



DICHTERBIJ DAN JE DENKT

Een verkennend onderzoek naar risicovol
zoekgedrag op pornowebsites

INHOUDSOPGAVE

I.	SAMENVATTING	03
2.	INLEIDING	05
	technologie om CSAM te detecteren	07
	huidig onderzoek	10
	onderzoeksvragen	10
3.	METHODE	11
	procedure	12
	risicotermen	12
	risicoscore	13
	zoekprofielen	13
	steekproef	15
4.	RESULTATEN	17
	pornowebsites	17
	zoekgedrag - categorieën & zoektermen	18
	zoekgeschiedenis	20
	beschrijving van de profielen	21
5.	DISCUSSIE	22
	aanbevelingen	25
6.	REFERENTIES	29
7.	BEGRIPPENLIJST	30



SAMENVATTING

Doel van het onderzoek

Veel beelden van seksueel misbruik van minderjarigen zijn op het openbare internet te vinden. De kans bestaat dat mensen hier, bedoeld en onbedoeld, mee in aanraking komen. Kijkers van misbruikbeelden zijn de eerste keer vaak niet doelbewust op zoek naar dergelijk materiaal, maar zijn in hun zoektocht naar porno wel bij misbruikbeelden uitgekomen en het vervolgens blijven kijken. Het is belangrijk om te voorkomen dat mensen tijdens het porno kijken terechtkomen bij dergelijke beelden. Om preventie effectief in te zetten, is meer kennis nodig over het zoekgedrag op pornowebsites wat mogelijk kan leiden tot strafbaar beeldmateriaal. Deze kennis kan helpen om vroegtijdig in te grijpen en interventies te ontwikkelen met als doel het voorkomen dat pornogebruikers in aanraking komen met misbruikbeelden.

Methode

Dit exploratieve onderzoek maakt gebruik van zoekprofielen die het gedrag van verschillende gebruikers op een pornowebsite nabootsen. Deze zoekprofielen gebruiken verschillende pornowebsites en categorieën of zoektermen als startpunt. Een voorbeeld is een zoekprofiel dat begint op een pornowebsite, klikt op de categorie 'teens' en vanuit daar verder klikt. Vervolgens is gekeken of het zoekprofiel terechtkomt bij twijfelachtig materiaal of misbruikbeelden. Twijfelachtig materiaal houdt in dat het mogelijk gaat om een minderjarige maar dat het op basis van de lichamelijke kenmerken niet geconcludeerd kan worden.

Belangrijkste bevindingen

De resultaten laten zien dat het risico om terecht te komen bij twijfelachtig materiaal of misbruikbeelden vooral afhangt van hoe de pornowebsite is opgebouwd, wat samen lijkt te hangen met de connecties naar andere pornowebsites. De pornowebsite die veel pop-ups vertoonde, liet een groot netwerk zien van aan elkaar gelinkte pornowebsites. Dit heeft geleid tot verschillende websites waarop twijfelachtig materiaal en misbruikbeelden is gevonden. Dit was ongeacht de specifieke pornocategorie waarop werd gezocht. Zowel het zoeken op jong (bijv. 'teens'), een extreme pornocategorie (bijv. 'brutal'), of een controle categorie (bijv. 'blonde') zorgde ervoor dat

er twijfelachtig materiaal of misbruikbeelden werd gevonden.

Conclusie & aanbevelingen

Het onderzoek laat zien dat het doorklikken op pornowebsites (middels advertenties en pop-ups) kan leiden tot risicovolle websites met twijfelachtig en strafbaar materiaal. Daarom is het belangrijk dat er goede regulatie komt op pornowebsites, waarbij pornowebsites de verantwoordelijkheid hebben om maatregelen te treffen die voorkomen dat er misbruikbeelden op hun platform staat of dat er gelinkt wordt naar dergelijk materiaal, en waarbij de overheid zorgt voor gepast beleid dat hier op toeziet. Omdat het toezicht verdeeld is over verschillende autoriteiten is het essentieel een gezamenlijke aanpak te ontwikkelen. Daarnaast is goede voorlichting belangrijk zodat mensen bewuste keuzes kunnen maken over hun kijkgedrag waardoor het risico dat mensen hier onbedoeld mee in aanraking komen, verminderd kan worden. De maatschappij heeft dan ook een gedeelde verantwoordelijkheid om het bekijken en verspreiden van misbruikbeelden tegen te gaan, zodat kinderen beschermd kunnen worden en (herhaald) slachtofferschap wordt voorkomen.

02

INLEIDING

De Nederlandse Politie ontving in 2022 meer dan 55.000 meldingen over personen die kijken naar beeldmateriaal van seksueel misbruik van minderjarigen.¹ Doordat veel misbruikbeelden op het Darkweb (een anoniem en verborgen gedeelte van het internet) te vinden zijn, is het aantal kijkers van misbruikbeelden vermoedelijk nog hoger. Het kijken van beeldmateriaal van seksueel misbruik van minderjarigen is strafbaar. Het zorgt ervoor dat het aanbod blijft bestaan en draagt daarmee bij aan (herhaald) slachtofferschap.

Er worden niet alleen nieuwe beelden vervaardigd maar bestaand materiaal blijft ook continu in omloop en is moeilijk offline te halen. Hierdoor worden minderjarigen steeds opnieuw slachtoffer van het misbruik. Dit kan zorgen voor een grote psychische last en kan gevolgen hebben voor het welzijn en functioneren van de minderjarige; met mogelijk levenslang traumatische gevolgen. Het is dan ook een maatschappelijk verantwoordelijkheid minderjarigen te beschermen tegen seksueel geweld in alle vormen. Ook (inter-)nationale verdragen zoals het VN- Kinderrechtenverdrag verplichten de Nederlandse staat hiertoe. Een preventieve aanpak die zich richt op de kijkers van misbruikbeelden is, naast goede hulp voor slachtoffers (en daders), een essentieel onderdeel van de integrale aanpak. Om preventie effectief in te zetten is meer kennis nodig over deze doelgroep en hoe zij terechtkomen bij misbruikbeelden.

1 Nationaal Rapporteur Mensenhandel en Seksueel Geweld tegen Kinderen, 2023

STOP IT NOW

Stop it Now zet zich in voor de preventie van seksueel kindermisbruik. Door middel van een anonieme en gratis hulplijn biedt Stop it Now advies, ondersteuning en hulp aan personen die zich zorgen maken over hun seksuele gevoelens en/of gedrag richting minderjarigen. Ook naasten en professionals kunnen contact opnemen met Stop it Now. Kijk voor meer informatie op www.stopitnow.nl.

Misbruikbeelden op het openbare web

Een veelvoorkomende misvatting is dat je actief op zoek moet gaan naar beeldmateriaal van seksueel misbruik van minderjarigen en dat dit uitsluitend gebeurt op het Darkweb. Echter, veel misbruikbeelden zijn te vinden op het openbare web (het normale, toegankelijke deel van het internet), zoals op pornowebsites, social media of forums. 39% van de personen die contact opneemt met Stop it Now in verband met zorgen over hun gedrag, zijn personen die kijken naar beeldmateriaal van seksueel misbruik van minderjarigen.² Een groot deel hiervan vertelt door het doorklikken op advertenties op strafbaar beeldmateriaal te zijn gestuit, terwijl ze oorspronkelijk op pornowebsites zochten naar categorieën zoals 'jong' of 'tieners',

Escaleren op pornowebsites

De helft van de personen die contact opneemt met Stop it Now met zorgen over het kijken naar misbruikbeelden zijn jonge mannen onder de 26 jaar.³ Het onderzoek naar jonge (potentiële) plegers van (online) seksueel kindermisbruik die contact opnemen met Stop it Now laat zien dat er bij veel jonge kijkers van misbruikbeelden sprake is van risicovol porno kijkgedrag of een escalatie hierin.⁴ Door herhaalde blootstelling aan pornografische beelden gaan deze personen op zoek naar steeds extremer beeldmateriaal om dezelfde mate van opwindning en spanning te ervaren. Hierdoor kunnen sommige personen escaleren naar beeldmateriaal van seksueel kindermisbruik. Deze personen zijn dan niet doelbewust op zoek geweest naar dit materiaal, maar zijn er door hun kijkgedrag wel mee in aanraking gekomen. Na deze eerste confrontatie zijn ze het materiaal blijven kijken.

Zoekgedrag op pornowebsites

Gezien de omvang en toegankelijkheid van porno op het internet is het zorgelijk dat personen via pornowebsites terecht kunnen komen bij misbruikbeelden. Er zijn duizenden pornowebsites online te vinden waar je niet voor hoeft te betalen. Het verdienmodel van zo'n website is vaak gebaseerd op advertenties. Wanneer iemand zoekt naar porno in bepaalde categorieën of specifieke zoektermen gebruikt, wordt deze informatie opgeslagen in cookies. Een algoritme kan deze informatie gebruiken om advertenties te tonen op basis van iemand zijn zoekgeschiedenis. Het is onbekend of een zoekgeschiedenis bestaande uit categorieën zoals 'teens' de kans vergroot dat er advertenties worden getoond die linken naar strafbaar materiaal.

