DE LA DÉTECTION AU SIGNALEMENT

Détection de CSAM en ligne : Procédures, mécanismes et limites

Diane Leblanc-Albarel

KU Leuven · COSIC

AFSIN · 15–17 septembre 2026

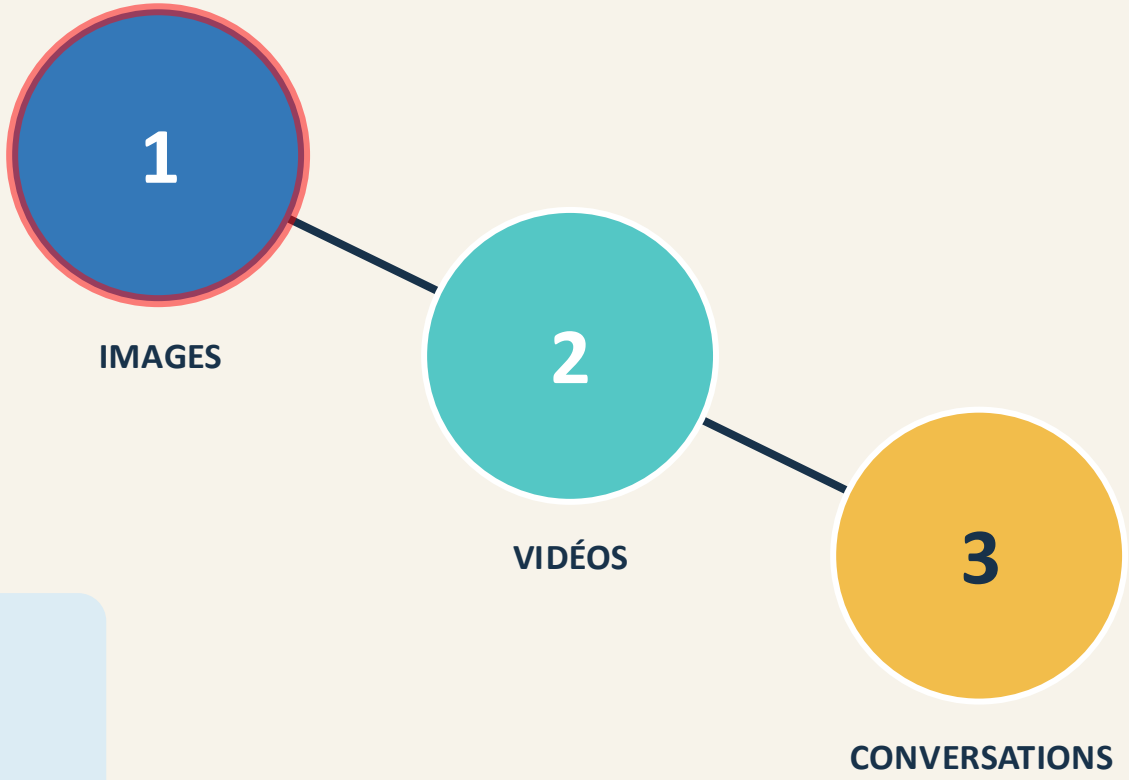
De quoi parle-t-on ?

Des contenus documentant ou représentant l'exploitation sexuelle de mineurs.

CSAM

Child Sexual Abuse Material

Matériel d'abus sexuels sur mineurs



Deux objectifs différents

Retrouver des contenus déjà identifiés
Repérer des contenus potentiellement nouveaux

NCMEC : un centre américain, des signalements mondiaux

Les plateformes américaines doivent lui signaler les cas connus, quel que soit le pays concerné.

1984
CRÉATION
Organisation privée
à but non lucratif

DROIT US
RÔLE LÉGAL
Signalement obligatoire
des plateformes américaines

NCMEC
National Center for Missing
& Exploited Children

1998
SYTEME DE SIGNALEMENT
Ouverture du système
national de signalement

QUEL QUE SOIT LE PAYS
Signalement à NCMEC,
même depuis l'étranger

DOUBLE RÔLE

OPÉRATIONNEL
reçoit, analyse et oriente
les signalements

INFLUENCE
intervient auprès du Congrès et défend
le maintien de capacités de détection

"End-To-End Encryption: Ignoring Abuse Won't Stop It"

De l’utilisateur à une éventuelle enquête

Une détection traverse plusieurs acteurs avant toute qualification judiciaire.

