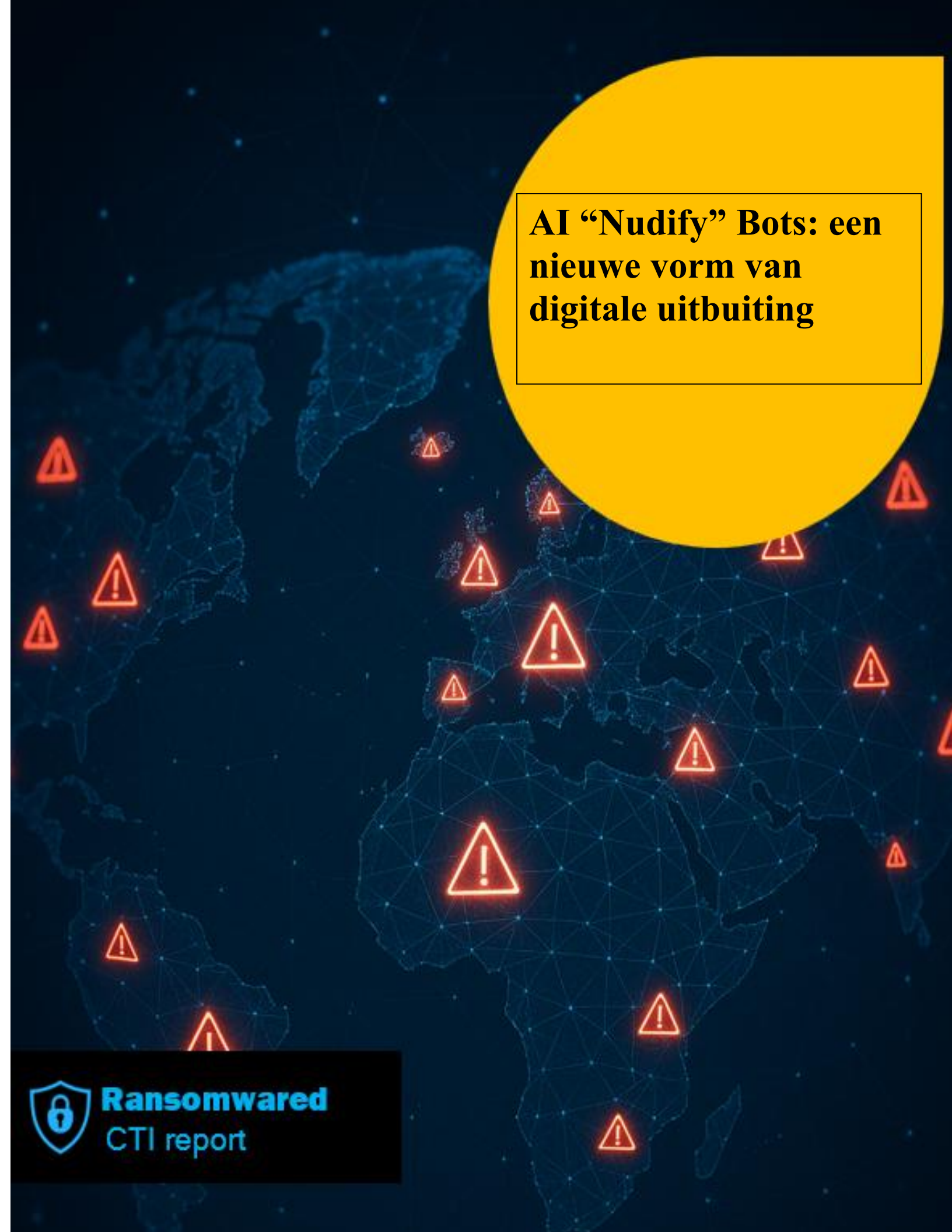
AI “Nudify” Bots: een nieuwe vorm van digitale uitbuiting
Ransomwared
CTI report

AI “Nudify” Bots: een nieuwe vorm van digitale uitbuiting

Door het Ransomwared CTI-team – Oktober 2025
TLP: GREEN – vrij deelbaar met bronvermelding

1.1 Inleiding

In de afgelopen twee jaar is er stilletjes een technologische wedloop ontstaan aan de randen van de AI-ontwikkelingen.
Terwijl de wereld gefascineerd toekijkt hoe generatieve AI teksten schrijft, beelden creëert en muziek componeert, wordt diezelfde technologie ook misbruikt om iets veel donkerders te doen: het digitaal “uitkleden” van mensen.

Zogeheten **“nudify-bots”** of **AI-nudificatieservices** gebruiken beeldgeneratie om van gewone foto’s of korte video’s realistisch ogende, seksueel expliciete versies te maken — zonder toestemming van de geportretteerde. Wat ooit technisch complex en arbeidsintensief was, is nu een geautomatiseerd proces geworden dat in een paar seconden via een telefoon kan worden uitgevoerd.

Gebruikers uploaden een foto of filmpje naar een site of een **Telegram-bot** en krijgen kort daarna een gemanipuleerd resultaat terug. Achter de schermen worden deze beelden bewaard, gekoppeld aan gegevens en soms gebruikt voor chantage.
Door de combinatie van open-source-modellen, goedkope cloudrekenkracht en anonieme cryptobetalingen is grootschalig misbruik spotgoedkoop en vrijwel risicoloos geworden.

Dit report beschrijft waarom deze ontwikkeling maatschappelijk en veiligheidskundig relevant is, hoe het businessmodel werkt en waarom het één van de urgentste digitale dreigingen van 2025 vormt.

1.2 De schaal van het fenomeen

Waar er in 2023 nog maar enkele van zulke bots bestonden, telden onderzoekers halverwege 2024 al **meer dan 50 actieve bots op Telegram**, goed voor naar schatting **3–4 miljoen maandelijkse gebruikers**.
Daarbovenop komen talloze klonen op Discord, eigen websites en darknet-marktplaatsen.
Elke keer dat een bot wordt verwijderd, verschijnen er binnen uren nieuwe exemplaren met andere namen en wallets.

De financiële prikkel is simpel: een paar euro per foto, kleine maandabonnementen en vrijwel geen handhavingsrisico.
De lage drempel trekt niet alleen criminelen aan, maar ook nieuwsgierige gebruikers die zich niet realiseren dat hun handeling strafbaar of schadelijk is.