2 Cijfers Stop it Now 2023

3 Cijfers Stop it Now 2023

4 Nationaal Rapporteur Mensenhandel en Seksueel Geweld tegen Kinderen & Stop it Now, 2023

In dit verkennend onderzoek worden zoekprofielen gebruikt die het gedrag van gebruikers op een pornowebsite simuleren. De zoekprofielen gebruiken verschillende pornowebsites en categorieën of zoektermen als startpunt. Vervolgens wordt onderzocht of het zoekprofiel terechtkomt bij risicovolle websites die twijfelachtig en mogelijk ook strafbaar beeldmateriaal bevatten. Met deze kennis kunnen gerichte interventies ontwikkeld en ingezet worden die zich richten op pornogebruikers voordat ze in contact komen met beeldmateriaal van seksueel misbruik van minderjarigen. Op deze manier kan vroegtijdig worden ingegrepen en voorkomen worden dat personen terechtkomen bij strafbaar materiaal.

TERMINOLOGIE

Hoewel er in de volksmond wordt gesproken over kinderporno, gebruikt Stop it Now de term 'beeldmateriaal van seksueel misbruik van minderjarigen', of verkort 'misbruikbeelden'. Porno heeft als doel seksuele opwindning te creëren en is niet strafbaar. Kinderporno is per definitie seksueel kindermisbruik en wel strafbaar. De term kinderporno dekt dus niet de ernst van het probleem. In Engelstalige literatuur wordt de vergelijkbare term en afkorting 'CSAM' gebruikt, wat staat voor 'child sexual abuse material'. Voor de leesbaarheid van het rapport zal deze afkorting worden gebruikt.

TECHNOLOGIE OM CSAM TE DETECTEREN

Om te onderzoeken welke zoekprofielen mogelijk leiden tot risicovolle websites wordt er in dit verkennend onderzoek gebruik gemaakt van technologie om internetpagina's te scannen. Deze technologie kan inzichtelijk maken hoe websites aan elkaar gelinkt zijn. In wetenschappelijk onderzoek is gebruik gemaakt van web crawler technologie om CSAM netwerken in kaart te brengen. Een web crawler begint bij een bekende website met CSAM en volgt de hyperlinks op een pagina en haalt vrijwel alle informatie van de pagina op.

Westlake, Bouchard & Frank (2011) onderzochten verschillende websites, waaronder blogs, waarvan bekend was dat er CSAM op staat. Bij het analyseren van de pagina's werd op basis van key words bepaald of de pagina CSAM bevat. Key words zijn woorden die geassocieerd zijn met CSAM en de web crawler detecteert of deze woorden voorkomen op een pagina. Key words waren onderverdeeld in code woorden (waarvan bekend is bij de politie dat CSAM-kijkers ze gebruiken om te zoeken naar strafbaar materiaal), softcore (bijv. 'young'), en hardcore (bijv. 'sex'). Ze vonden dat websites met CSAM een veel hogere frequentie van deze hardcore key words bevatten in vergelijking met controle websites.⁵ Al is het onduidelijk in hoeverre deze woorden in staat zijn om CSAM websites te onderscheiden van websites met pornografie.

Combinatie van hash waarden en key words meest effectief

Westlake, Bouchard, & Frank (2012) bouwen voort op voorgaand onderzoek door naast key words

5 Westlake, Bouchard & Frank, 2011

ook te kijken naar de afbeeldingen op de pagina's. Hiervoor hebben ze verschillende methoden onderzocht om CSAM te detecteren en websites met (mogelijk) strafbaar materiaal in kaart te brengen. Hiervoor hebben ze web crawler technologie gebruikt gebaseerd op:

1. Enkel hash values
2. Enkel key words
3. Hybride vorm met hash values én key words

Enkel het gebruik van key words maakt het niet mogelijk om de afbeeldingen te analyseren op inhoud en om te verifiëren of het gaat over misbruik van minderjarigen. Daarom wordt er gebruik gemaakt van hashes. Een hash is een code die uniek is voor een afbeelding. Je kunt het zien als een digitale vingerafdruk. Een hash database bevat hashes van bekend CSAM. Door middel van een exacte match met een hash in de database kan bekend CSAM gedetecteerd worden. Om te onderzoeken hoeveel key words nodig waren om een website met CSAM te vinden, zijn er testen gedaan met de web crawler. Er werd gevonden dat 7 key words voldoende betrouwbaar waren om bij CSAM uit te komen, waarbij er zo min mogelijk false positives⁶ waren. Het gebruik van key words kan leiden naar websites met nieuw materiaal maar heeft tegelijk het risico om veel false positives op te leveren. De key words kunnen namelijk ook verwijzen naar websites die geen CSAM bevatten. Key words waren onderverdeeld in drie categorieën. Categorie 1: code woorden die gebruikt worden door kijkers van CSAM om strafbaar materiaal te zoeken. Categorie 2: meer algemene woorden die ook gevonden worden op websites die geen CSAM bevatten, zoals 'boy' of 'girl'. Categorie 3: woorden met een expliciete seksuele betekenis, zoals 'sex' of 'anal'.

De resultaten van Westlake, Bouchard, & Frank (2012) laten zien dat enkel het gebruik van hash values niet voldoende werkte, omdat deze afbeeldingen niet veel voorkwamen op openbare websites. Websites met code woorden die gebruikt worden door kijkers van CSAM om strafbaar materiaal te zoeken (categorie 1) kwamen vaker voor. Op jongens gerichte key words werden vaker gevonden op websites met CSAM dan op meisjes gerichte key words (al was dit verschil niet significant). De resultaten toonden aan dat de hybride vorm beter in staat was om websites te detecteren die CSAM bevatten.⁷

Betrouwbaarheid web crawler technologie

Pornowebsites bevatten veel key words die ook te vinden zijn op websites met CSAM. Hoe goed is web crawler technologie in staat dit onderscheid te maken? Westlake, Bouchard, & Frank (2017) onderzochten de validiteit van web crawlers door websites met CSAM te vergelijken met seks gerelateerde (seksuele voorlichting of pornografie) en seks ongerelateerde (sport) websites. Specifieke code woorden, die gebruikt worden door kijkers van CSAM om strafbaar materiaal te zoeken, werden in dezelfde mate gevonden op de verschillende typen websites al kwamen ze vrij weinig voor. Er werd ook gekeken naar meer thematische woorden niet direct gerelateerd aan CSAM (boy, girl), seks-georiënteerde woorden (pussy, oral), expliciete woorden gerelateerd aan seksueel misbruik (torture, rape) en niet-expliciete woorden gerelateerd aan persoonlijke eigenschappen (innocent, smooth). Thematische, seks-georiënteerde, expliciete en niet-expliciete

6 Een onterechte positieve uitkomst waarbij er geen CSAM op de website voorkomt.

7 Westlake, Bouchard, & Frank, 2012

woorden kwamen in veel hogere frequentie voor op de websites met CSAM in vergelijking met de andere websites. Op websites met pornografie komen seks-georiënteerde en expliciete woorden wel voor maar in mindere mate dan op CSAM websites. Daarnaast hadden CSAM websites meer afbeeldingen per pagina in vergelijking met de seks gerelateerde websites. Specifiek de websites met pornografie hadden gemiddeld meer video's, terwijl websites met CSAM dus vaker afbeeldingen bevatten. Afbeeldingen van de hash database werden enkel minimaal gevonden op de seks gerelateerde websites.

Het gebruik van hash values en key words bleek dus effectief in het onderscheiden van de verschillende typen websites.⁸ Westlake, Bouchard, & Frank (2017) suggereren dat key words betrouwbaar zijn maar dat de constante ontwikkeling op dit gebied (nieuwe woorden die gebruikt worden) de validiteit beperkt. Daarom is het belangrijk dat key words goed up-to-date zijn en regelmatig aangevuld worden met de meest gebruikte termen.

Daarnaast is het belangrijk dat een set van gerelateerde key words gecombineerd wordt (dus ook thematische en seks-georiënteerde woorden).

Key words

Onderzoek met web crawlers laat zien dat een combinatie van hash waarden en key words effectief is om CSAM websites te detecteren. Het is hierbij wel belangrijk om key words te gebruiken die mogelijk kunnen wijzen op CSAM. Een onderzoek naar bestandsnamen van CSAM op het openbare web laat zien dat websites die het materiaal weergeven in de meeste gevallen expliciete woorden gebruiken om de inhoud van het materiaal aan te geven.⁹ Dus foldernamen bevatten vaak woorden geassocieerd met kinderen (bijv. boy of girl), lichaamskenmerken (bijv. naked), en/of seksuele handelingen (bijv. sex).

