierats med hjälp av den frivilliga rapporteringen enligt interimsförordningen. Detta bekräftas i rapporter från brottsbekämpande myndigheter, enligt vilka många ärenden har inletts enbart på grundval av frivillig rapportering¹³.

2.2.3. Antalet dömda förövare

De flesta medlemsstater uppfyllde sina skyldigheter, men två medlemsstater lämnade inte några uppgifter eller något skäl till varför de inte lämnat rapport enligt artikel 8.1 c i interimsförordningen.

Flera medlemsstater lämnade inte någon statistik för rapporteringsperioden i enlighet med den bestämmelsen och angav som skäl att

- uppgifter ännu inte är tillgängliga (Belgien och Spanien),
- det i den centrala databas som används för att registrera brott inte krävs att den ursprungliga hänvisningens typ registreras (Irland),
- uppgifter inte har samlats in (Tyskland).

Medlemsstaterna lämnade mycket varierande uppgifter om antalet förövare som dömts och rapporteringsperioderna var inte enhetliga, vilket framgår av tabell nr 8 nedan.

Tabell nr 9: Antalet dömda förövare

Land	Rapporteringsperiod	Antal fällande domar	Kommentarer
------	---------------------	----------------------	-------------

¹³ Se t.ex. förteckningen över exempel på ärenden över hela EU som inleddes tack vare frivillig rapportering från organisationer i den [konsekvensbedömning](#) som åtföljer förslaget till förordning för att förebygga och bekämpa sexuella övergrepp mot barn (se särskilt bilaga 7).

Österrike	2021	850	Uppgifterna skiljer inte mellan brott som begåtts på och utanför nätet.
Belgien	-	-	Inga uppgifter tillgängliga.
Bulgarien	2021 till 2022	52	
Kroatien	-	-	Inga uppgifter tillgängliga.
Cypern	2022	0	Inga fällande domar hittills.
Tjeckien	1 januari 2022 till 31 juli 2022	20	
Danmark	2 augusti 2021 till 20 januari 2023	224	
Estland	2021	2	Endast fällande domar till följd av NCMEC-rapporter ingår.
Finland	2021	240	
Frankrike	4 augusti 2021 till 3 augusti 2022	820	
Tyskland	-	-	Inga uppgifter tillgängliga.
Grekland	2021 till 2022	62	
Ungern	2021 till 2022	126	
Irland	-	-	Inga uppgifter tillgängliga.
Italien	2021 till 2022	5 835	Uppgifterna skiljer inte mellan brott som begåtts på och utanför nätet.
Lettland	2021 till 2022	33	Uppgifterna skiljer inte mellan brott som begåtts på och utanför nätet.
Litauen	1 januari 2021 till 30 juni 2022	10	
Luxemburg	2022	11	Uppgifterna skiljer inte mellan brott som begåtts på och utanför nätet.
Malta	-	-	Inga uppgifter har lämnats/rapporterats.
Nederländerna	2021	217	
Polen	Andra halvåret 2021 till första halvåret 2022 / 2022 / Första halvåret 2023	185 / 194 / 81	
Portugal	-	-	Inga uppgifter har lämnats/rapporterats.
Rumänien	-	-	Inga uppgifter har lämnats/rapporterats.
Slovakien	2021	10	
Slovenien	2021 till 2022	45	
Spanien	2022	-	Inga uppgifter tillgängliga.
Sverige	2022	55	

Det är viktigt att notera att antalet fällande domar inte motsvarar antalet fällda förövare, eftersom en person kan fällas för ett eller flera brott som rör sexuella övergrepp mot barn på nätet.

Dessutom är statistiken över fällande domar som rapporterats för en viss period inte nödvändigtvis kopplad till de rapporter som har mottagits under samma period (dvs. en fällande dom under 2022 kan t.ex. vara kopplad till en rapport från 2021 eller 2020 och en rapport från 2022 kan leda till en fällande dom först under 2023 eller senare). Detta påpekades uttryckligen av flera medlemsstater i deras rapporter (Irland, Luxemburg, Sverige).

I vissa fall sammanställdes ingen statistik som visade huruvida rapporter om misstänkta handlingar (t.ex. via NCMEC) ledde till fällande domar eller med andra ord om de fällande domarna var resultatet av information från en tillhandahållare eller en offentlig organisation (Österrike, Lettland). Estland var det enda land som uttryckligen bekräftade att statistiken enbart omfattar fällande domar till följd av NCMEC:s rapporter. Det kan också vara så att rapporterna ledde till andra gärningsmän, som utreddes och dömdes i samband med utredningarna (Österrike).