1.3 Waarom dit een CTI-onderwerp is

Op het eerste gezicht lijkt dit geen klassiek cyberdreigingsvraagstuk.
Er is geen malware, geen exploit, geen ransomware.
Toch raakt het drie kerngebieden van cyber threat intelligence:

1. **Digitale identiteit en vertrouwen.** Beelden zijn overtuigende bewijzen. Door gemanipuleerde beelden te verspreiden, kunnen aanvallers reputaties schaden, mensen chanteren of organisaties destabiliseren.
2. **Infrastructuur.** Dezelfde netwerken van bots, cryptowallets en geanonimiseerde hosting worden ook gebruikt voor phishing, fraude en ransomware.
3. **Incidentrespons.** SOC's en CERT's krijgen steeds vaker meldingen van beeldmanipulatie-chantage. Zonder kennis van dit fenomeen kunnen zij slachtoffers niet adequaat helpen.

Kortom: AI-nudificatie is niet enkel een moreel probleem, maar een **operationele dreiging** die de integriteit van digitale identiteiten ondermijnt.

1.4 De sociale ingenieurslaag

Elke nudify-dienst begint met een lokmiddel: een "grappige" of "onschuldige" uitdaging.
Berichten met teksten als *"Laat AI je outfit beoordelen"* of *"Bekijk jezelf in virtuele mode"* circuleren op sociale media.
Wie meedoet, maakt een account aan en uploadt een foto. De operator verzamelt ondertussen e-mail, IP-adres en andere gegevens die zogenaamd nodig zijn.

Diezelfde informatie kan later worden gebruikt voor **afpersing**: *"Betaal of we publiceren het resultaat."*
CTI-onderzoekers duiden dit als **image-based social engineering**: emotionele manipulatie in plaats van technische exploitatie.

1.5 De schijn van onschuld

Veel diensten presenteren zich als "filters" of "AI-kunst".
Gebruiksvriendelijke interfaces en nep-privacyverklaringen ("we slaan niets op") wekken vertrouwen.
In werkelijkheid worden zowel bron- als resultaatbeelden opgeslagen om nieuwe modellen mee te trainen.
Zo groeit ongemerkt een steeds grotere database van echte gezichten, die weer als trainingsdata voor volgende generaties misbruik-AI dient.

1.6 Van nieuwsgierigheid naar criminaliteit

Niet elke gebruiker heeft kwade bedoelingen, maar intentie doet er weinig toe zodra beelden online zijn gezet.
Eenmaal geüpload betekent **verlies van controle**.
En in steeds meer landen geldt: het maken of bezitten van seksuele deepfakes zonder toestemming is strafbaar, ook als ze door AI zijn gemaakt.

CTI-teams moeten zich hiervan bewust zijn, want incidenten die lijken op "pesterij" kunnen juridisch worden gekwalificeerd als digitale aanranding of smaad.

1.7 Psychologische en reputatieschade

Slachtoffers melden gevoelens van schaamte, paniek en machteloosheid.
De schade is niet financieel, maar existentieel: vertrouwen in omgeving en zelfbeeld verdwijnt.
Organisaties kunnen indirect slachtoffer worden: een gemanipuleerde foto van een medewerker of politicus kan een communicatiecrisis veroorzaken nog vóór de waarheid bekend is.

1.8 Economische motor en darknet-markt

Het verdienmodel is gestroomlijnd:

- lage hostingkosten (open-source-AI op huurservers);
- microbetalingen via crypto;
- automatische verwerking.

Ransomwared onderzoekers zagen advertenties waarin aanbieders **0,000060 BTC (~€ 6)** vroegen per minuut bewerkt videomateriaal.
Dat bewijst dat er een actieve **marktstructuur** bestaat.
Door anonimiteit en cryptotransacties is het risico voor daders minimaal.

1.9 Van foto naar video

Nieuwe modellen verwerken inmiddels **bewegend beeld**.
Korte filmpjes van straat- of winkelcamera's kunnen worden omgezet in gemanipuleerde clips waarin tientallen mensen voorkomen.
De dreiging verschuift zo van individueel misbruik naar **massa-uitbuiting van publieke beelden**.

1.10 Kinderbescherming – het zwaarste risico

Het meest verontrustend is de opkomst van **AI-gegenereerde kindermisbruikbeelden**.
Modellen kunnen jeugdige gezichten of lichamen synthetiseren zonder dat ooit een echt kind is gefotografeerd.
Dat klinkt minder schadelijk, maar vergroot juist de markt en de vraag, en wordt gebruikt om echte kinderen te chanteren.

Sommige landen, waaronder Nederland en Duitsland, hebben het bezit van zulke beelden al strafbaar gesteld, echt of synthetisch.
Internationale harmonisatie blijft echter achter, waardoor daders veilige havens vinden.

1.11 Detectie en verwijdering

De combinatie van snelheid, decentrale opslag en encryptie maakt opsporing moeilijk.
Beelden worden geüpload, gekloond en gedeeld via IPFS of tijdelijke links.

Hash-herkenning faalt zodra bestanden minimaal worden gewijzigd.
Een effectieve aanpak vereist **samenwerking tussen platforms, wetgeving en AI-onderzoek**.

1.12 Regelgeving in beweging

De **EU-AI-verordening** (AI Act) zal "AI-systemen die personen manipuleren door beeldgeneratie zonder toestemming" als hoog risico bestempelen.
De **Digital Services Act (DSA)** verplicht platforms reeds tot snelle verwijdering van niet-consensuele intieme beelden.
Toch is internationale handhaving traag en gefragmenteerd.

1.13 Strategische gevolgen

- **Voor overheden:** ondermijning van vertrouwen in publieke figuren.
- **Voor bedrijven:** merkschade en personeelsrisico's.
- **Voor politie en justitie:** noodzaak tot nieuwe opsporingsmethoden rond AI-bewijs en cryptostromen.
- **Voor cybersecurity-teams:** uitbreiding van het dreigingsbeeld naar menselijk misbruik en perceptie-aanvallen.

1.14 De menselijke factor

Bewustwording is de eerste verdediging.
Mensen moeten beseffen dat een foto uploaden naar een onbekende site vergelijkbaar is met het overhandigen van je identiteitsbewijs aan een onbekende.
Ouders en scholen moeten jongeren uitleggen dat **slachtoffer zijn geen schuld betekent**, en dat er altijd hulp bestaat.