Daarnaast is het belangrijk om te weten hoe personen naar deze bestanden zoeken. Om meer zicht te krijgen op key words die gebruikt worden om naar misbruikbeelden te zoeken, heeft Steel (2009) gekeken naar queries¹⁰ op een peer-to-peer netwerk. Er werd gevonden dat key words die niet direct gerelateerd zijn aan seksueel misbruik wel gebruikt worden om te zoeken naar CSAM.¹¹ Voorbeelden zijn 'child' en 'teen'. Als iemand zoekt op de term 'teen' wil dat echter niet per definitie zeggen dat ze op zoek zijn naar strafbaar beeldmateriaal. De onderzoekers hebben voor drie populaire woorden (underage, young, teen) onderzocht welke woorden hiermee vaak samen voorkomen. De term 'young' lijkt in dit opzicht wat ambigu omdat het vaak voorkomt met woorden zoals 'love' en 'fight'. De termen 'teen' en 'underage' komen daarentegen vaker samen voor met woorden zoals 'rape', 'pedo' of 'preteen'. De combinatie van deze woorden suggereert wel strafbaar materiaal en kunnen dus indicatoren zijn dat iemand op zoek is naar CSAM.

8 Westlake, Bouchard, & Frank, 2017

9 Guerra & Westlake, 2021

10 Een query is een zoekopdracht, meestal een woord of een combinatie van woorden, om specifieke informatie te zoeken.

11 Steel, 2009

HUIDIG ONDERZOEK

Onderzoeken die CSAM netwerken in kaart brengen maken gebruik van websites als beginpunt waarvan bekend is dat er CSAM op staat. Maar deze onderzoeken geven geen informatie over hoe iemand voor het eerst terechtkomt bij een website die CSAM bevat. In het huidige onderzoek worden daarom websites met pornografie als startpunt gebruikt om te onderzoeken hoe mensen terechtkomen bij risicovolle websites met mogelijk strafbaar materiaal. Om dit inzichtelijk te maken is het belangrijk om ook het grensgebied te onderzoeken; dus wanneer beelden twijfelachtig zijn. Twijfelachtig materiaal betekent dat het mogelijk wel gaat om een minderjarige, maar dat het op basis van de lichamelijke kenmerken niet met honderd procent zekerheid te concluderen is dat het een minderjarige is. Onderzoek heeft aangetoond dat een combinatie van hash waarden en key words het meest effectief is om CSAM websites te detecteren. Daarom wordt in het huidige onderzoek ook gebruik gemaakt van zowel hash waarden als key words. Er zal hiervoor een combinatie van verschillende key words worden gebruikt die mogelijk kunnen wijzen op CSAM.

ONDERZOEKSVRAGEN

Het doel van dit onderzoek is om inzicht te krijgen in het zoekgedrag op pornowebsites dat kan leiden tot twijfelachtig materiaal en/of CSAM. Meer kennis over het zoekgedrag en de paden vanuit pornowebsites naar CSAM kan helpen om vroegtijdig in te begrijpen en interventies te ontwikkelen met als doel het voorkomen dat pornogebruikers in aanraking komen met CSAM. Dit heeft geleid tot de volgende onderzoeksvraag:

Welk zoekgedrag op pornowebsites kan leiden tot twijfelachtig materiaal en CSAM?

Om deze vraag te kunnen beantwoorden, is er gekeken naar verschillende factoren die hier mogelijk bij betrokken zijn, zoals het type website, categorieën en zoektermen, en een zoekgeschiedenis. Om dit in kaart te brengen worden de volgende deelvragen onderzocht:

- 1. Is er een verschil tussen pornowebsites in de kans om bij twijfelachtig materiaal en CSAM terecht te komen?*
- 2. Is er een verschil tussen categorieën en zoektermen in de kans om bij twijfelachtig materiaal en CSAM terecht te komen?*
- 3. Beïnvloedt de zoekgeschiedenis van een pornogebruiker de kans om bij twijfelachtig materiaal en CSAM terecht te komen?*

03

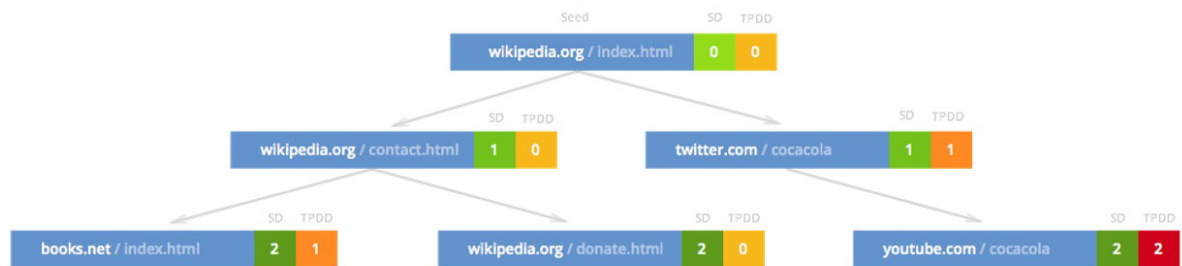
METHODE

Voor dit verkennend onderzoek heeft Offlimits samengewerkt met Web-IQ, die een methode heeft ontwikkeld om het “digitale pad” van pornowebsites naar twijfelachtig materiaal en CSAM in kaart te brengen door gericht websites te scannen. Dit is een techniek waarbij in tegenstelling tot webcrawling slechts specifiek benodigde data (HTML data¹ en hashes) van het internet wordt binnengehaald en opgeslagen om te worden geanalyseerd. Data die niet nodig is, wordt dus niet opgehaald. Indien er een afbeelding op de pagina staat, wordt hiervan een hash gemaakt. De afbeelding zelf wordt niet opgehaald of opgeslagen. Hiermee voldoet de technologie aan de AVG.² Waarbij de juridische basis, bewaartermijn en veiligheids- en toegangseisen streng worden nageleefd.

Het beginpunt van de scan is een URL van een websitepagina, en van elke pagina die wordt gescand worden naast specifieke benodigde data ook de hyperlinks opgehaald. Deze hyperlinks leiden naar een volgende pagina waarvan specifieke data vervolgens ook wordt gescand en geanalyseerd. Hierdoor ontstaat er een netwerk van aan elkaar gelinkte internetpagina's. Om te voorkomen dat dit netwerk te groot wordt, is er een seed distance (SD) gebruikt van 4. De seed distance is het aantal links dat gevolgd moet worden. Na 4 links wordt het scannen van de pagina's gestopt.

1 HTML staat voor 'HyperText Markup Language' en is de codeer taal om een webpagina te maken.

2 Algemene Verordening Gegevensbescherming



In bovenstaand voorbeeld is van wikipedia.org/index.html naar books.net/index.html de seed distance 2.

De hash methode

Een hash is zoals eerder beschreven een unieke code voor een afbeelding. Indien er tijdens het scannen een hash is opgehaald, wordt deze ter controle opgestuurd naar de Hash Check Service (HCS) die voor het huidige onderzoek is gebruikt. Deze database bevat hashes van afbeeldingen op basis van lijsten van o.a. de Nederlandse politie, Interpol en internationale autoriteiten zoals het NCMEC die door specialisten zijn aangemerkt als CSAM. Alleen een exacte match met een hash in deze database betekent dat er bekend CSAM is gevonden. Indien er een match was met de HCS heeft een analist van het Meldpunt Kinderporno de afbeelding bekeken om de match te valideren. Er zijn echter beperkingen in het gebruik van een hash database. Ten eerste kan er enkel gezocht worden naar afbeeldingen en op dit moment nog geen video's. Dit omdat er nog geen gekwalificeerde hash lijsten van CSAM video's beschikbaar zijn in de HCS. Daarnaast kan er alleen worden gezocht naar bekend materiaal waarvan de hash in deze database staat en zal nieuw CSAM niet gedetecteerd worden. Ook kan een kleine wijziging in een afbeelding een nieuwe hash opleveren waardoor de afbeelding niet meer wordt gedetecteerd.

PROCEDURE

In dit verkennend onderzoek wordt een computerprogramma aangezet die het gedrag van verschillende gebruikers op pornowebsites simuleert. Dit worden in dit onderzoek 'zoekprofielen' genoemd. De zoekprofielen gebruiken verschillende pornowebsites en categorieën of zoektermen als startpunt. Een voorbeeld van een zoekprofiel is een gebruiker die begint op een pornowebsite en klikt op de categorie 'teens' om daar een video te bekijken. Vanuit deze beginpagina worden vervolgens de links opgehaald en geanalyseerd totdat de seed distance van 4 is bereikt. Dit zoekgedrag op pornowebsites is onderverdeeld in verschillende zoekprofielen. Het overzicht van de zoekprofielen met het specifieke zoekgedrag is te vinden in Tabel 1.

RISICOTERMEN

In dit verkennend onderzoek worden key words gebruikt om een indicatie te geven van het risico dat een pagina CSAM bevat, en zal in het vervolg 'risicotermen' worden genoemd. De risicotermen die zijn gebruikt, zijn gebaseerd op lijsten van Interpol en de politie. Er is voor dit onderzoek

gekozen om een onderverdeling te maken in laag en hoog risicoterminen. Voor deze onderverdeling is gekeken naar de HTML data van de gevonden pagina's om te zien hoe vaak en in welke context de termen voorkomen. Laag risicoterminen zijn termen die indirect gerelateerd zijn CSAM, zoals 'lolita' en 'young', maar niet per definitie betekenen dat het om CSAM gaat. Deze termen komen veel voor op pornowebsites. Hoog risicoterminen zijn termen die CSAM suggereren, zoals '15 years old' en 'little boy'. Deze termen komen in mindere mate voor op pornowebsites.