I de flesta fall utgår man från att det antal fällande domar som rapporterats avser mål som avslutats efter att de sannolikt har överklagats genom rättssystemet. I en medlemsstat (Danmark), där siffrorna sammanställs enligt det senaste avgörandet, är siffrorna inte slutgiltiga, eftersom avgörandena kan ha överklagats i efterhand.

I vissa fall görs i uppgifterna från de nationella it-systemen som rapporteras av medlemsstaterna ingen skillnad mellan brott som begås på och utanför nätet (Österrike, Luxemburg, Lettland).

Eftersom medlemsstaternas rapporter är så pass varierande och insamlingen av statistik skiljer sig åt på nationell nivå går det inte att få en helhetsbild av antalet förövare som dömts för sexuella övergrepp mot barn på nätet i EU. För närvarande går det inte heller att utifrån de tillgängliga uppgifterna entydigt koppla fällande domar till rapporter som lämnats av tillhandahållare och organisationer som agerar i allmänhetens intresse mot sexuella övergrepp mot barn under specifika rapporteringsperioder enligt förordningen.

2.3. Utvecklingen av tekniska framsteg

Den teknik som för närvarande används för att upptäcka sexuella övergrepp mot barn på nätet omfattar teknik och verktyg för att upptäcka ”känt” (dvs. tidigare upptäckt) och ”nytt” (dvs. inte tidigare upptäckt) material med sexuella övergrepp mot barn samt kontaktsökning med barn (s.k. grooming).

Exemplen nedan är några av de vanligaste verktygen, men det är ingen uttömmande förteckning. Många av dessa verktyg görs tillgängliga för tillhandahållare, brottsbekämpande myndigheter och andra organisationer som har ett påvisat berättigat intresse. Verktygen kombineras oftast med mänsklig granskning för att säkerställa största möjliga noggrannhet.

Detta avsnitt omfattar även nya tekniska framsteg i fråga om artificiell intelligens.

2.3.1. Spårning av känt material med sexuella övergrepp mot barn

Befintlig teknik för att spåra känt material med sexuella övergrepp mot barn bygger enbart på automatisk analys av innehåll¹⁴ och är vanligtvis baserad på hashing. Hashingsteknik är en typ av digitalt fingeravtryck. Tekniken skapar en unik digital signatur (så kallad hash) av en bild som sedan jämförs med andra signaturer (hasher) av andra foton för att hitta kopior av samma bild. Med denna teknik upptäcks endast matchande hasher, inte material som inte matchar hashen. Hashvärden är inte heller reversibla och kan därför inte användas för att återskapa en bild.

Det finns många varianter och implementeringar av hashingsteknik. Bland de verktyg som används för att spåra känt material med sexuella övergrepp mot barn ingår i) Microsoft PhotoDNA, ii) Google CSAI Match, iii) Apple NeuralHash + Private Set Intersection, iv) Meta SSN++, v) PDQ och TMK+PDQF, vi) MD5 Hash generator (Skype) och vii) Safer (Thorn).

Det mest använda verktyget är Microsoft PhotoDNA, vilket används av över 150 organisationer¹⁵. PhotoDNA har använts i mer än tio år och har hög noggrannhet. Andelen falska positiva resultat uppskattas till högst 1 på 50 miljarder, baserat på testning¹⁶. Felfrekvensen i PhotoDNA är mycket låg på grund av teknikens natur. Med tekniken upptäcks enbart kopior av tidigare identifierat innehåll. Den ursprungliga PhotoDNA spårar känt material med sexuella övergrepp mot barn i bilder, men finns det också en version för att spåra material i videor.

Tekniken utvecklas och förbättras kontinuerligt. I maj 2023 meddelade Microsoft att nya matchningsfunktioner har införts som möjliggör snabbare sökning (cirka 350 gånger snabbare) och som samtidigt sänker kostnaderna för matchningsprocessen utan att försämra noggrannheten. Enligt Microsoft möjliggör det nya biblioteket också en mer omfattande detektering av vända eller roterade bilder. Dessutom rapporterade Internet Watch Foundation (IWF) att man nyligen förbättrat sin hashingsteknik¹⁷.