1.15 Aanbevolen acties (samenvatting)

Doelgroep	Belangrijkste actie
Individuen	Beperk het delen van persoonlijke beelden, gebruik privacy-instellingen, zoek regelmatig via reverse-image-search, meld misbruik.
Organisaties	Neem procedures op in het incident-responsplan, train personeel, bied opvang aan slachtoffers.
Platforms	Rate-limiting, perceptuele hashdetectie, snelle verwijderteams.
Beleidsmakers	Harmoniseer wetgeving over synthetische NCII.
Onderzoekers	Deel alleen geanonimiseerde data, geen exploiteerbare modellen.

1.16 Slotbeschouwing

De opkomst van nudify-bots is geen geïsoleerd incident maar een symptoom van een bredere trend:
elke innovatie die creativiteit democratiseert, democratiseert ook misbruik.
De grens tussen kunst en aanranding wordt bepaald door intentie en regelgeving, niet door technologie.

Voor CTI-specialisten betekent dit dat de “menselijke laag” — identiteit, perceptie en toestemming — even belangrijk is geworden als firewalls en antivirus.
Het Ransomwared-team waarschuwt: **de instrumenten om realiteit te vervalsen zijn al vrijgegeven**.
Onze taak is niet om vooruitgang te stoppen, maar om **weerbaarheid en ethiek** op te bouwen.

2. De ‘rontgenbril grap’– Hoe een oude grap werkelijkheid word

2.1 Een onschuldige grap uit vroeger tijden

Iedereen kent hem wel: de grap over die “magische bril” waarmee je iedereen zogenaamd naakt kunt zien.
Je zet hem op, kijkt rond, en lacht. En dan komt de clou: als je hem afzet, blijken de buren nog steeds naakt te zijn — “de bril is stuk.”
Een flauwe, ouderwetse grap, lang onschuldig vermaak.

Maar anno 2025 is dat niet meer grappig.
De “bril” bestaat namelijk echt — niet fysiek, maar digitaal.
Wat ooit een onmogelijke fantasie was, kan nu worden uitgevoerd door kunstmatige intelligentie.

2.2 De metafoor van de bril

In termen van cybersecurity staat die bril symbool voor **technologische capaciteit**, en de ogen erachter voor **menselijke intentie**.

Iedereen beschikt tegenwoordig over een smartphone met een krachtige camera, en vrijwel iedereen heeft toegang tot AI-modellen die beelden kunnen bewerken.
Wat vroeger sciencefiction was, is nu letterlijk binnen handbereik.
De magische bril is vervangen door een app of een webservice die in luttele seconden een “uitgeklede” versie van iemand kan genereren.

Het is een treffende metafoor voor drie kernpunten:

- **Macht:** technologie maakt het mogelijk om visuele waarheid te manipuleren.
- **Toegankelijkheid:** het vergt geen expertise meer, slechts nieuwsgierigheid.
- **Anoniem misbruik:** de gebruiker kan handelen zonder sporen na te laten.

Samen maken die factoren van een grap een wapen.

2.3 Waarom deze vergelijking werkt

Mensen begrijpen complexe dreigingen beter via herkenbare beelden.
Waar termen als “generatieve modellen” of “diffusie-algoritmen” abstract klinken, maakt de brilgrap het tastbaar.
Iedereen voelt intuïtief aan dat **iemand zonder toestemming naakt bekijken** verkeerd is.
Iedereen begrijpt dat een apparaat dat dat mogelijk maakt, gevaarlijk is.
Daarmee is de morele kern direct duidelijk — zonder dat je technische details hoeft te geven.

Voor CTI-teams en beleidsmakers is dit belangrijk:
de bril is niet alleen een metafoor, maar ook een communicatiemiddel om uit te leggen wat deze technologie maatschappelijk betekent.

2.4 Hoe een grap uitmondt in uitbuiting

Zonder in technische details te treden, kan de dynamiek van misbruik als volgt worden samengevat:

1. **De uitnodiging.**
   Op sociale media verschijnt een link: *"Bekijk jezelf in virtuele lingerie"*, *"Ontdek hoe AI je ziet"* of *"Doe mee met de AI-fashion challenge"*.
2. **De upload.**
   De gebruiker — vaak onwetend —maakt een account aan "voor privacy".
   Daarmee deelt ze e-mailadres, IP-adres en soms betaalgegevens.
3. **Het resultaat.**
   Ze ontkleed zich en zet de camera aan zodat de AI een analyse kan maken. Binnen enkele seconden verschijnt een gemanipuleerd beeld dat de gebruiker zogenaamd "in stijl" toont.
4. **De opslag.**
   Zowel het originele als het gemanipuleerde beeld worden bewaard.
5. **De chantage.**
   Enkele dagen later volgt een e-mail: *"Betaal of we publiceren dit."*

Deze werkwijze is meermaals journalistiek en forensisch aangetoond.
Het is pure **social engineering via nieuwsgierigheid** — een psychologische val, geen hack.

2.5 Wie achter de schermen zit

De actoren achter nudify-diensten zijn divers, maar vallen in enkele categorieën:

- **Commerciële exploitanten:** zien het als businessmodel en verkopen toegang via abonnementen of cryptobetalingen.
- **Trollgemeenschappen:** verspreiden gemanipuleerde beelden "voor de lol".
- **Chantagegroepen:** gebruiken de resultaten voor afpersing of reputatievernietiging.
- **Politieke of ideologische actoren:** richten zich op journalisten of publieke figuren om hen in diskrediet te brengen.

Hun infrastructuur — anonieme hosting, cryptowallets, end-to-end-encryptie — overlapt sterk met andere cybercrime-activiteiten.
CTI kan hier waarde toevoegen door deze overlap in kaart te brengen: wallet-analyse, domeinherkomst, en netwerktracering.

2.6 Illustratief scenario

Een voorbeeld uit samengestelde praktijkgevallen:

"Lara, 26 jaar, werkt in marketing."
Ze post regelmatig onschuldige vakantiefoto's op Instagram.

Via een DM ontvangt ze een link naar een “AI-style mirror” waarmee ze haar gewenste outfit kan “zien op een catwalk”.
Ze maakt een account aan en gaat naakt voor de camera van haar Telefoon staan. De app instrueert haar om bepaalde houdingen aan te nemen om de AI het beste beeld te geven.
Twee weken later krijgt ze een e-mail met een expliciete foto van haarzelf en de boodschap: *“We hebben meer. Betaal $200 of het gaat online.”*

De site is inmiddels offline, de domeinhouder onvindbaar.
De expliciete foto duikt kort daarna op een forum op.
Lara’s werkgever ontvangt vragen van journalisten.
De schade is permanent.