1

UTILISATEUR / CONTENU

Publication ou partage

2

PLATEFORME PRIVÉE

Détecte, transmet un signalement

Reçoit, examine et oriente le rapport

3

AUTORITÉS COMPÉTENTES

Qualifient le rapport et décident d’une éventuelle enquête

4

Une détection peut conduire à un signalement. Ce rapport n’est pas une conclusion judiciaire.

Où le contenu peut-il devenir visible ?

Le chiffrement de bout en bout ne supprime pas tous les points d'observation mais il les déplace.

1 CÔTÉ SERVEUR

Ex. Gmail : détection automatisée possible sur une pièce jointe accessible au service.

2 APRÈS SIGNALEMENT

WhatsApp reçoit jusqu'à 5 messages récents envoyés par le compte signalé.

3 SUR LE TERMINAL

Détection avant chiffrement ou après déchiffrement : client-side scanning potentiel.

Le scan automatique dépend du service

MESSAGERIES PRIVÉES

Facebook

Photos/vidéos publiées ou en direct : scan automatique

Messenger : scan automatique des médias hors E2EE

Instagram

Photos, Reels, vidéos en direct : scan automatique

Messages privés : scan automatique

WhatsApp

Photos de profil/groupe : scan automatique

Messages E2EE : scan après signalement

TikTok

Vidéos publiques/privées/en direct : scan automatique

Messages privés : scan automatique du texte + contenus

Google

Drive, Photos, YouTube : scan automatique des médias

Gmail (paramètre par défaut) : scan automatique d'images

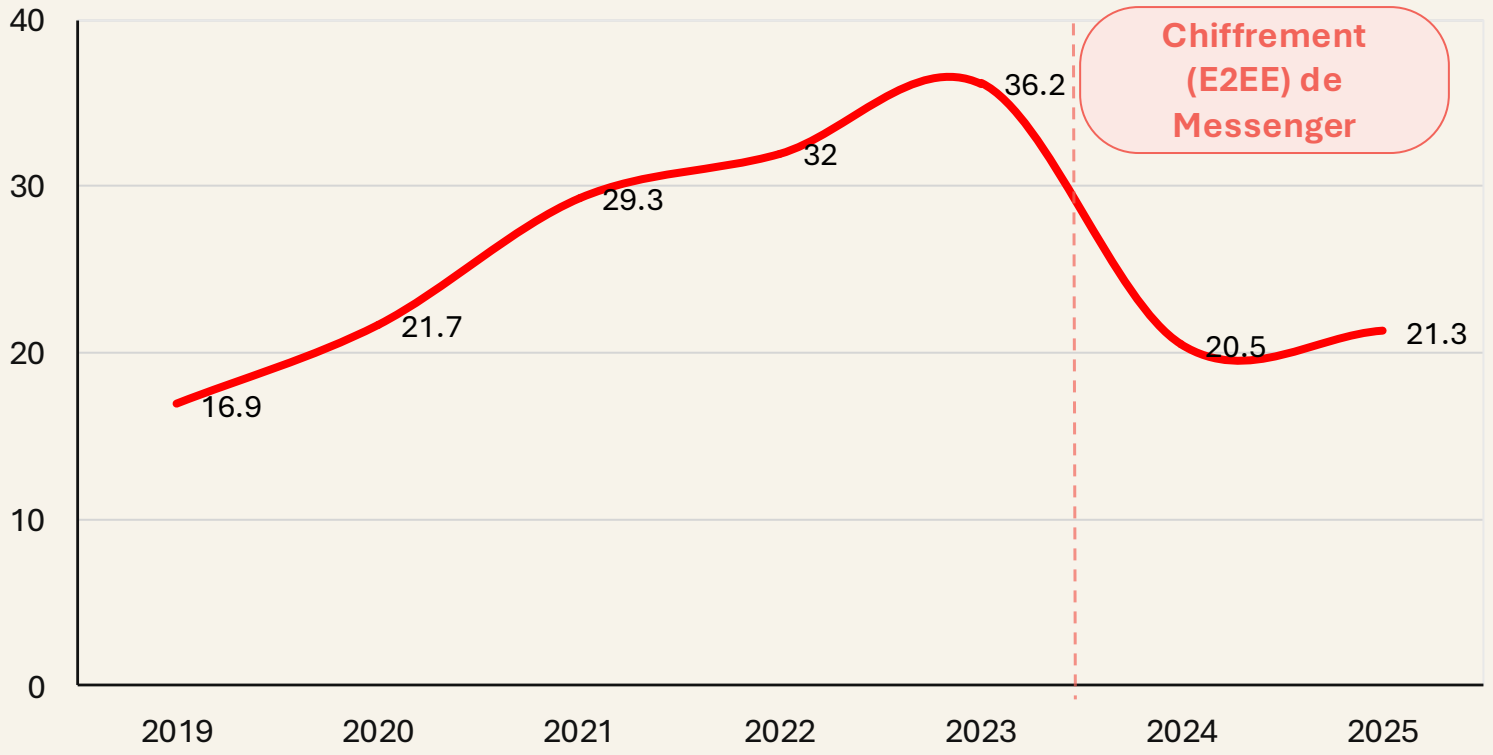
Microsoft

OneDrive, Outlook, Bing, Xbox : scan automatique

Outlook : pièces jointes ; Teams : outils plus ciblés

Les signalements ont changé d'échelle

Rapports annuels reçus par NCMEC, en millions.



NCMEC relie la baisse à l'E2EE de Messenger : rapports Facebook, 17,8 M → 8,6 M → 4,9 M.