2.3.2. Spårning av nytt material med sexuella övergrepp mot barn

Den teknik som för närvarande används för att upptäcka nytt material med sexuella övergrepp mot barn omfattar klassificerare och artificiell intelligens (AI) som analyserar bilder och videor för att upptäcka innehållsmönster som matchar mönster som genererats utifrån tidigare identifierat material med sexuella övergrepp mot barn. En klassificerare är en algoritm som sorterar data i angivna klasser eller informationskategorier genom mönsterigenkänning. Klassificerare behöver data att tränas på och deras noggrannhet förbättras ju mer data de matas med.

¹⁴ Tillhandahållarna anser inte att metadata är ett effektivt verktyg för att upptäcka material med sexuella övergrepp mot barn. Se särskilt s. 10–11 i Pfefferkorn, R., Stanford Internet Observatory, ”Content-Oblivious Trust and Safety Techniques: Results from a Survey of Online Service Providers”, 9 september 2021.

¹⁵ Microsoft, [Digital Crimes Unit](#).

¹⁶ [Utlåtande av Hany Farid, PhotoDNA-utvecklare, till House Committee on Energy and Commerce, *Fostering a Healthier Internet to Protect Consumers*, 16 oktober 2019.](#)

¹⁷ Internet Watch Foundation (IWF), [The Annual Report 2022](#), s. 129–133.

Verktyg för att upptäcka nytt material med sexuella övergrepp är exempelvis i) Safer (Thorn), ii) Google Content Safety API, iii) Facebooks AI-teknik¹⁸, iv) Amazon Rekognition och v) Hive AI for visual content.

Forskning har visat att automatiserade verktyg och system som klassificerar är de mest användbara metoderna för att upptäcka material med sexuella övergrepp mot barn¹⁹. För att spåra nytt material med sexuella övergrepp mot barn ligger noggrannheten för närvarande långt över 90 %. Thorn anger till exempel att företagets klassificerare kan ställas in på en 99-procentig precision (för både känt och nytt material med sexuella övergrepp mot barn), vilket innebär 0,1 % falska positiva resultat²⁰. Dessa värden kommer sannolikt att förbättras i takt med ökad användning och återkoppling.

2.3.3. Spårning av grooming

I verktygen för att spåra grooming (kontaktsökning med barn) i textbaserad kommunikation används teknik enbart i syfte att hitta mönster som pekar på eventuella konkreta element som kan leda till misstanke om sexuella övergrepp mot barn på nätet, utan att det är möjligt att uttyda själva innehållet. Tekniken tillämpas på textbaserade chattkonversationer. Konversationerna bedöms utifrån olika egenskaper och tilldelas en övergripande sannolikhetsbedömning, som visar den uppskattade sannolikheten för att konversationen utgör grooming. Dessa bedömningar fungerar sedan som avgörande faktorer fastställda av enskilda företag för att flagga konversationer som kräver ytterligare mänsklig granskning.

Exempel på verktyg som används för textdetektering är i) Microsofts Project Artemis²¹, ii) Amazon Rekognition, iii) Twitchs Spirit AI-teknik (baserad på NLP, textklassificerare)²², iv) Metas internt konstruerade klassificerare baserad på maskininlärning och ”rangordning” (kombinerar intern språkanalysteknik med metadata), v) Roblox chattfiltrering och²³ vi) Thorn och Tech Coalitions tekniska lösning baserad på maskininlärning och klassificerare²⁴.

På samma sätt som vid identifiering av nytt material med sexuella övergrepp mot barn kräver även identifiering av grooming att tekniken tränas med sådant innehåll. Tillgången till sådana träningsdata är fortfarande den största utmaningen för att kunna utveckla och förbättra tekniken.

¹⁸ Mer information om Facebooks verktyg för att med hjälp av artificiell intelligens och maskininlärning proaktivt upptäcka nakenbilder på barn och tidigare okänt innehåll som exploaterar barn finns [här](#) och [här](#).

¹⁹ Pfefferkorn, R.: ”Content-Oblivious Trust and Safety Techniques: Results from a Survey of Online Service Providers”, *Journal of Online Trust and Safety*, februari 2022, s. 1.

²⁰ Thorn, ”[Thorn’s Automated Tool to Remove Child Abuse Content at Scale Expands to More Platforms through AWS Marketplace](#)”, 24 maj 2021.

²¹ Microsofts Project Artemis utvecklades i samarbete med The Meet Group, Roblox, Kik och Thorn.