Dit voorbeeld laat zien: het slachtoffer hoeft geen publieke figuur te zijn.
Gewoon bestaan op internet is al genoeg.

2.7 Waarom mensen erin trappen

Mensen handelen vaak vanuit nieuwsgierigheid, competitie of sociale druk.
“AI-filters” worden geassocieerd met mode en entertainment.
Daardoor voelen ze niet als gevaarlijk aan.
Criminelen maken daar gebruik van door misbruik te verpakken als trend.

De beloning — likes, aandacht of nieuwsgierigheid — activeert dezelfde dopaminecircuits als gokken of gamen.
Zo ontstaat een **sociale besmetting**, waarin duizenden mensen in korte tijd beelden uploaden zonder na te denken.

2.8 De illusie van toestemming

Gebruikers denken vaak: *“Ik heb de foto zelf geüpload, dus ik heb toestemming gegeven.”*
Maar **toestemming tot uploaden** is niet **toestemming tot bewerking of publicatie**.
Die nuance verdwijnt in gebruiksvoorwaarden vol misleidende zinnen als “AI-fictie, geen echte mensen”.
Juridisch gezien blijft het niet-consensueel.

2.9 Onomkeerbaarheid

Eenmaal online is permanent online.
Zelfs als de dienst beweert beelden te verwijderen, worden ze intussen opgeslagen, gecrawld of gedeeld.
Gemodificeerde hashes zorgen ervoor dat detectiesystemen mislukken.
Zo blijft elk slachtoffer onderdeel van een steeds groter wordend datanetwerk.

Dat is de donkere kant van de brilgrap:
eens gezien, blijft gezien — eens gedeeld, blijft gedeeld.

2.10 De trivialisering van kwaad

Wat ooit uren Photoshop vergde, kan nu met één prompt.
Het morele obstakel — inspanning — is verdwenen.
Hannah Arendt noemde dat de **"banaliteit van het kwaad"**: kwaad wordt gewoonte als het weinig moeite kost.
AI heeft misbruik geautomatiseerd en emotioneel gedistantieerd.
De gebruiker hoeft niet langer te beslissen *of* iets moreel is; de knop doet het voor hem.

Voor CTI is dit een **dreigingsdemocratisering**:
gevaarlijke capaciteiten worden massaal beschikbaar voor laagdrempelig misbruik.

2.11 Het "dual use"-probleem

Generatieve AI is per definitie dubbel inzetbaar.
Hetzelfde model dat artsen helpt of mode simuleert, kan ook worden misbruikt.
Zolang open-source-modellen vrij circuleerden zonder ingebouwde beperkingen, rustte de ethiek bij de gebruiker — een zwakke schakel.
De cybersecuritysector pleit daarom voor **verantwoord modelbeheer**: traceerbare versies, ethische licenties en ingebouwde promptfilters.

2.12 Collaterale schade in datasets

Gemodificeerde beelden belanden uiteindelijk in datasets die door onderzoekers worden gebruikt.
Zo vervuilt misbruikdata de wetenschap.
Wanneer universiteiten per ongeluk zulke beelden in hun trainingssets opnemen, riskeren zij juridische en reputatieschade.
Een enkele nudify-foto kan zo een wereldwijde nasleep krijgen.

2.13 Humor als camouflage

Online circuleren nudify-beelden vaak als memes.
Men lacht, deelt en bagatelliseert.
Die humor werkt als moreel schild: wie lacht, voelt zich niet schuldig.
Zo ontstaat **moraalvervaging**, vergelijkbaar met het vroege internet-trollingtijdperk.
CTI-communicatie moet dat narratief doorbreken: dit is geen grap, maar **digitale aanranding**.

2.14 Organisatorische blindheid

Veel organisaties zien dit als "socialmedia-probleem", niet als veiligheidsincident.
Fout: een gemanipuleerde video van een medewerker kan reputatie, aandelen of politieke besluitvorming beïnvloeden.
Beeldmanipulatie is ook een instrument voor **informatieoorlogvoering**.
Het verdient een plek in threat modeling, naast phishing en malware.

2.15 Wat de bril ons leert

1. **Verlangen bouwt sneller dan ethiek.** Alles wat mensen willen, zal iemand proberen te maken.
2. **Zichtbaarheid is kwetsbaarheid.** Alles wat zichtbaar is, kan worden gekopieerd en verdraaid.
3. **Lachen is geen toestemming.** Humor kan misbruik maskeren, maar neemt de schade niet weg.

2.16 Overgang naar de volgende fase

Dezelfde mechanismen gelden voor video, maar op grotere schaal.
Als de bril vroeger één persoon liet zien, “zien” moderne AI-modellen nu hele menigten.
De volgende hoofdstukken behandelen hoe **video-opnames in straten, winkelcentra en publieke ruimtes** nieuwe vormen van misbruik mogelijk maken — en hoe we daarop kunnen reageren.

3. Wanneer elke camera een potentieel wapen wordt

3.1 De volgende fase van zichtbaarheid

De moderne samenleving leeft in **beweging**: overal hangen camera's, smartphones filmen continu, drones zweven boven evenementen, en zelfs brillen en auto's registreren beelden.
Dat betekent dat er niet alleen meer wordt gezien — maar ook meer **kan worden misbruikt**.

Dezelfde technologie die één foto kan manipuleren, kan nu **videobeelden frame per frame bewerken**.
Een korte opname in een winkelcentrum, park of straat kan worden verwerkt tot een "AI-video" waarin mensen worden uitgekleed of in compromitterende houdingen worden geplaatst.
Wat vroeger een privé-inbreuk was, wordt nu een publiek risico.

3.2 Van momentopname naar continue observatie

Een foto toont één fractie van een seconde.
Een video van 60 seconden bevat honderden beelden — elk een datapunt dat kan worden aangepast.

Voor misbruikers betekent dit:

- **Meer realisme.** Beweging maakt de vervalsing geloofwaardiger.
- **Meer slachtoffers.** Eén opname van een drukke plek kan tientallen mensen tonen die zonder toestemming worden verwerkt.

Vanuit CTI-oogpunt betekent dat een enorme schaalvergroting: één actie kan leiden tot massale reputatie- en privacy-schade.