< 0,5 M **en 2012**

32 M en 2022 : une multiplication par plus de 60 en dix ans.

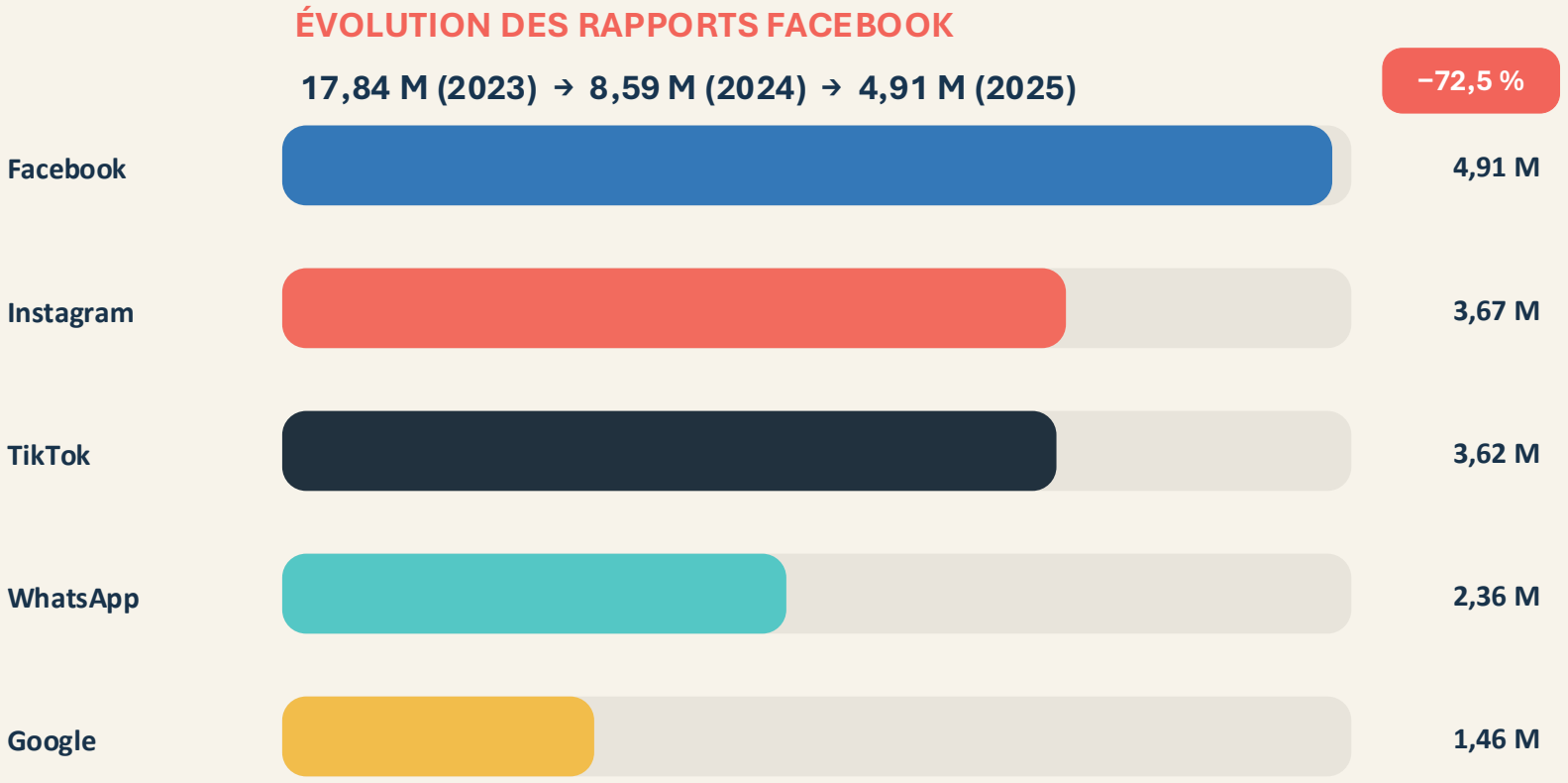
NATIONAL CENTER FOR MISSING & EXPLOITED CHILDREN®

Google Meta Microsoft SNAPCHAT

L'essentiel des rapports vient des plateformes en ligne.

Les volumes sont concentrés entre quelques plateformes