RISICOSCORE

Om een indicatie te geven van het risico dat een pagina CSAM bevat, is er een risicoscore toegevoegd. De risicoscore is gebaseerd op het aantal risicoterminen en links naar bekend CSAM, en kan gebruikt worden om een overzicht te maken van de meest risicovolle pagina's. Deze risicoscore classificeert de HTML pagina's van een website als volgt:

- Risicoterm in de URL: +10
- Laag risicoterm op de pagina: +10 (met een maximum van 50)
- Hoog risicoterm op de pagina: +20 (met een maximum van 100)
- Combinatie van risicoterminen: +20 (met een maximum van 100)
- Linkt naar een plaatje met CSAM: +50

Het totaal aantal gescoorde punten wordt gedeeld door het totaal mogelijk punten (=310). Als een pagina bijvoorbeeld een laag risicoterm en een hoog risicoterm bevat dan heeft die pagina een risicoscore van $(10+20)/310 \times 100\% = 10\%$ (afgerond).

Dit onderzoek is exploratief van aard en er is niet eerder gebruik gemaakt van een risicoscore. Er is gekozen voor een score die een hogere uitkomst geeft indien er meer risicoterminen op de pagina staan (met name als dit hoog risicoterminen zijn), indien specifieke combinaties van termen voorkomen, of indien er bekend CSAM op de pagina staat. Een combinatie van risicoterminen zijn bijvoorbeeld de termen 'child' en 'abuse', 'kid' en 'nudist', of 'girl', 'tiny', en 'fuck'. Op deze manier is er een overzicht gemaakt van de meest risicovolle websites gebaseerd op de pagina's met de hoogste risicoscores. Deze websites geven een grotere kans om CSAM te bevatten.

ZOEKPROFIELEN

Er zijn voor dit onderzoek drie verschillende pornowebsites gebruikt als startpunt. De beschrijvingen van de pornowebsites zijn gebaseerd op observaties van het Meldpunt Kinderporno.

Pornowebsite 1: Een populaire pornowebsite die veel wordt gebruikt. De website toont weinig advertenties naar andere pornowebsites. Woorden die CSAM suggereren zijn verbannen op de website en ook maakt de website gebruik van waarschuwingsberichten.

Pornowebsite 2: Een populaire pornowebsite die veel wordt gebruikt. De website toont advertenties naar andere pornowebsites, maar blijft wel vaak binnen hetzelfde domein. Woorden die CSAM suggereren zijn niet verbannen en er worden geen waarschuwingsberichten getoond.

Pornowebsite 3: Deze website toont veel advertenties naar andere pornowebsites en kenmerkt zich doordat er veel pop-ups openen zodra ergens op de pagina wordt geklikt. Dit kan zijn op het moment dat er wordt geprobeerd om een video te openen, maar ook bij het wegklikken van de pagina. Op deze manier worden er steeds nieuwe pornowebsites geopend. Woorden die CSAM suggereren zijn niet verbannen en er worden geen waarschuwingsberichten getoond.

Verschillende categorieën en zoektermen

Er is gekozen voor een aantal specifieke categorieën en zoektermen. Vanuit de ervaring van Stop it Now weten we dat personen die contact opnemen met de hulplijn in verband met het kijken naar CSAM aangeven dat ze via pornowebsites zijn afgedwaald naar strafbaar materiaal. Ze zijn bijvoorbeeld begonnen met categorieën zoals 'teens' en hebben doorgeklikt op advertenties die opvolgend werden getoond. Ook geven veel kijkers aan dat ze steeds extremere beelden zijn gaan kijken en dat er ook interesse is voor ander extreem pornografisch materiaal (zoals hardcore of dieren). Om een vergelijking te maken is er ook een controle categorie of zoekterm gebruikt die geen associatie heeft met 'jong'. Dit heeft ertoe geleid dat de zoekprofielen opgedeeld kunnen worden in drie typen:

1. Zoekprofielen die als startpunt een categorie of zoekterm gebruiken gerelateerd aan jong (bijv. 'teens' of 'young girls').

2. Zoekprofielen die als startpunt een categorie of zoekterm gebruik gerelateerd aan extreme porno (bijv. 'brutal').

3. Zoekprofielen die een controle categorie of zoekterm gebruiken als startpunt (bijv. 'milf'³).

Uit onderzoek naar CSAM netwerken blijkt dat websites vaak overheersend gericht zijn op specifieke content, bijvoorbeeld websites die voornamelijk beeldmateriaal van jongens bevat of juist enkel van meisjes.⁴ Daarom zijn er binnen type 1 (gerelateerd aan jong) ook twee profielen ingezet met specifiek een zoekopdracht naar 'young boys' en 'young girls' om te onderzoeken of dit tot verschillende resultaten zou leiden.

Zoekgeschiedenis

Een pornogebruiker bouwt een zoekgeschiedenis op afhankelijk van de specifieke content waarop deze zoekt. Dit gebeurt in principe altijd als iemand browsst op het internet waarbij de activiteit op een website wordt opgeslagen in zogenaamde 'cookies'. Informatie in cookies kan bijvoorbeeld worden gebruikt om advertenties te tonen op basis van iemand zijn interesse. Om meer zicht te krijgen op het effect van een zoekgeschiedenis, is er een zoekprofiel gestart die klikt op de categorie 'teen' en die dit elke 3 uur herhaalt voor een langere periode. Dit profiel bouwt dus een

3 Een veel gebruikte pornocategorie: 'Mother I'd Like to Fuck'

4 Frank, Westlake & Bouchard, 2010

zoekgeschiedenis op waarin het regelmatig zoekt op teens. Ook is er een zoekprofiel die op elke website heeft geklikt waarop eerdere zoekprofielen twijfelachtig of strafbaar materiaal hebben gevonden. Dit zoekprofiel doorloopt dus meerdere keren het proces met steeds een andere website als beginpunt. Hierdoor heeft dit zoekprofiel ook een zoekgeschiedenis bestaande uit verschillende bezochte websites.

Er zijn in totaal 45 zoekprofielen gestart. Niet alle profielen hebben data opgehaald, bijvoorbeeld omdat er geen links werden gedetecteerd op de pagina. Deze zijn niet meegenomen in de verdere analyse. Ook de zoekprofielen die horen bij pornowebsite 1 en 2 die niet leiden tot twijfelachtig of strafbaar materiaal zijn niet meegenomen in het overzicht. Dit betekent dat er enkel zoekprofielen van pornowebsite 3 worden weergegeven. Tabel 1 geeft een overzicht van de selectie van zoekprofielen met zoekgedrag die zijn beschreven in dit rapport.

Tabel 1. Overzicht van zoekprofielen en zoekgedrag met startpunt pornowebsite 3

Zoekprofiel	Zoekgedrag
Profiel 1	Klikt op categorie: teen
Profiel 2	Klikt op categorie: blonde
Profiel 3	Klikt op categorie: brutal
Profiel 4	Klikt op categorie: teen, herhaald zich*
Profiel 5	Vult in zoekterm: young girls
Profiel 6	Vult in zoekterm: young boys
Profiel 7	Vult in zoekterm: loli**
Profiel 8	Vult in zoekterm: milf
Profiel 9	Klikt op alle websites waarop (potentieel) CSAM is aangetroffen

* Het zoekprofiel heeft de categorie 'teen' elke 3 uur aangeklikt over een langere periode.

** Deze term suggereert jong en wordt specifiek gebruikt om te zoeken naar jonge meiden.

STEEKPROEF

Alle opgehaalde hashes van de afbeeldingen zijn gescand door de Hash Check Service (HCS) om te controleren op bekend CSAM. Dit heeft echter een grote beperking want onbekend CSAM wordt hierdoor niet gedetecteerd. Daarom is per zoekprofiel een steekproef gemaakt van afbeeldingen op basis van de meest risicovolle websites. Om een beeld te krijgen van de meest risicovolle websites is gekeken naar de risicoscore van de pagina's. Op basis van de meest risicovolle websites is een steekproef van afbeeldingen geselecteerd. De afbeeldingen zijn vervolgens gecontroleerd door een analist van het Meldpunt Kinderporno om te onderzoeken of er twijfelachtig en strafbaar materiaal tussen zit. Twijfelachtig materiaal werd als zodoende aangemerkt indien het mogelijk wel gaat om een minderjarig persoon maar op basis van

de lichamelijke kenmerken niet met 100% zekerheid geconcludeerd kan worden dat het een minderjarige is.

Methode steekproef

Voor het maken van de steekproef zijn twee methodes gebruikt.

1. Risicoscore per profiel

Er is een lijst gemaakt van websites op basis van de hoogste risicoscores van de pagina's die zijn opgehaald bij het betreffende profiel. Vervolgens zijn van deze websites 10-15 afbeeldingen geselecteerd, afhankelijk van het aantal websites dat werd gevonden. Indien er weinig hoog risico websites waren, werd de steekproef aangevuld met afbeeldingen van websites met lager risico om wel voldoende afbeeldingen mee te nemen om het aantal voor de steekproef te behalen (N=300). Het komt veel voor dat de specifieke pagina's die zijn opgehaald bij het betreffende profiel een lage risicoscore hebben, terwijl de bijbehorende website wel een hoge gemiddelde risicoscore heeft (op basis van alle pagina's). Daarom is gekozen om een andere methode te gebruiken voor de tweede steekproef.