3.3 De mythe van "openbare ruimte = toestemming"

Veel mensen denken dat wie in een winkelstraat of park loopt, automatisch gefilmd mag worden en dus "geen privacy" heeft.
Dat is een misverstand.
Zichtbaarheid in de openbare ruimte is géén toestemming voor **digitale bewerking** of **seksualisering**.

Een camerabeeld dat legaal is opgenomen kan — zodra het wordt gemanipuleerd — veranderen in **illegaal materiaal**.
Dat onderscheid is cruciaal voor beleidsmakers en juristen: rechtmatig filmen wordt onrechtmatig zodra AI het beeld misbruikt.

3.4 Nieuwe aanvalsvlakken

Vector	Beschrijving	Kans	Impact
Livestream-scraping	Het automatisch downloaden van openbare livestreams (TikTok, YouTube, etc.) en hergebruiken als input.	Hoog	Ernstig; honderden gezichten per stream.
CCTV-lekken	Onbeveiligde beveiligingscamera's die via internet te bekijken zijn.	Middel	Groot; structurele datalekken.
Sociale challenges	Trends waarbij mensen zichzelf filmen op straat ("day in the life") leveren gratis grondstof.	Zeer hoog	Hoog; vooral jongeren kwetsbaar.
Dronebeelden	Stock- of hobbybeelden van festivals en stranden.	Middel	Groot; vaak veel mensen herkenbaar.

De conclusie is eenvoudig: **elke camera is potentieel misbruikbaar**.
Niet het filmen zelf is het probleem, maar het ongecontroleerde gebruik van de beelden erna.

3.5 Voorbeeldscenario: het winkelcentrum

Een videograaf maakt een artistieke opname in een groot winkelcentrum.
Honderden bezoekers komen vluchtig in beeld.
Later uploadt hij zijn montage naar een openbaar videoplatform.

Enkele dagen later duikt de video opnieuw op, bewerkt door een zogenaamde "AI-styling service".
Iedereen in het beeld lijkt nu halfnaakt.
De clip wordt gedeeld in een besloten forum waar gebruikers betalen in **Bitcoin** — €6 per minuut gemanipuleerd materiaal.

Geen van de gefilmde personen weet van het bestaan van die versie.
Toch is hun gezicht zichtbaar, hun lichaam gemanipuleerd, hun waardigheid aangetast.

Voor CTI-onderzoekers is dit scenario relevant omdat de **workflow volledig geautomatiseerd** is:
geen handmatige Photoshop meer, maar één script dat duizenden frames omzet.

3.6 Beweging versterkt geloofwaardigheid

Video's hebben een psychologisch voordeel: ze "voelen echt".
Beweging, schaduw, gezichtsuitdrukking — ons brein gelooft wat het ziet.
AI-modellen gebruiken dat tegen ons.

Zelfs met forensische software is het moeilijk te bewijzen dat een clip nep is.
En tegen de tijd dat experts dat hebben aangetoond, is de video al duizenden keren gedeeld.
Voor slachtoffers is dat onherstelbare reputatieschade.

3.7 Secundaire slachtoffers

Een gemanipuleerde video tast niet alleen de direct afgebeelde persoon aan.
Ook omstanders in hetzelfde beeld kunnen worden verdacht, bespot of betrokken bij online pesterijen.
Dit **besmettingseffect** vergroot de schade en maakt incidentrespons complex: één clip kan tientallen levens beïnvloeden.

3.8 Kruising met kindveiligheid

Publieke ruimtes zijn vol kinderen — speeltuinen, winkelstraten, scholen.
Wanneer videobeelden van zulke plekken worden verwerkt door nudify-AI, ontstaat automatisch het risico op **AI-gegeneerde kinderporno**.
Ook al was het origineel onschuldig, de output is strafbaar.

Er bestaan al darknet-aanbieders die openlijk adverteren met “AI family park compilations” of “beach AI sets” voor crypto.
Die termen verbergen wat het werkelijk is: digitale uitbuiting van minderjarigen.

3.9 Waarom detectie van video nog moeilijker is

Het herkennen van gemanipuleerde foto’s is al lastig; bij video komt daar nog bij:

- **Frame-consistentie.** Moderne modellen zorgen dat belichting en beweging kloppen, waardoor geen zichtbare fouten meer zijn.
- **Compressie.** Door het opnieuw uploaden op sociale platforms verdwijnen forensische sporen.
- **Bestandsgrootte.** Grote bestanden kunnen niet proactief door elk platform worden geanalyseerd.
- **Variabele hashes.** Elke mini-aanpassing verandert de digitale vingerafdruk.

Zonder gezamenlijke dataset en samenwerking tussen platforms is handmatige detectie onbegonnen werk.

3.10 Convergentie met surveillance en commerciële data

Winkels en gemeenten gebruiken videodata voor crowd control en marketing.
Dat levert dagelijks terabytes aan beeld op.
Als zulke systemen worden gehackt of slecht beveiligd, is dat een goudmijn voor misbruikers.

CTI-monitoring heeft al verschillende gevallen gezien van onbeveiligde IP-camera’s met livebeelden van winkelstraten of schoolpleinen.
Eén download en die beelden kunnen voor nudify-training worden misbruikt.

Preventie:

- versleutelde verbindingen,
- wachtwoordrotatie,
- netwerksegmentatie,
- logging van toegang.

Deze maatregelen sluiten aan bij de verplichtingen onder **NIS2** en de **AVG**.

3.11 De economische motor

De ondergrondse markt voor gemanipuleerde video's volgt een bekend patroon:

- **Lage instapkosten.** Open-source-modellen draaien op goedkope GPU's.
- **Schaalbaar product.** Eén opname kan duizenden beelden opleveren.
- **Vaste klantenkring.** Dezelfde afnemers kopen telkens nieuw materiaal.
- **Anonieme betaling.** Crypto en prepaidkaarten verhullen geldstromen.

Door deze schaalbaarheid en anonimiteit blijft het misbruik rendabel, zelfs met lage marges per video.

3.12 Normalisering door gewenning

Hoe vaker mensen geconfronteerd worden met deepfakes of AI-erotiek, hoe minder ze schrikken.
Die **desensitisatie** is gevaarlijk: het maakt misbruik "normaal".
De samenleving verliest empathie voor slachtoffers en verlegt morele grenzen.

Daarom moeten voorlichtingscampagnes niet sensatiegericht zijn, maar **normstellend**:
dit is geen grap of "filterfout", dit is digitale aanranding.