Rapports adressés à NCMEC par les principaux fournisseurs en 2025.



75,6 %

des rapports ESP proviennent de ces cinq fournisseurs

(>2000 ESP in total)

21,18 M rapports → 61,8 M fichiers

29,4 M images

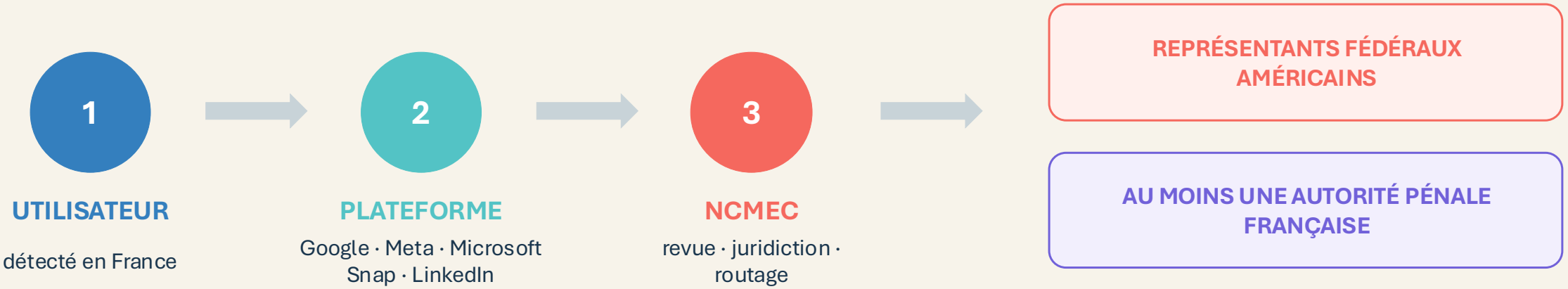
26,3 M vidéos

+ 6,1 M autres fichiers

Un rapport ≠ un fichier ≠ un compte ≠ un utilisateur ≠ une enquête ≠ une condamnation.