2. Gemiddelde risicoscore alle profielen

Er is een lijst gemaakt van websites met de hoogste gemiddelde risicoscore op basis van alle pagina's van een website die door alle profielen zijn opgehaald. Dit zijn dus de meest risicovolle websites op basis van alle opgehaalde data. Van de hoog risico websites die zijn opgehaald in het betreffende profiel worden 10-15 afbeeldingen geselecteerd.

De steekproef begint steeds met de meest risicovolle websites (risicoscore >20), maar de websites nemen af in risicoscore om voldoende websites mee te nemen waarop een steekproef van afbeeldingen kan worden gemaakt (N=200). De profielen verschillen in de specifieke websites en pagina's die worden opgehaald en dus ook in de gemiddelde risicoscore. In totaal zijn er per profiel 500 afbeeldingen gecheckt (combinatie van methode 1 en 2).

Betrouwbaarheid risicoscore

Om te testen of een hoge risicoscore ook daadwerkelijk tot meer twijfelachtig en strafbaar materiaal leidt, zijn er steekproeven gedaan van pagina's met een lage en een hoge risicoscore. Deze steekproeven zijn gemaakt op basis van de pagina's van een aantal profielen. Een analist van het Meldpunt Kinderporno heeft geanalyseerd of er twijfelachtig of strafbaar materiaal op de pagina's stond. Pagina's met een hoge risicoscore lijken vaker twijfelachtig of strafbaar materiaal te bevatten in vergelijking met pagina's met een lage risicoscore. Hierdoor werd de risicoscore van de pagina's als een goede maatstaf gezien voor het maken van een selectie van afbeeldingen.

04

RESULTATEN

PORNOWEBSITES

Onderzoeksvraag 1: Is er een verschil tussen pornowebsites in de kans om bij twijfelachtig materiaal en CSAM terecht te komen?

Op pornowebsite 1 en 2 waren de links voornamelijk naar websites gerelateerd aan hetzelfde domein, en ook werden veel websites opgehaald die niet relevant zijn voor dit onderzoek (bijv. nieuwspagina's). Er is een random steekproef (N=1000) gemaakt van afbeeldingen op basis van de profielen van pornowebsite 1 en 2 en hier was geen twijfelachtig of strafbaar materiaal op gevonden.

Vanuit pornowebsite 2 waren de meeste outlinks naar een forum waar mensen seksuele verhalen kunnen posten in verschillende thema's. Dit biedt tegelijk een platform voor 'taboo' verhalen, zoals minderjarig, incest, verkrachting en seks met dieren. Pagina's die zijn gevonden door de profielen zijn steekproefsgewijs geselecteerd op risicotermen en gecheckt op inhoud. Hierin werd duidelijk dat er in sommige verhalen seksuele scenario's beschreven werden met minderjarigen, dieren of een verkrachting. Verhalen met minderjarigen (zoals het beschrijven van fantasieën) zijn op dit moment niet strafbaar volgens de Nederlandse wetgeving, mits het gaat over fictieve verhalen. Het daadwerkelijk uitvoeren van seksuele handelingen met minderjarigen is wel strafbaar. De profielen met pornowebsite 3 als startpunt halen aanzienlijk grotere hoeveelheden pornowebsites op vergeleken met pornowebsite 1 en 2. Zo haalt pornowebsite 1 vrijwel geen andere pornowebsites op, pornowebsite 2 rond de 40, en pornowebsite 3 meer dan 2000. De profielen met pornowebsite 3 als startpunt hebben geresulteerd in twijfelachtig en strafbaar materiaal.

Conclusie

Er is een aantoonbaar verschil tussen pornowebsites in de connecties die ze hebben met

andere (porno-)websites. Van de onderzochte pornowebsites kwamen alleen de profielen van pornowebsite 3 uit bij websites met twijfelachtig en strafbaar materiaal. Het risico op strafbaar materiaal lijkt op basis van deze resultaten dus af te hangen van de opbouw van de pornowebsite.

ZOEKGEDRAG - CATEGORIEËN & ZOEKTERMEN

Onderzoeksvraag 2: Is er een verschil tussen categorieën en zoektermen in de kans om bij twijfelachtig materiaal en CSAM terecht te komen?

Er waren nauwelijks verschillen tussen de categorieën van de profielen van pornowebsite 1 en 2. Aangezien deze websites weinig risico lieten zien, zijn deze voor de verdere analyse niet meegenomen. Onderstaande gaat daarom enkel over de profielen van pornowebsite 3. Voor pornowebsite 3 zijn verschillende categorieën en zoektermen als startpunt gebruikt. Het totaal aantal opgehaalde afbeeldingen en de resultaten van de check met de Hash Check Service (HCS) zijn weergegeven in Tabel 2. De resultaten van de steekproeven zijn weergegeven in Tabel 3.

Tabel 2. Resultaten HCS voor zoekprofielen van pornowebsite 3

Profiel	Totaal afbeeldingen opgehaald	Totaal HCS hits (% van totaal afbeeldingen)	HCS hits gecheckt door analist			
			Volwassen	Twijfelachtig	CSAM	Niet gevonden
Categorie: teen	28.990	8 (0.027%)	8	0	0	0
Categorie: brutal	340.465	38 (0.011%)	28	7	2	0
Categorie: blonde	440.645	61 (0.013%)	48	10	3	0
Zoekterm: Young girls	787.288	163 (0.020%)	77	19	3	64
Zoekterm: young boys	750.436	186 (0.024%)	85	22	4	75
Zoekterm: loli	800.400	183 (0.022%)	82	22	5	74
Zoekterm: milf	551.942	55 (0.009%)	46	8	1	0

Tabel 3. Resultaten steekproeven voor zoekprofielen van pornowebsite 3

	Steekproef 1 (N=300)			Steekproef 2 (N=200)		
Profiel	Twijfelachtig	CSAM	Totaal	Twijfelachtig	CSAM	Totaal
Categorie: teen	6	4	10 (3.3%)	1	2	3 (1.5%)
Categorie: brutal	0	0	0 (0%)	1	0	1 (0.5%)
Categorie: blonde	0	0	0 (0%)	1	2	3 (1.5%)
Zoekterm: young girls	1	0	1 (0.3%)	2	0	2 (1%)
Zoekterm: young boys	2	0	2 (0.7%)	5	2	7 (3.5%)
Zoekterm: loli	0	0	0 (0%)	4	4	8 (4%)
Zoekterm: milf	0	0	0 (0%)	4	0	4 (2%)

De resultaten laten zien dat alle profielen de potentie hebben om twijfelachtig en/of strafbaar materiaal te vinden. Dus ongeacht of iemand zoekt op 'teens', een extreme categorie of een controle categorie bestaat de kans dat je binnen 4 kliks bij twijfelachtig of strafbaar materiaal terechtkomt. Het onderzoek is exploratief van aard en de steekproeven zijn klein, hierdoor is het niet mogelijk om harde conclusies te trekken over een specifieke categorie of zoekterm. Wel lijkt het totaal aantal Hash Check Service (HCS) hits in verhouding met het totaal aantal opgehaalde afbeeldingen in Tabel 2 groter te zijn bij categorieën die een associatie hebben met jong. Echter, een groot aantal URL's zijn niet teruggevonden waardoor er niet gecheckt kan worden wat er op de URL te zien was. Dit kan bijvoorbeeld komen doordat het materiaal is verwijderd. In Tabel 3 wordt het vaakst CSAM gevonden door de profielen die zoeken op 'teen', 'young boys', en 'loli'. Dit zijn categorieën of zoektermen die een associatie hebben met jong.

Conclusie

Op pornowebsite 3 kwamen de zoekprofielen binnen 4 kliks bij websites met twijfelachtige beelden of CSAM terecht, ongeacht de gebruikte categorie of zoekterm. Op basis van de HCS hits en de steekproeven lijken de zoekprofielen die relateren aan jong vaker dergelijke beelden te vinden.

ZOEKGESCHIEDENIS

Onderzoeksvraag 3: Beïnvloedt de zoekgeschiedenis van een pornogebruiker de kans om bij twijfelachtig materiaal en CSAM terecht te komen?

Het hebben van een zoekgeschiedenis leidde bij de profielen van pornowebsite 1 en 2 niet tot meer risicovolle pagina's. Bij pornowebsite 1 werden er wel meer websites opgehaald, maar die waren niet relevant voor het onderzoek (bijv. nieuwspagina's). Daarom zijn deze voor de verdere analyse niet meegenomen. Onderstaande gaat daarom enkel over de profielen van pornowebsite 3.

Voor pornowebsite 3 is een vergelijking gemaakt met een zoekprofiel dat een enkele keer klikt op de categorie 'teen', een zoekprofiel die dit elke 3 uur herhaalt voor een langere periode, en een zoekprofiel die op elke website klikt waarop twijfelachtig of strafbaar materiaal is gevonden. De resultaten hiervan zijn weergegeven in Tabel 4 en Tabel 5.