3.13 Verantwoordelijkheid van organisaties

Bedrijven, scholen en gemeenten die video-opnames maken, hebben ook verantwoordelijkheid:

- **Transparantie:** duidelijk aangeven waarom en hoe er gefilmd wordt.
- **Beperk bewaartijd:** verwijder opnames zodra ze niet meer nodig zijn.
- **Snelle verwijdering bij misbruik:** actieve meldprocedures voor deepfakes.
- **Training:** personeel leren herkennen wanneer gemanipuleerd beeld circuleert.

Volgens de **AVG** en **Digital Services Act** kunnen organisaties aansprakelijk worden gesteld als hun beelden bron worden van misbruik.

3.14 Koppeling met CTI-domeinen

Voor CTI-teams is videomanipulatie relevant binnen vier gebieden:

1. **Technisch:** infrastructuur en indicatoren van misbruikdiensten.
2. **Psychologisch:** hoe beelden worden gebruikt voor chantage of desinformatie.
3. **Juridisch:** bewijsbeheer en internationale samenwerking.
4. **Beleid:** adviseren over regelgeving en modeltoegang.

Zo wordt CTI een brug tussen technologie, mens en beleid.

3.15 Verdediging in lagen

Niveau	Maatregel	Doel
Capture	CCTV beveiligen, toegang beperken, encryptie.	Datalekken voorkomen.
AI-verwerking	Watermerken, logging, ethische modellicenties.	Misbruik traceerbaar maken.
Distributie	Snelle meldpunten, API's voor verwijdering.	Verspreiding beperken.
Wetgeving	Internationale samenwerking.	Handhaving verbeteren.
Educatie	Publieke bewustwording.	Slachtoffers beschermen.

Een integrale aanpak combineert techniek, beleid en cultuur.

3.16 Ethische dilemma's

Journalisten en onderzoekers willen waarschuwen, maar niet exploiteren.
Verantwoorde communicatie betekent:

- geen expliciete beelden tonen;
- slachtoffers anonimiseren;
- context geven;
- duidelijke labeling ("synthetisch beeld").

Onderzoekers moeten datasets beveiligen en niet openbaar maken zonder controle.
Transparantie mag nooit ten koste gaan van veiligheid.

3.17 Strategische gevolgen

Door de overvloed aan gemanipuleerde video's begint het publiek **authentieke beelden te wantrouwen**.
Dit heet het "leugenaarsdividend": als alles nep kan zijn, gelooft niemand nog wat echt is.
Dat tast rechtspraak, journalistiek en democratie aan.

Daarom moeten landen investeren in **digitale waarmerkingssystemen**:

- bronhandtekeningen,
- watermerken,
- blockchain-gebaseerde herkomstcontrole.

Zo kan men aantonen dat een video authentiek is.

3.18 Naar realtime misbruik

De volgende technologische stap is **live AI-filtering** — real-time aanpassing van beeld via AR-brillen of livestreams.
Hoewel dit nu nog experimenteel is, toont het een zorgwekkend toekomstbeeld:
misbruik “ter plekke”, zonder dat er ooit een bestand wordt opgeslagen.

Zonder tijdige regelgeving kan dit leiden tot **live-perceptie-aanvallen**, waarbij mensen worden geseksualiseerd op het moment zelf.

3.19 Belangrijkste inzichten

1. Elke camera is potentieel misbruikbaar.
2. AI-gemanipuleerde video ondermijnt geloof in bewijs.
3. Economische prikkels drijven de darknet-markt.
4. Kinderveiligheid is het zwaarst getroffen domein.
5. CTI moet zich richten op **perceptieve veiligheid** — bescherming van waarheid.

4. Het gevaar van AI-gegenereerde kindermisbruikbeelden

4.1 Inleiding

Van alle dreigingen die door generatieve AI mogelijk zijn gemaakt, is er geen enkele zo moreel verwoestend als de productie van **AI-gegenereerde kindermisbruikbeelden** (in het Engels vaak aangeduid als *AI-generated child sexual imagery*, of **AI-CSI**).

Deze synthetische beelden lijken op foto's of video's van kinderen in seksuele contexten, maar worden volledig door kunstmatige intelligentie gemaakt.
Ze worden vaak aangeduid als "niet-echt", omdat er geen direct slachtoffer zou zijn.
Toch veroorzaken ze **echte schade** – psychologisch, maatschappelijk en juridisch.

Ze:

- **onderhouden en vergroten de vraag** binnen kindermisbruiknetwerken;
- **worden gebruikt voor chantage of grooming**, waarbij echte kinderen worden bedreigd;
- **bemoeilijken politiewerk**, omdat duizenden synthetische bestanden echte opsporing vertragen;
- **ondermijnen grenzen** tussen fantasie en criminaliteit, waardoor de drempel voor misbruik lager wordt.

Voor de CTI-gemeenschap betekent dit dat kindbescherming niet langer alleen een sociaal of juridisch domein is, maar ook een **cybersecurity-vraagstuk**.

4.2 Van fotomontage naar machinale verbeelding

Voordat AI bestond, manipuleerden daders bestaande foto's met Photoshop: hoofden van kinderen werden op volwassen lichamen geplakt.
Dat was technisch onvolmaakt en herkenbaar.

Generatieve AI-modellen hebben die grens volledig weggevaagd.
Met zogenoemde *diffusion models* en *fine-tuning* kunnen daders realistische lichamen, gezichten en bewegingen creëren zonder ooit een echte foto te gebruiken.
Een prompt volstaat.

De werkwijze is vergelijkbaar met die van "nudify-bots":

1. **Beeldverzameling** – Foto's van kinderen worden online verzameld, vaak uit openbare profielen of schoolwebsites.
2. **Modeltraining** – Open-source-AI wordt aangepast om kinderlijke kenmerken te genereren.
3. **Synthetische productie** – Nieuwe beelden of video's worden gemaakt op verzoek of in bulk.
4. **Verkoop en distributie** – De resultaten worden verhandeld op darknet-fora in ruil voor cryptovaluta.

Een fysieke misdaad lijkt op het eerste gezicht afwezig, maar het **effect op slachtoffers en samenleving is reëel**.

4.3 De gevaarlijke mythe van “er is geen echt slachtoffer”

Sommigen beweren dat AI-kindermisbruikbeelden minder erg zijn omdat er geen echt kind is gefotografeerd. Dat argument is vals.