Un signalement français reste aussi dans le circuit américain

Les rapports passent par NCMEC et services américains



Détecté en Europe → signalé aux États-Unis → rendu disponible aux autorités compétentes

La durée de vie du rapport reste mal documentée

La règle publiée pour la plateforme ne fixe pas le calendrier de conservation chez NCMEC.

PLATEFORME

1 AN MINIMUM

après le CyberTip, pour le contenu fourni dans le rapport

Certains éléments peuvent être conservés plus longtemps selon les finalités et les règles applicables.

NCMEC

DURÉE MAXIMALE PUBLIQUE

NON TROUVÉE

Aucune règle publique générale trouvée imposant la suppression des rapports finalement non confirmés.

Cela ne démontre pas une conservation permanente ; cela révèle une absence de transparence publique sur le cycle de vie du rapport.

Des traitements largement automatisées

La revue humaine peut souvent intervenir après un traitement automatique.

NCMEC · TRAITEMENT DES RAPPORTS ~ 153 000 / JOUR

IMAGES ET VIDÉOS

DÉTECTION AUTOMATISÉE

FICHIER RECONNU

TRAITEMENT AUTOMATISÉ

FICHIER NON RECONNU

EXAMEN HUMAIN

SI NOUVEL AJOUT À LA BASE :

3 VALIDATIONS HUMAINES

GOOGLE + YOUTUBE · 2025 – 8,8 M – 24 200 / JOUR

99,67 %

AUTOMATIQUE

0,33 %

MANUEL

1

SANCTION / RETRAIT DU CONTENU

2

SIGNALEMENT À NCMEC

SI RECOURS

1. Vérification automatisée
2. Si recours rejeté, vérification humaine et décision finale

Automatisation des deux côtés. la sanction peut préceder le signalement.

En Europe, le scan des messages est permis par une exception

La dérogation autorise temporairement certaines détections volontaires.

RÈGLE

ePRIVACY

Confidentialité des communications privées

EXCEPTION TEMPORAIRE

RÈGLEMENT (UE) 2021/1232

DÉTECTION VOLONTAIRE

Détecter · retirer · signaler
Rétablie jusqu'au 3 avril 2028

Elle autorise certains traitements. Elle n'impose pas un scan général.

Le cadre permanent reste à négocier.

Le CSAR cherche un cadre permanent

Les trois institutions de l'UE convergent sur un Centre européen, mais pas encore sur les moyens.

Obligations des services

Évaluer et réduire les risques
liés à leurs services

Signaler les abus identifiés
au Centre européen

Retirer les contenus
sur ordre des autorités

CSAR

CADRE PERMANENT

Missions du Centre européen

Vérifier les signalements
et écarter les erreurs manifestes

Transmettre les signalements
aux autorités nationales et à Europol

Aider les victimes
à faire retirer les contenus

UN POINT DE CONVERGENCE · DEUX QUESTIONS

CENTRE EUROPÉEN

Réduire la dépendance à NCMEC

QUESTION 1

DÉTECTION OBLIGATOIRE ?

QUESTION 2

ASSURANCE DE L'ÂGE ?

En France : soutenir l'objectif, encadrer les moyens

Deux résolutions parlementaires, puis une position officielle au Conseil de l'UE.

SÉNAT · RÉSOLUTION N° 77

AVIS SUR LE PROJET CSAR

ADOPTÉE LE 20 MARS 2023

« Appelle à la suppression des dispositions relatives à la recherche indifférenciée [...] »

Encadrer les injonctions de détection

- contenus déjà identifiés
- audit indépendant
- chiffrement protégé
- âge respectueux de la vie privée

ASSEMBLÉE · RÉSOLUTION T.A. N° 186

AVIS D'INITIATIVE

ADOPTÉE LE 3 DÉCEMBRE 2025

« L'Europe ne peut plus externaliser la sécurité de ses enfants à un organisme étranger »

Réduire la dépendance à NCMEC

- créer un Centre européen
- moyens de détection suffisants
- protéger la vie privée et les données