Tabel 4. Resultaten HCS hits voor zoekprofielen van pornowebsite 3

Profiel	Totaal afbeeldingen opgehaald	Totaal HCS hits (% van totaal afbeeldingen)	HCS hits gecheckt door analist			
			Volwassen	Twijfelachtig	CSAM	Niet gevonden
Categorie: teen	28.990	8 (0.027%)	8	0	0	0
Categorie: teen (elke 3 uur herhaald)	461.475	73 (0.015%)	55	12	3	3
Klikt op alle risicovolle sites	1.342.923	323 (0.024%)	262	35	11	15

Tabel 5. Resultaten steekproeven voor zoekprofielen van pornowebsite 3

	Steekproef 1 (N=300)			Steekproef 2 (N=200)		
Profiel	Twijfelachtig	CSAM	Totaal	Twijfelachtig	CSAM	Totaal
Categorie: teen	6	4	10 (3.3%)	1	2	3 (1.5%)
Categorie: teen (elke 3 uur herhaald)	0	0	0 (0%)	5	0	5 (2.5%)
Klikt op alle risicovolle sites	4	0	4 (1.3%)	1	0	1 (0.5%)

Een zoekprofiel die herhaaldelijk zoekt op 'teen' haalt meer afbeeldingen op en heeft daardoor ook meer hits met de Hash Check Service (HCS) vergeleken met het zoekprofiel dat eenmalig zoekt op 'teen'. De resultaten van de steekproeven zijn niet eenduidig. Doordat het kleine steekproeven betreft, is het niet mogelijk om harde conclusies te trekken over de zoekgeschiedenis. Het zoekprofiel dat als startpunt begint bij risicovolle websites en hier meerdere van aanklikt, haalt aanzienlijk meer afbeeldingen op vergeleken met de andere profielen. Dit profiel toont meer hits met de HCS, maar niet meer twijfelachtige of strafbare beelden op basis van de steekproef.

Conclusie

Op basis van de gebruikte zoekprofielen lijkt er geen eenduidig bewijs te zijn dat een zoekgeschiedenis vaker leidt tot twijfelachtige of strafbare beelden.

BESCHRIJVING VAN DE PROFIELEN

Risicotermen

Tabel 6 geeft een overzicht van het aantal pagina's met hoog risicotermen per profiel. Hier is te zien dat profiel 1 (categorie: 'teen') veruit de meeste risicovolle pagina's heeft. Op basis van de gemiddelde risicoscore van alle pagina's blijkt dat profiel 6 (zoekterm: 'young boys') de hoogste score heeft, gevolgd door profiel 1 (categorie: 'teen').

Tabel 6. Risicoscore en risicotermen per profiel

Profiel	Aantal pagina's met hoog risico termen	Gemiddelde risicoscore alle pagina's
Profiel 1 Categorie: teen	16.499	6.99
Profiel 7 Zoekterm: loli	6.930	6.34
Profiel 2 Categorie: blonde	5.768	6.15

Profiel 9 Klikt op alle risicovolle sites	2.081	5.15
Profiel 6 Zoekterm: young boys	1.056	7.1
Profiel 3 Categorie: brutal	929	5.05
Profiel 5 Zoekterm: young girls	670	3.85
Profiel 8 Zoekterm: milf	498	4.66
profiel 4 Categorie: teen (elke 3 uur herhaald)	368	6.28

Random steekproef

Er is een random steekproef gedaan van afbeeldingen op basis van alle websites die zijn opgehaald door profiel 1 (categorie: 'teen'). Op basis van 2720 afbeeldingen zijn er 22 twijfelachtig of strafbaar, dit is 0.8% van de steekproef. De steekproeven van profiel 1 op basis van de meest risicovolle websites gaven een percentage van 3.3% (steekproef 1) en 1.5% (steekproef 2). Steekproeven op basis van de meest risicovolle websites geven dus een grotere kans op twijfelachtig of strafbaar materiaal.

Opvallende bevindingen

Tijdens het controleren van de pagina's door een analist van het Meldpunt Kinderporno was opgevallen dat bepaalde URL's tijdens het bekijken, refereerden naar een andere website met twijfelachtig of strafbaar materiaal. Dit gebeurde door middel van een pop-up. Door het kijken van een pagina die niet direct strafbaar materiaal bevat, kan er dus een pop-up worden geopend die leidt naar een website waar wel strafbaar materiaal op staat. Het viel tevens op dat de websites met twijfelachtig en strafbaar materiaal erg overeen kwamen in opzet, vormgeving en domeinnaam. Een andere opvallende bevinding is dat de websites die twijfelachtig of strafbaar materiaal bevatten in veel gevallen Aziatische pornowebsites zijn. Dit zijn websites die bestaan uit pornografische beelden met Aziatische personen.

05

DISCUSSIE

Deze exploratieve studie heeft onderzocht welk zoekgedrag op pornowebsites mogelijk kan leiden naar twijfelachtig en strafbaar materiaal. De resultaten suggereren dat het risico voornamelijk afhangt van de opbouw van een pornowebsite wat mogelijk samenhangt met het aantal connecties dat een website heeft naar andere pornowebsites.

Zo vertoont pornowebsite 3 veel pop-ups naar andere pornowebsites wat leidde tot een groot netwerk van aan elkaar gelinkte websites. Een deel van deze websites had een hoge risicoscore op basis van de opgehaalde pagina's en ook is er op een aantal van deze websites twijfelachtig en strafbaar materiaal gevonden. Dit was ongeacht de specifieke pornocategorie waarop werd gezocht. Zowel het zoeken op jong (bijv. 'teens'), een extreme pornocategorie (bijv. 'brutal'), of een controle categorie (bijv. 'blonde') zorgde ervoor dat er twijfelachtig materiaal of misbruikbeelden werd gevonden.

Er worden in dit verkennend onderzoek relatief lage aantallen (potentieel) CSAM gevonden middels de Hash Check Service (HCS.) Ander onderzoek vond ook dat op openbare websites weinig HCS hits worden gevonden.¹ Ook in de steekproeven werden kleine aantallen (potentieel) CSAM gevonden. Echter, hierbij dient rekening te worden gehouden met het feit dat het kleine steekproeven zijn en dat het werkelijke aantal CSAM dat op pornowebsites te vinden is, mogelijk hoger zal zijn. Op basis van een random steekproef van afbeeldingen werd een percentage gevonden van 0.8% (potentieel) CSAM. Dit komt neer op 232 twijfelachtige of strafbare afbeeldingen. Het totaal aantal strafbare afbeeldingen op pornowebsites is dus vermoedelijk vele malen hoger.

¹ Westlake, Bouchard, & Frank, 2012; 2017

Dit verkennend onderzoek laat zien dat mensen in hun zoektocht naar porno, ook onbedoeld, terecht kunnen komen bij CSAM. Dit sluit aan bij een recent onderzoek op het Darkweb waarin kijkers van misbruikbeelden hebben aangegeven dat ze CSAM zijn tegengekomen op het openbare internet, waarvan 32% op pornowebsites.² Uit eerder onderzoek weten we ook dat mensen vaak niet actief op zoek gaan naar CSAM maar tijdens het porno kijken er wel mee in aanraking kunnen komen.³ Een groot deel van de CSAM kijkers op het Darkweb geeft aan dat ze de eerste keer onbedoeld zijn geconfronteerd met het materiaal, bijvoorbeeld door het per ongeluk te zijn tegengekomen, terwijl ze aan het zoeken waren naar ander extreem materiaal, of doordat iemand het aan hen stuurde.⁴ Door confrontatie met deze beelden bestaat de kans dat ze dit materiaal daarna blijven bekijken. Het is dus belangrijk om te voorkomen dat mensen tijdens het porno kijken in aanraking komen met CSAM.

Beperkingen huidig onderzoek

Er zijn een aantal beperkingen aan dit onderzoek. Zo is een groot deel van de Hash Check Service (HCS) hits door een analist van het Meldpunt Kinderporno beoordeeld als volwassen pornografie. De HCS is gebaseerd op lijsten van autoriteiten waarvan is beoordeeld dat het minderjarigen betreft. Er zijn echter lijsten waarin ook volwassen pornografie voorkomt. Dit zal in de toekomst verder opgeschoond moeten worden zodat het enkel hashes betreft van CSAM. Tijdens het analyseren van de steekproeven zijn ook veel URL's gevonden waarbij de afbeelding of link niet meer bestond. Dit kan betekenen dat de afbeelding offline is gehaald, waardoor het niet meer te achterhalen is of het ging om strafbaar beeldmateriaal. Hierdoor geven de HCS hits een onvolledig beeld.