²² Mer information finns på https://safety.twitch.tv/s/article/Our-Work-to-Combat-Online-Grooming?language=en_US.

²³ Roblox filtrerar inlägg och chattar för spelare som är tolv år eller yngre från olämpligt innehåll och för att förhindra att personlig information som hemadresser lämnas ut. Detta filtreringssystem täcker alla kommunikationsområden på Roblox, både offentliga och privata. Roblox, ”[Safety Features: Chat, Privacy & Filtering](#)”, hämtad i juli 2023.

²⁴ Tech Coalition, ”[New Technology to Help Companies Keep Young People Safe](#)”, 20 juni 2023.

Thorn har i samarbete med Tech Coalition och dess medlemmar lanserat ett nytt initiativ som syftar till att utveckla tekniska lösningar för att identifiera och hantera försök till grooming på nätet. Det är ett värdefullt initiativ som kan användas på en rad olika plattformar som erbjuder textbaserad kommunikation. Tekniken baseras på Thorns arbete med en NLP-klassificerare (*natural language processing*) eller maskininlärningsmodell, som upptäcker och kategoriserar innehåll eller beteenden på nätet som faller inom definierade ”klasser” kopplade till grooming (såsom exponering för sexuellt material eller försök till ett personligt möte med en minderårig) samt en övergripande bedömning av hur en konversation är relaterad till grooming²⁵.

2.3.4. Nya utmaningar i form av dialogrobotar och konst/bild-generatorer med artificiell intelligens

Utvecklingen och lanseringen av AI-dialogrobotar som ChatGPT (en stor språkmodell, LLM, som utvecklats av OpenAI) och konst-/bildgeneratorer som DALL-E²⁶ och Midjourney²⁷ har väckt stor uppmärksamhet, främst på grund av deras förmåga att snabbt ge färdiga svar eller skapa realistiska bilder som kan användas i många olika sammanhang. Dessa nya verktyg blev snabbt mycket populära och används flitigt. Ledande produkter finansieras och utvecklas av teknikföretag som Microsoft och Google, den nya tekniken förfinas och förbättrade versioner lanseras regelbundet.

Samtidigt som denna teknik erbjuder stora möjligheter för både företag och allmänheten kan de också utgöra en risk för dem. Farhågorna är att sådana produkter kan utnyttjas av kriminella för brottsliga syften, bland annat sexuell exploatering av barn.

Som Europol rapporterade är all information som ChatGPT tillhandahåller visserligen fritt tillgänglig på internet, men det gör det betydligt enklare för illvilliga aktörer ”att utan förkunskaper lära sig mer om ett stort antal potentiella brott, allt från hur man bryter sig in i ett hem till terrorism, cyberbrott och sexuella övergrepp mot barn”. På så sätt kan dessa aktörer bättre förstå och därefter utföra dessa typer av brott²⁸.

Enligt OpenAI:s regler får ChatGPT inte svara på frågor om sexuellt, hatiskt och våldsamt innehåll eller innehåll som främjar självskadebeteende. Men dessa skyddsåtgärder kan kringgåås ganska enkelt genom promptkonstruktion²⁹. Den senaste utvecklingen av AI-dialogrobotar (t.ex. av Snapchat) visar hur dessa kan överskrida gränsen till stötande eller farliga interaktioner,

²⁵ Tech Coalition, [”New Technology to Help Companies Keep Young People Safe”](#), 20 juni 2023.

²⁶ DALL-E är ett AI-system som kan skapa realistiska bilder och konst utifrån en beskrivning i naturligt språk.

²⁷ Midjourney är ett program och en tjänst för generativ artificiell intelligens som genererar bilder utifrån beskrivningar i naturligt språk.

²⁸ Europol, [ChatGPT – The impact of Large Language Models on Law Enforcement](#), 2023, ISBN 978-92-95220-57-7, s. 7.

²⁹ Swanson, S.M., [”ChatGPT Generated Child Sex Abuse When Asked to Write BDSM Scenarios”](#), *Vice*, 6 mars 2023; Mitchell, A., [”ChatGPT gives sick child sex abuse answer, breaking its rules”](#), *New York Post*, 24 juli 2023; Europol, [ChatGPT – The impact of Large Language Models on Law Enforcement](#), 2023, ISBN 978-92-95220-57-7, s. 5.

däribland sexuella övergrepp mot barn³⁰. Nu när fler företag överväger att testa AI-dialogrobotar på sina plattformar (Instagram, eventuellt WhatsApp och Messenger) krävs noggranna analyser av hur användarna påverkas, särskilt barn och ungdomar.