- **Echte gezichten worden vaak gebruikt** als basismateriaal.
- **De vraag naar misbruikcontent groeit**: wie gewend raakt aan synthetische beelden, zoekt vaak verder.
- **Synthetische beelden dienen als chantage-materiaal** tegenover echte kinderen (“we hebben al foto’s van je”).
- **Opsporingsdiensten raken overspoeld** door nepbestanden, waardoor echte slachtoffers later worden gevonden.

De schade is dus niet hypothetisch. Ze tast fundamentele normen aan over wat menselijkheid, bescherming en toestemming betekenen.

4.4 De ondergrondse markt

Op het dark web worden al **AI-CSI-diensten** aangeboden, vergelijkbaar met andere cybercriminele markten.

- Aanbieders adverteren met termen als *“custom age synthesis”* of *“fictional animation”*.
- Betalingen verlopen via Bitcoin of Monero.
- Er bestaan escrow-systemen, reviews en per-minuut-prijzen — vaak ongeveer **0,000060 BTC (circa €6)** per minuut gegenereerde video.
- Veel verkopers voegen een zin toe als *“for artistic or fictional purposes only”* om juridische aandacht te ontwijken.

CTI-analyse toont dat deze markten infrastructuur delen met phishing- en ransomwaregroepen: versleutelde communicatie, Tor-hosting en cryptowalletclusters.
Door financiële sporen te volgen, kunnen opsporingsdiensten deze ecosystemen in kaart brengen.

4.5 Juridische lacunes

Hoewel vrijwel elk land het bezit van echte kindermisbruikbeelden verbiedt, is de wetgeving over **synthetische** varianten ongelijkmatig.

Regio / Land	Wettelijke status van AI-CSI	Opmerkingen
Europese Unie (algemeen)	Meestal strafbaar via nationale wetgeving; harmonisatie loopt.	AI-wetgeving zal misbruik verder reguleren.
Nederland / Duitsland	Strafbaar zodra de persoon “op een minderjarige lijkt”, echt of niet.	Gelijkgesteld aan echt beeldmateriaal.
Verenigd Koninkrijk	De Online Safety Act (2024) verbiedt productie en verspreiding van synthetische kindermisbruikbeelden.	Handhaving via Ofcom.
Verenigde Staten	Verboden onder federale obsceniteitswetgeving.	Interpretatie verschilt per staat.

Regio / Land	Wettelijke status van AI-CSI	Opmerkingen
Azië / Oceanië	Sterk wisselend; sommige landen hebben nog geen specifieke bepalingen.	Misbruikers wijken hierheen uit.

Door deze verschillen kunnen daders servers en wallets in “juridische schaduwzones” plaatsen.
Voor CTI betekent dit: grensoverschrijdende samenwerking is essentieel.

4.6 Psychologische en maatschappelijke gevolgen

AI-CSI veroorzaakt schade op vier niveaus:

1. **Individueel:** Slachtoffers van wie de gelijkenis wordt gebruikt, lijden aan angst, schaamte en wantrouwen.
2. **Gezinnen:** Ouders of verzorgers ervaren schuldgevoel en sociale stigmatisering.
3. **Samenleving:** De normalisering van kindermisbruikbeelden verzwakt morele normen.
4. **Instituten:** Scholen, AI-bedrijven en cloudaanbieders lopen reputatie- en aansprakelijkheidsrisico’s.

Het verlies van empathie en vertrouwen in digitale veiligheid raakt het fundament van de samenleving.

4.7 Hoe daders zich verbergen

AI-CSI-netwerken gebruiken technieken die lijken op die van geavanceerde cybercrime:

- **Steganografie:** afbeeldingen worden verstopt in andere bestanden.
- **Tijdelijke hosting:** bestanden verdwijnen automatisch na download.
- **Lokale generatie:** daders gebruiken offline modellen om detectie te vermijden.
- **Codenamen:** projecten worden onschuldig genoemd (“dataset A1” of “research pack”).

Voor CTI-analisten is patroonherkenning in blockchain-transacties en serverstructuren cruciaal om deze verborgen ketens te doorbreken.

4.8 Detectie en ontwrichting

Technisch:

- Ontwikkel AI-detectors die synthetisch misbruik kunnen herkennen zonder legitieme data te schenden.
- Implementeer hash- en vergelijkingsmodellen op uploadplatforms.
- Deel *indicators of compromise* (IOC’s) onder TLP:AMBER-afspraken.

Financieel:

- Traceer cryptotransacties en koppel wallets aan verdachte markten.
- Werk samen met exchanges via *Know Your Transaction*-protocollen.

Infrastructuur:

- Meld en sluit hostingproviders en CDNs die misbruik faciliteren.
- Pas KYC-regels toe op GPU-verhuur en AI-cloudservices.

Wetgeving:

- Maak internationale verdragen die AI-CSI gelijkstellen aan echte kindermisbruikbeelden.
- Stimuleer samenwerking via Europol en Interpol.

4.9 Hulp en slachtofferondersteuning

Wanneer iemand ontdekt dat er AI-gegenereerd misbruikmateriaal van hem of zijn kind circuleert, moet snel en zorgvuldig worden gehandeld:

1. **Verzamel bewijs** – noteer links, tijdstippen, hashes. Download niets.
2. **Meld aan nationale instanties** – zoals het Meldpunt Kinderporno of de politie.
3. **Contacteer hostingplatforms** – met een duidelijk verwijderverzoek.
4. **Zoek psychologische steun** – via Slachtofferhulp of vertrouwenspersonen.
5. **Voorkom verdere verspreiding** – waarschuw scholen of werkgevers.

CTI-teams die zulke meldingen ontvangen moeten empathisch handelen: de menselijke kant gaat vóór technische nieuwsgierigheid.

4.10 Verantwoordelijkheid van AI-ontwikkelaars

Bedrijven en open-sourcegemeenschappen kunnen veel doen om misbruik te voorkomen:

- **Datasets opschonen:** verwijder afbeeldingen van minderjarigen uit trainingsdata.
- **Prompt-filters:** blokkeer expliciete of leeftijdsgebonden opdrachten.
- **Licenties:** geef modellen uit onder ethische voorwaarden.
- **Transparantie:** registreer gebruik zonder beelden op te slaan, zodat misbruik traceerbaar is.

Onder de komende **EU-AI-wet** vallen zulke modellen waarschijnlijk in de categorie *hoog risico* – wat rapportage- en auditplichten met zich meebrengt.