Door de omvang van de dataset en beperkte technische middelen waren de steekproeven klein en was het niet mogelijk om ze compleet te randomiseren. Daarom kunnen de zoekprofielen niet objectief vergeleken worden en laten ze vooral zien dat elke zoekterm de potentie heeft om bij risicovolle websites en (potentieel) CSAM terecht te komen. Daarnaast vormden de risicoterminen een grote uitdaging in dit onderzoek omdat op pornowebsites veel seks gerelateerde woorden voorkomen en ook termen zoals 'teen' en 'young'. Door het grote aantal pagina's dat wordt opgehaald, is het niet mogelijk om van elke pagina de context te achterhalen. Hierdoor is het lastig om een goede risico indicatie te maken van een pagina. Toch laat eerder onderzoek zien dat key words die niet direct gerelateerd zijn aan seksueel misbruik wel in combinatie gebruikt kunnen worden om naar CSAM te zoeken.⁵ Om zo betrouwbaar mogelijk het risico in te schatten is dan ook een lijst met verschillende key words gebruikt, in combinatie met het checken van hashes, wat in een ander onderzoek effectief was om websites met CSAM te detecteren.⁶

Er zijn daarnaast ook een aantal technologische beperkingen. Het kan zijn dat het nodig is om te registreren op bijvoorbeeld een forum voordat iemand toegang kan krijgen tot strafbaar materiaal. Het zoekprofiel kan niet verder als deze op een inlogpagina of een betaalmuur stuit. Verder bleek het in de praktijk lastig om een pornogebruiker na te bootsen. De technologie scant de pagina's

2 Suojellaan Lapsia, Protect Children ry., 2024

3 Nationaal Rapporteur Mensenhandel en Seksueel Geweld tegen Kinderen & Stop it Now, 2023

4 Insoll et al., 2022

5 Steel, 2009

6 Westlake, Bouchard, & Frank, 2012; 2017

voor links en er wordt hierbij niet bewust een keuze gemaakt welke video er wordt bekeken. Een pornogebruiker maakt wel een bewuste keuze; als iemand graag naar jonge vrouwen kijkt zal deze gebruiker video's aanklikken waarop in de beschrijving of preview te zien is dat het gaat om een jonge vrouw. Het algoritme toont video's die passen bij deze voorkeur en zo kan iemand in theorie steeds verder afdwalen in video's van jonge vrouwen. In dit onderzoek is getracht een zoekgeschiedenis na te bootsen van iemand die vaak zoekt op de categorie 'teen'. Toch blijft dit beperkt tot het klikken op deze categorie en is dit geen realistische zoekgeschiedenis van een pornogebruiker.

Mogelijkheden vervolgonderzoek

Voor vervolgonderzoek is het belangrijk om het zoekgedrag van een pornogebruiker realistischer te kunnen nabootsen. Bijvoorbeeld door het ontwikkelen van technologie die 'intelligenter' kan zoeken op specifieke video's, zoals het zoeken naar woorden bij video's die wijzen op jong en vervolgens alleen op de links van deze video's klikken.

Er is veel in ontwikkeling op het gebied van Artificial Intelligence (AI). Zo wordt AI gebruikt om tools te ontwikkelen die CSAM kunnen classificeren op basis van kenmerken in de video waardoor het op den duur minder afhankelijk zou zijn van menselijke analisten. Wel zal altijd de menselijke check nodig zijn om te bepalen of het daadwerkelijk strafbaar materiaal betreft. Het gebruiken van AI om CSAM te classificeren had voor dit onderzoek een aanwinst geweest, omdat op die manier in een korte tijd een veel grotere dataset kan worden geclassificeerd. Hierdoor kan het totaal aantal CSAM per profiel worden vastgesteld waardoor een betrouwbare vergelijking mogelijk is.

De profielen van pornowebsite 3 kwamen uit bij websites met twijfelachtig materiaal en CSAM. Pornowebsite 3 wordt gekenmerkt door het openen van pop-ups en liet veel connecties zien met andere pornowebsites. Ook werd geobserveerd door de analist dat de websites met twijfelachtig en strafbaar materiaal erg overeen kwamen wat betreft opzet en vormgeving. Dit roept verschillende vragen op: hoe zijn deze websites precies aan elkaar verbonden, komen ze vanuit dezelfde eigenaar en host, hoe ziet het verdienmodel eruit, etc. Om websites die CSAM bevatten beter te kunnen detecteren, kan het waardevol zijn om de structuur en connecties van deze websites nader te onderzoeken. Daarnaast zou het interessant zijn om te onderzoeken of op basis van HTML data patronen ontdekt kunnen worden die websites die CSAM bevatten onderscheiden van andere pornowebsites.

AANBEVELINGEN

Het is belangrijk dat mensen veilig porno kunnen kijken en dit kan alleen als er meer regulatie en toezicht komt op pornowebsites wat er voor zorgt dat er geen strafbare beelden op deze websites te vinden zijn. Maar dit is niet gemakkelijk te realiseren omdat dit veel praktische vraagstukken geeft. Bijvoorbeeld: Wie gaat dat reguleren? Hoe kun je zeker weten dat een website veilig is? En hoe ga je ervoor zorgen dat elke pornowebsite hieraan meewerkt? Toch zijn er een aantal maatregelen die websites zelf kunnen nemen. Als mensen deze maatregelen op een website kunnen herkennen, kunnen ze zelf een bewuste keuze maken wel of niet op de website te blijven. Hieronder volgen een aantal suggesties.

Keurmerk veilige pornowebsite

Experts vanuit verschillende disciplines kunnen met elkaar in gesprek gaan over de maatregelen die pornowebsites kunnen nemen om ervoor te zorgen dat er geen misbruikbeelden op hun website staan. De sector kan samenwerken voor het opstellen van richtlijnen in de vorm van een 'code'. Het is hierbij belangrijk dat pornowebsites er zelf ook bij betrokken zijn. Pornowebsites kunnen er dan voor kiezen om zich hierbij aan te sluiten en deze 'code' te ondertekenen. Dit zou zich in de toekomst verder kunnen ontwikkelen tot een 'keurmerk' die een website kan krijgen indien deze aan bepaalde voorwaarden voldoet. Er moet dan wel een onafhankelijke en deskundige partij zijn die hierop kan toetsen en herhaaldelijk kan monitoren.

Check met de HCS

Het is mogelijk om beeldmateriaal op een website te laten checken op bekend CSAM door een hash database, zoals de Hash Check Service (HCS). Indien er strafbaar materiaal op de website wordt gedetecteerd, kan dit materiaal worden verwijderd. Op deze manier kan er relatief eenvoudig worden voorkomen dat er bekend CSAM op een website te vinden is. Echter, met de nieuwe Europese Digital Service Act (DSA) is het nog niet mogelijk om een check met de HCS verplicht te stellen. Wel kunnen platformen of pornowebsites er zelf voor kiezen om zich aan te sluiten bij de HCS. Dit kan via het Meldpunt Kinderporno.

Verificatie uploaders en betrokkenen

Een ID verificatie kan worden ingesteld om de uploader van pornografische beelden te verifiëren. Indien er strafbare beelden worden geüpload, is dit dus gekoppeld aan een persoon. Al moeten er goede maatregelen zijn met betrekking tot de privacy van de gebruiker. Verificatie is ook belangrijk in het voorkomen dat seksueel beeldmateriaal zonder toestemming wordt geüpload op pornowebsites. Alle betrokken personen in de beelden moeten zich kunnen identificeren en toestemming verlenen. Daarnaast is toezicht hier ook van belang. Het gaat in dit geval om bijzondere persoonsgegevens die vallen onder de AVG waarvoor de Autoriteit Persoonsgegevens verantwoordelijk is. Aangezien er twijfelachtige beelden te vinden zijn op pornowebsites moet er dus gecontroleerd worden dat het niet om minderjarigen gaat.

Moderatie

Moderatie valt onder de verplichtingen van de Digital Service Act (DSA). Door goede moderatie kan worden voorkomen dat strafbaar materiaal op een platform terechtkomt. Ook kunnen websites gebruikers de mogelijkheid bieden om video's te rapporteren zodat deze snel bekeken en verwijderd kunnen worden, bijvoorbeeld als het illegale beelden betreft of als beeldmateriaal zonder toestemming online is gezet. Het is belangrijk dat er goed toezicht is op deze moderatie, zeker gezien er websites zijn met twijfelachtig en strafbaar materiaal. De Autoriteit Consument & Markt (ACM) is verantwoordelijk voor het toezicht op de moderatie. De Autoriteit online Terroristisch en Kinderpornografisch Materiaal (ATKM) houdt toezicht op de verspreiding van seksueel beeldmateriaal van minderjarigen.

Waarschuwingsberichten

Een platform kan preventieve maatregelen nemen gericht op personen die mogelijk op zoek zijn naar misbruikbeelden. Zo kunnen termen die wijzen op minderjarige personen of ander strafbaar gedrag verbannen worden op een pornowebsite. Op het moment dat personen hierop zoeken

krijgen ze een automatisch waarschuwingsbericht te zien, bijvoorbeeld dat hun zoekopdracht strafbaar materiaal suggereert en wat de mogelijke consequenties hiervan zijn. Hierbij kan worden verwezen naar een hulporganisatie, zoals Stop it Now, die hulp biedt aan personen die misbruikbeelden kijken. Websites zoals Pornhub doen dit al.

Wat kunnen pornogebruikers zelf doen?