För dessa nya verktyg krävs också lämpliga skyddsåtgärder så att de inte missbrukas för att producera AI-genererat ”deep fake”-material med sexuella övergrepp mot barn³¹. Med tanke på AI-verktygens snabba utveckling blir det sannolikt snart betydligt enklare att generera bilder som inte går att skilja från verkliga bilder. Om verklighetstroget datorgenererat material blir vanligt på nätet uppstår flera viktiga utmaningar i kampen mot sexuella övergrepp mot barn, eftersom polisen då inte bara får det svårt att utreda och lagföra sådana brott, utan också att identifiera de verkliga offren³².

Forskning har visat att tillgång till material om sexuella övergrepp mot barn ofta är första steget mot att begå övergrepp i realiteten, oavsett om materialet skildrar övergrepp och utnyttjande som är verkliga eller som har genererats på ett verklighetstroget sätt³³. Att begränsa spridningen av AI-genererat, manipulerat material om sexuella övergrepp mot barn är därför ytterst viktigt som förebyggande åtgärder mot förövare. En annan farhåga är att personer som utsätter barn för grooming kan använda ChatGPT:s avancerade textgenereringsfunktioner i kombination med befintliga kostnadsfria generatorer för text-till-bild-AI för att snabbt och enkelt skapa innehåll till falska profiler och trovärdiga konversationer för att rikta in sig på barn på nätet. Även om ChatGPT i sig inte uppmuntrar människor till grooming på nätet kan vem som helst mata konversationerna de har med barn på nätet genom AI-teknik för att göra sig mer övertygande och trovärdiga för sina offer, vilket underlättar manipulation.³⁴ Generativ AI skulle kunna öka antalet fall av grooming på nätet och till och med ”automatisera grooming av barn i stor skala”³⁵.

3. SLUTSATSER

Genomförandeåtgärder som har vidtagits av tillhandahållare

Tillhandahållarnas rapportering visade att de med hjälp av en mängd teknik och processer för detektering har spårat och rapporterat sexuella övergrepp mot barn på nätet enligt interimsförordningen. Samtliga tillhandahållare uppgav att de skickat dessa rapporter till

³⁰ Fowler, G.A., ”Snapchat tried to make a safe AI. It chats with me about booze and sex”, *The Washington Post*, 14 mars 2023; Vincent, J., ”Instagram is apparently testing an AI chatbot that lets you choose from 30 personalities”, *The Verge*, 7 juli 2023.

³¹ Crawford, A., Smith, T., ”Illegal trade in AI child sex abuse images exposed”, BBC, 27 juni 2023.

³² Thiel, D., Stroebel, M., och Portnoff, R. (2023). *Generative ML and CSAM: Implications and Mitigations*. Stanford Digital Repository. Tillgänglig på <https://purl.stanford.edu/jv206yg3793>. <https://doi.org/10.25740/jv206yg3793>. s. 2.

³³ Protect Children, ”Protect Children’s research in the dark web is revealing unprecedented data on CSAM users”, 6 juni 2021; RAINN, ”What is Child Sexual Abuse Material (CSAM)”, 25 augusti 2022.

³⁴ Breck Foundation, ”Is artificial intelligence putting children at risk?”, 9 februari 2023, uppdaterad 3 april 2023.

³⁵ Butler, J., ”AI tools could be used by predators to ‘automate child grooming’, eSafety commissioner warns”, *The Guardian*, 19 maj 2023.

NCMEC. När det gäller vilken typ och volym av personuppgifter som tillhandahållarna hanterat visade rapporterna en mängd insamlade trafikuppgifter och en varierad detaljnivå på de behandlade datavolymer. Kommissionen har därför inte tillgång till enhetliga uppgifter om tillhandahållarna på EU-nivå för den aktuella rapporteringsperioden (juli 2021 till 31 januari 2023).

Tillhandahållarna lämnade inte in uppgifter om antal och andel fel (falska positiva resultat) uppdelat på olika slags teknik, vilket tyder på att de använder en skiktad metod för att upptäcka sexuella övergrepp mot barn på nätet i kombination med mänsklig granskning. Men samtidigt har de infört ett antal olika skyddsåtgärder för att begränsa och minska felprocenten i spårningen. Tillhandahållarna uppgav också att de har riktlinjer för datalagring och dataskyddsåtgärder fastställda i sina integritetspolicier eller integritetsförklaringar, vilka stöds av standardiserade dataskyddsåtgärder för branschen.