4.11 De rol van CTI-organisaties

Ook al is kindermisbruikbestrijding geen primaire taak van een SOC, CTI-teams kunnen helpen:

- **Threat-hunting:** opsporen van wallets, servers en marktplaatsen.
- **Informatie-uitwisseling:** IOC's delen met politie en NGO's.
- **Opleiding:** analisten trainen in veilig melden en omgaan met gevoelige data.
- **Samenwerking:** brug slaan tussen cybersecurity en kinderbescherming.

Zo draagt CTI bij aan een veiliger internet zonder dat het zelf beelden hoeft te bekijken of op te slaan.

4.12 Verantwoorde communicatie

Journalisten en onderzoekers moeten waarschuwen zonder te sensationaliseren.
Dat betekent:

- geen expliciete beelden tonen;
- slachtoffers beschermen;
- duidelijke taal gebruiken ("AI-gegenereerd misbruikmateriaal" in plaats van "virtuele kinderporno");
- altijd verwijzen naar hulpkanalen.

Sensatie voedt nieuwsgierigheid; nuchtere uitleg bevordert preventie.

4.13 Beleidsaanbevelingen

1. **Gelijke strafbaarheid** voor synthetisch en echt misbruikmateriaal.
2. **Verplichte AI-impactanalyses** bij de release van nieuwe modellen.
3. **Publieke financiering** van detectieonderzoek binnen opsporingsdiensten.
4. **Uitbreiding van meldpunten** voor AI-misbruik.
5. **Campagnes** met de boodschap: *"Gegenereerd of niet, dit is misbruik."*
6. **Aansprakelijkheid van platforms** bij nalatigheid in moderatie.

Een gecombineerde aanpak van wet, technologie en educatie is de enige structurele oplossing.

4.14 Vooruitblik

De volgende generatie AI-tools kan **realtime videomanipulatie** uitvoeren — zelfs tijdens livestreams.
Dan ontstaat een nieuwe uitdaging: bewijsmateriaal dat verdwijnt zodra de stream stopt.
Proactieve maatregelen zijn nodig, zoals:

- **Digitale handtekeningen** in camera's om authenticiteit te waarborgen.
- **Realtime detectie-algoritmes** in streamingdiensten.
- **Hardware-beperkingen** die voorkomen dat beelden van minderjarigen zonder toestemming worden verwerkt.

Door nu te reguleren, voorkomen we dat we straks opnieuw achter de feiten aanlopen.

4.15 Moreel perspectief

De strijd tegen AI-kindermisbruik gaat over meer dan technologie; het gaat over menselijkheid.
Een samenleving die toestaat dat de onschuld van kinderen door algoritmes wordt verhandeld, verliest haar morele kompas.

Elk familiefilmpje, elke schoolfoto en elke vakantieherinnering hoort een herinnering te blijven — geen grondstof voor uitbuiting.
De taak van cybersecurityprofessionals is niet alleen systemen beschermen, maar ook mensen.
Dat vraagt om kennis, empathie en moed.

4.16 Conclusie – een gezamenlijke verantwoordelijkheid

AI-gegenereerde kindermisbruikbeelden vormen een samensmelting van drie gevaren:

1. onbeperkte AI-toegang,
2. economische prikkels voor misbruik,
3. anonimiteit via crypto en darknet.

De aanpak moet even breed zijn als het probleem zelf:

- **Juridisch:** universele strafbaarstelling en versnelde internationale samenwerking.
- **Technisch:** detectie- en watermerk-technologieën.
- **Financieel:** blokkeren van betaalstromen.
- **Sociaal:** onderwijs, bewustwording en hulpverlening.
- **Menselijk:** bescherming van slachtoffers en van onderzoekers die ermee te maken krijgen.

Het Ransomwared CTI-team waarschuwt: dit is geen toekomstmuziek.
Deze markten bestaan nú al.
Wie wegkijkt, maakt zich medeplichtig aan stil leed.

De grap over de bril — ooit onschuldig — is veranderd in een spiegel van onze moraal.
Wat we kiezen te zien, en vooral wat we kiezen níet te zien, zal bepalen hoe humaan de AI-tijdperk werkelijk wordt.

About Ransomwared

Ransomwared is a European-based cybersecurity initiative committed to protecting organizations against the evolving threat of ransomware. Our mission is to **disrupt the economics of cyber extortion** by providing intelligence, technology, and rapid response capabilities that empower defenders to outpace attackers.

At the core of our work is a **next-generation, AI-enhanced, autonomous SOC (Security Operations Center)** that operates 24/7. This SOC continuously ingests global threat intelligence, analyzes attacker behaviors, and autonomously correlates patterns against enterprise telemetry. By leveraging **machine learning models trained on ransomware TTPs**—including those used by groups such as **Akira**—we provide real-time detection, predictive defense, and automated containment actions.

How We Stay Ahead of Threats Like Akira

- **AI-Driven Threat Intelligence**: Our models are continuously refined with data from ransomware campaigns, CVE exploit chains, and underground ecosystems, enabling proactive detection of new attack variants.
- **24/7 Autonomous SOC**: Operating around the clock, our SOC doesn't just monitor—it autonomously correlates anomalies, isolates compromised endpoints, and enforces adaptive security controls in real time.
- **Behavioral Defense**: By mapping techniques to **MITRE ATT&CK**, we detect ransomware campaigns even when adversaries change infrastructure, binaries, or ransom note formats.
- **Continuous Learning**: Every incident enriches our AI and SOC capabilities, strengthening defenses not only for individual organizations but across the entire Ransomwared community.

Our Broader Mission

- **Threat Intelligence Reports**: In-depth CTI reporting (like this Akira analysis) that provides technical, operational, and strategic insights.
- **Vulnerability-to-Exploit Correlation**: Automated pipelines that link CVEs with ransomware campaigns within hours of disclosure.
- **Resilience by Design**: Guidance for implementing Zero Trust, immutable backups, and robust incident response frameworks.

Our Vision

We believe the fight against ransomware will not be won by reacting to incidents, but by **out-automating adversaries**. By combining advanced AI, a 24/7 autonomous SOC, and a culture of open intelligence sharing, Ransomwared helps organizations move from reactive defense to **proactive resilience**—ensuring that ransomware groups like Akira lose their strategic advantage.

For more information, resources, and access to our threat intelligence services, visit:
🌐 **www.ransomwared.eu**