Dit onderzoek laat zien dat pop-ups naar andere, onbekende pornowebsites mogelijk kunnen leiden naar strafbaar materiaal. Websites die veel pop-ups genereren, zijn dus mogelijk risicovoller en kunnen beter worden vermeden. Het is aan te raden om binnen een vertrouwde website te blijven en bij twijfel een website weg te klikken. Ook kunnen pornogebruikers ervoor kiezen om betaalde websites te gebruiken, zoals platformen die ethische porno aanbieden. Bij ethische porno worden alle betrokken personen, zowel de acteurs/actrices als de cameraman, eerlijk betaald. Ethische porno wordt gemaakt met consent van alle betrokkenen, laat vaak meer diversiteit zien (in leeftijd, gender, seksuele oriëntatie, etc.), gebruikt veiligheidsmaatregelen (zoals soa testen en anticonceptie), en is transparant (bijv. over de performers, regisseur en crew maar ook 'behind the scenes').

Het is daarnaast verstandig om een website van tevoren te checken op veiligheid. Op een website wordt vermeld hoe het platform omgaat met privacy en welke maatregelen worden genomen om het platform veilig te houden. Dat betekent dat een pornogebruiker oplettend kan zijn om bovenstaande veiligheidsmaatregelen op een pornowebsite te zoeken en op basis daarvan een bewuste keuze te maken. Het is hiervoor wel belangrijk dat mensen hierover geïnformeerd worden.

Voorlichting

Het is belangrijk dat er goede voorlichting beschikbaar is voor zowel jongeren als volwassenen over veilig pornogebruik. Dit zou deel moeten zijn van een breder pakket aan seksuele voorlichting voor jongeren over gezonde en positieve seksualiteit en het tegengaan van online seksueel grensoverschrijdend gedrag. Voorlichting kan gaan over stereotypen die te zien zijn in porno, maar ook waar je op kunt letten voor het inschatten van risico's en veiligheid, en hoe je bewuste keuzes kunt maken over de pornowebsite die je wilt gebruiken en het type porno waar je naar kijkt. Hierbij is het belangrijk dat er een plek is waar mensen goede informatie kunnen vinden maar ook laagdrempelig vragen kunnen stellen en eventuele twijfels of zorgen kunnen bespreken.

Conclusie

Dit verkennend onderzoek laat zien dat het doorklikken op pornowebsites (middels pop-ups) kan leiden tot risicovolle websites met twijfelachtig en strafbaar materiaal. Daarom is het belangrijk dat er goede regulatie komt op pornowebsites waarmee voorkomen kan worden dat pornogebruikers in aanraking komen met beeldmateriaal van seksueel misbruik van minderjarigen. Het is de verantwoordelijkheid van pornowebsites om te zorgen dat ze alle maatregelen treffen die voorkomen dat er strafbaar materiaal op hun platform te vinden is of dat er gelinkt wordt naar websites met strafbaar materiaal. De Nederlandse overheid moet zorgen voor gepast beleid dat hierop toeziet. Omdat het toezicht verdeeld is over verschillende autoriteiten is het essentieel een gezamenlijke aanpak te ontwikkelen. Daarbij is goede voorlichting belangrijk, wat deel kan zijn van een integrale aanpak die online seksueel grensoverschrijdend gedrag tegen gaat. De juiste kennis bieden, aan jongeren in het bijzonder, kan mensen helpen om bewuste keuzes te maken

waardoor ze risico's kunnen verminderen. Hierdoor kunnen mensen online op een veilige manier hun seksualiteit ontdekken, wordt het bekijken en verspreiden van misbruikbeelden tegengegaan, en wordt (herhaald) slachtofferschap voorkomen.

REFERENTIES

1. Frank, R., Westlake, B., & Bouchard, M. (2010). The structure and content of online child exploitation networks. Proceedings of ISI-KDD '10: ACM SIGKDD Workshop on Intelligence and Security Informatics, 3, 1-9. <https://doi.org/10.1145/1938606.1938609>
2. Guerra, E., & Westlake, B. G. (2021). Detecting child sexual abuse images: Traits of child sexual exploitation hosting and displaying websites. Child Abuse & Neglect, 122, 105336. <https://doi.org/10.1016/j.chiabu.2021.105336>
3. Insoll, T., Ovaska, A.K., Nurmi, J., Aaltonen, M., & Vaaranen-Valkonen, N. (2022). Risk factors for child sexual abuse material users contacting children online: Results of an anonymous multilingual survey on the dark web. Journal of Online Trust & Safety, 1(2), 1-24. <https://doi.org/10.54501/jots.v1i2.29>
4. Nationaal Rapporteur Mensenhandel en Seksueel Geweld tegen Kinderen (2023). Monitor seksueel geweld tegen kinderen 2018-2022. Nationaal Rapporteur.
5. Nationaal Rapporteur Mensenhandel en Seksueel Geweld tegen Kinderen & Stop it Now (2023). Maak het bespreekbaar: Onderzoek naar jonge (potentiële) daders van seksueel kindermisbruik, Nationaal Rapporteur.
6. Westlake, B. G., Bouchard, M., & Frank, R. (2011). Finding the key players in online child exploitation networks. Policy & Internet, 3(2), 1-32. <https://doi.org/10.2202/1944-2866.1126>
7. Westlake, B., Bouchard, M., & Frank, R. (2012). Comparing methods for detecting child exploitation content online. European Intelligence and Security Informatics Conference, 156-163. doi: 10.1109/EISIC.2012.25
8. Westlake, B., Bouchard, M., & Frank, R. (2017). Assessing the validity of automated webcrawlers as data collection tools to investigate online child sexual exploitation. Sexual Abuse, 29(7), 685-708. <https://doi.org/10.1177/1079063215616818>
9. Steel, C. M. S. (2009). Child pornography in peer-to-peer networks. Child Abuse & Neglect, 33(8), 560-568. <https://doi.org/10.1016/j.chiabu.2008.12.011>
10. Suojellaan Lapsia, Protect Children ry. (2024). Tech platforms used by online child sexual abuse offenders: Research report with actionable recommendations for the tech industry.

BEGRIPPENLIJST

♦ **Beeldmateriaal van seksueel misbruik van minderjarigen**

Afbeeldingen of video's met een seksuele strekking, waarbij iemand is betrokken die de leeftijd van achttien jaar nog niet heeft bereikt. Afbeeldingen of video's worden als zodoende aangemerkt als er sprake is van een expliciete seksuele houding, als er seksuele handelingen worden getoond, of als een minderjarige zich in een zodanige houding of omgeving bevindt waardoor de afbeelding of video een onmiskenbare seksuele strekking heeft.

♦ **CSAM**

De Engelstalige term en afkorting voor beeldmateriaal van seksueel misbruik van minderjarigen. CSAM staat voor 'child sexual abuse material'.

♦ **Darkweb**

Het Darkweb is een anoniem en verborgen gedeelte van het internet dat enkel toegankelijk is met specifieke software.

♦ **Hash / Hash database**

Een hash is een code die uniek is voor een afbeelding. Je kunt het zien als een digitale vingerafdruk. Een hash database bevat hashes van afbeeldingen op basis van lijsten die door autoriteiten zijn aangemerkt als CSAM. Door middel van een exacte match met een hash in de database kan bekend CSAM worden gedetecteerd.

♦ **Risicotermen**

Voor dit onderzoek worden risicotermen gebruikt om een indicatie te geven van het risico dat er CSAM op de pagina staat. De risicotermen zijn gebaseerd op lijsten van autoriteiten en zijn woorden die geassocieerd zijn met CSAM, bijvoorbeeld 'little girl'. In wetenschappelijk literatuur wordt dit meestal 'key words' genoemd.

♦ **Twijfelachtig beeldmateriaal**

Twijfelachtig beeldmateriaal betekent dat het mogelijk wel gaat om een minderjarige maar dat het op basis van de lichamelijke kenmerken niet met honderd procent zekerheid te concluderen is dat het een minderjarige is.

♦ **Web crawler**

Een webcrawler is een computerprogramma die websites systematisch scant door op hyperlinks te klikken en informatie van pagina's op te halen.

♦ Zoekgeschiedenis

In dit onderzoek wordt met zoekgeschiedenis al het zoekgedrag bedoelt van een gebruiker om naar pornografische content te zoeken. Bijvoorbeeld de specifieke pagina's die iemand heeft bezocht of de zoektermen die zijn gebruikt. In principe heeft iedereen die browsert op het Internet een zoekgeschiedenis waarbij de activiteit op een website wordt opgeslagen in zogenaamde 'cookies'. Informatie in cookies kan bijvoorbeeld worden gebruikt om advertenties te tonen op basis van iemand zijn interesse.

♦ Zoekprofiel

In dit onderzoek wordt een computerprogramma aangezet, ook wel zoekprofiel genoemd, die het gedrag van verschillende gebruikers op pornowebsites simuleert. De zoekprofielen gebruiken verschillende pornowebsites en categorieën of zoektermen als startpunt. Zo kan een zoekprofiel beginnen op een pornowebsite waar deze klikt op de categorie 'teen' om vervolgens de links op te halen en verder te klikken.

COLOFON

Referentie: Stop it Now (2024) ***Dichterbij dan je denkt:
Een verkennend onderzoek naar risicovol zoekgedrag
op pornowebsites***

Kelly van den Heuvel

Stop it Now, Offlimits

www.stopitnow.nl

Offlimits

Nieuwezijds Voorburgwal 120-3

1012 SH Amsterdam

020 261 5232

www.offlimits.nl