Genomförandeåtgärder som har vidtagits av medlemsstaterna

Enligt interimsförordningen (artikel 8) är medlemsstaterna skyldiga att lämna in viktig statistik om fall av sexuella övergrepp mot barn på nätet som har upptäckts och rapporterats till deras brottsbekämpande myndigheter, men också antalet barn som har identifierats som offer och antal dömda förövare. Eftersom medlemsstaterna i de flesta fall lämnade uppgifter för olika rapporteringsperioder gick det inte att dra jämförbara slutsatser om det totala antalet rapporter om upptäckta sexuella övergrepp mot barn på nätet som mottagits på EU-nivå. Dessutom är de rapporter som medlemsstaterna tar emot och rapporterar inte nödvändigtvis åtgärdbara rapporter, dvs. rapporter som skulle kunna användas för utredningar, eller motsvarar inte antal rapporterade fall. Endast ett fåtal medlemsstater har angett den typ av tillhandahållare på vars tjänst de sexuella övergreppen mot barn på nätet upptäcktes. I vissa fall görs i den nationella statistiken ingen skillnad mellan brott som utretts utifrån rapporter från tillhandahållare och andra organisationer som agerar i allmänhetens intresse mot sexuella övergrepp mot barn och brott som utretts utifrån andra rapporter.

Det gick inte heller att ur de mottagna rapporterna utläsa uppdelningen per kön av det totala antalet barn som identifierats som offer för sexuella övergrepp mot barn på nätet i EU. Skälen är exempelvis att de rapporterade uppgifterna gäller för olika perioder, olika åldersgränser används för att definiera barn som offer för sexuella övergrepp på nätet, statistik på denna detaljnivå inte samlas in på nationell nivå på grund av tekniska eller andra begränsningar, ingen skillnad görs mellan barn som fallit offer för sexuella övergrepp på och utanför nätet osv. Vissa medlemsstater tar även med barn som har genererat eget material i sin statistik. Ännu viktigare är att statistiken ofta inte skiljer mellan offer som identifierats utifrån rapporter från tillhandahållare och organisationer som agerar i allmänhetens intresse mot sexuella övergrepp mot barn enligt förordningen och de som identifierats utifrån andra skäl och sätt.

Uppgifterna om antalet förövare som har dömts är också bristfälliga. Det kan ibland vara så att uppgifter inte är tillgängliga eftersom källan till den ursprungliga hänvisningen inte är registrerad i databaserna och det därför inte görs skillnad mellan gärningsmän som har dömts på grund av rapporter som lämnats in enligt förordningen och andra rapporter. I vissa fall skiljer inte heller

uppgifterna i de nationella it-systemen mellan brott som begås på och utanför nätet. Statistiken över fällande domar som rapporterats för en viss period är inte heller alltid kopplad till de rapporter som har mottagits under den aktuella perioden, utan kan avse rapporter från tidigare perioder. Insamlad statistik om antalet fällande domar kan också skilja sig åt från antalet dömda förövare (eftersom en förövare kan ha fler fällande domar).

På grund av den varierande statistik som lämnats av medlemsstaterna, som inte alltid verkar samla in uppgifterna på ett systematiskt och korrekt sätt, och alla andra ovannämnda faktorer går det inte att få en heltäckande bild av de rapporter som mottagits om sexuella övergrepp mot barn på nätet, antalet barn som identifierats som offer för dessa brott eller antalet förövare som dömts på EU-nivå i enlighet med förordningen. Det faktum att det för de flesta medlemsstater verkar vara en betydande skillnad mellan det antal rapporter som NCMEC uppger sig ha skickat till medlemsstaten och det antal rapporter som medlemsstaten uppger att den har tagit emot tyder på att datainsamlingen och rapporteringen inte är heltäckande. Vissa medlemsstater bekräftade att deras behöriga myndigheter genomgår strukturella förändringar eller omorganisationer för att inrätta nya avdelningar med ansvar för att utreda brott som rör sexuella övergrepp mot barn på nätet. Nya it-system håller också på att införas och de nationella myndigheterna har ombetts ändra sina registreringsförfaranden och sin statistik i vissa medlemsstater. Detta bör skapa bättre förutsättningar för mer korrekt statistik från medlemsstaterna i framtiden. Kommissionen kommer under alla omständigheter att utnyttja sina befogenheter enligt fördragen för att se till att medlemsstaterna fullgör sina rapporteringsskyldigheter enligt interimsförordningen.

Allmänt

Sammantaget visar denna rapport på betydande skillnader i rapporteringen av uppgifter om bekämpning av sexuella övergrepp mot barn på nätet enligt interimsförordningen från både tillhandahållare och medlemsstater. En ökad standardisering av tillgängliga uppgifter och rapporteringen av dem, såsom i förslaget till förordning om förebyggande och bekämpande av sexuella övergrepp mot barn³⁶, skulle bidra till en bättre bild av relevanta åtgärder för att bekämpa dessa brott. Det verkar också som att det krävs ytterligare insatser från tillhandahållarnas och medlemsstaternas sida för att se till att uppgifter samlas in och rapporteras enligt kraven i interimsförordningen.

Tillgängliga uppgifter visar att det inom ramen för det nuvarande frivilliga systemet för spårning och rapportering kan vara så att material som automatiskt flaggas som möjliga sexuella övergrepp mot barn vid en mänsklig granskning sedan visar sig inte vara det. Det kan bero på att det saknas en gemensam uppsättning hashkoder och andra indikatorer för att upptäcka sådant material som bekräftats vara olagligt i EU, eller bero på olika rättsliga standarder i olika jurisdiktioner, i synnerhet mellan EU och Förenta staterna när det gäller de relevanta definitionerna. Uppgifterna pekar också på stora variationer i antalet begäranden om prövning och i utfallet av dessa prövningar. Här går det inte att dra några slutsatser, eftersom det saknas information om framför allt antalet begärda prövningar och skälen till återställandet.

³⁶ Förslag till Europaparlamentets och rådets förordning om fastställande av regler för att förebygga och bekämpa sexuella övergrepp mot barn, [COM/2022/209 final](#).

När det gäller kraven enligt artikel 9.2 om villkoren för behandlingen visar den information som lämnats att den använda tekniken motsvarar tekniska tillämpningar som har utformats endast i syfte att upptäcka och avlägsna material om sexuella övergrepp mot barn på nätet och rapportera det till brottsbekämpande myndigheter och till organisationer som agerar i allmänhetens intresse mot dessa övergrepp. Det lämnades ingen information om i vilken utsträckning de använda verktygen överensstämde med den senaste tekniken och användes på ett sätt som inkräktar på privatlivet så lite som möjligt och inte heller om en konsekvensbedömning avseende dataskydd enligt artikel 35 i förordning (EU) 2016/679 och ett förhandssamråd enligt artikel 36 i samma förordning hade genomförts i förväg.

När det gäller förordningens proportionalitet är den centrala frågan i vilken utsträckning interimsförordningen uppnår den jämvikt som unionslagstiftaren eftersträvade mellan å ena sidan målet av allmänt intresse att effektivt bekämpa de mycket allvarliga brott som det är fråga om och behovet av att skydda barns grundläggande rättigheter (värdighet, integritet, förbud mot omänsklig eller förnedrande behandling, privatliv, barnets rättigheter) och å andra sidan skyddet av de grundläggande rättigheterna för användare av de tjänster som omfattas (privatliv, skydd av personuppgifter, yttrandefrihet, effektiva rättsmedel). De tillgängliga uppgifterna är för knapphändiga för att man här ska kunna dra några definitiva slutsatser. Det skulle varken vara möjligt eller lämpligt att tillämpa en numerisk standard vid bedömningen av proportionaliteten i fråga om antalet räddade barn med tanke på de betydande negativa effekter som sexuella övergrepp har på ett barns liv och rättigheter. Mot bakgrund av ovanstående finns det dock inget som tyder på att undantaget inte skulle vara proportionerligt.

Trots att de tillgängliga uppgifterna är bristfälliga och inte ger någon inblick i användningen av frivilliga rapporter i ett stort antal medlemsstater visar de tillgängliga uppgifterna tydligt att tusentals barn har identifierats under rapporteringsperioden, mer än två tusen fällande domar har meddelats och miljontals bilder och videor har hindrats från att spridas, vilket har lett till en minskning av sekundär viktimisering. Man kan därför dra slutsatsen att frivillig rapportering i hög grad bidrog till att skydda ett stort antal barn, även från pågående övergrepp, och det verkar som om interimsförordningen är effektiv.