FRANCE AU CONSEIL DE L'UE

DÉROGATION TEMPORAIRE

APPROUVÉE LE 23 JUILLET 2026

« Forte préoccupation quant à [...] l'exclusion des communications chiffrées »

DANS LE PÉRIMÈTRE DE LA DÉTECTION VOLONTAIRE

APPROBATION

Volontariat jusqu'en avril 2028

PRÉOCCUPATIONS

- Souhaite ne pas exclure toute détection dans l'E2EE
- Regrette que les difficultés d'accès des enquêteurs aux données restent insuffisamment traitées

PhotoDNA : de Microsoft à un déploiement mondial

Pour ouvrir la boîte technique, commençons par PhotoDNA : une fonction pour les images déjà connues.

PhotoDNA

Microsoft
+ Hany Farid
2009

fonction fermée

FOURNIE À NCMEC

SERVICE CLOUD · 2014

puis déployée auprès de plateformes et d'organisations qualifiées

Retrouver des copies d'images déjà connues

MICROSOFT UNIQUEMENT

12,3 milliards

d'éléments scannés en 2022 par les services couverts

Spécification complète et code non publics · aucun benchmark indépendant historiquement reproductible retrouvé

« PhotoDNA est généralement présentée comme extrêmement fiable. Pourtant, ses affirmations publiques ne sont pas toujours cohérentes entre elles. »

Une affirmation non vérifiée devient une référence

« 1 sur 50 milliards » : une déclaration au Congrès américain devient une information présentée comme vérifiée.

2019 — **Hany Farid**
Congrès américain

> *"False alarm rates for PhotoDNA that we developed a decade ago are about 1 in 50 billion."*

2021 — **EPRS**
PE 662.598

> *"The developer of the technology* ***has testified*** *that the expected error rate […] is 1 in 50 billion."*

2022 — **Commission européenne**
SWD(2022) 209

> *"No more than 1 in 50 billion,* ***based on testing****."*

2023 — **Commission européenne**
COM(2023) 797

> *"PhotoDNA's error rate* ***remains exceedingly low****."*

2025 — **Commission européenne**
COM(2025) 740

> *"The rate of false positives is* ***estimated, on the basis of testing****, to be no more than 1 in 50 billion."*

Le cas irlandais montre les conséquences après la détection

Un signalement écarté peut néanmoins laisser des données dans le circuit.

SEULEMENT

20 %

considérés comme constituant du CSAM

> 11 % de faux signalements

≈ 69 % : ni décrits comme CSAM avéré ni comptés comme faux signalements

> 11 % DES RAPPORTS

SIGNALEMENTS ÉCARTÉS

photos familiales ou de plage citées

APRÈS LA MISE HORS DE CAUSE

Données toujours conservées

selon l'Irish Council for Civil Liberties (ICCL)

Le dommage potentiel peut apparaître après la détection : transmission, qualification et conservation.

On ne sait pas quel mécanisme de détection a déclenché chacun de ces rapports.

Meta et Google : des erreurs constatées en pratique

Deux indicateurs : décisions corrigées après recours et détections non confirmées.

META · DONNÉES 2023

≈ 3,6 M

contenus ayant fait l'objet d'une action

RECOURS

réexamen

≈ 0,32 %

contenus restaurés ou mesures sur comptes annulées

≈ 1 sur 310

GOOGLE · HACHAGE · 2024

1 834

contenus détectés comme CSAM

dénominateur déclaré

EXAMEN

humain

0.54 %

10 contenus non confirmés comme CSAM

≈ 1 sur 183

Pourquoi une évaluation indépendante ?

En sécurité, on teste une garantie en essayant réellement de mettre le système en défaut.

AFFIRMATIONS PUBLIQUES

« 1 sur 50 milliards »

DOCUMENTS PUBLICS

preuves non reproductibles
+ informations contradictoires

→

CRYPTANALYSE

attaquer le
système déployé

→

ATTAQUE RÉUSSIE

limite démontrée

Échec d'une attaque ≠ preuve de sécurité

White-Box Attacks on PhotoDNA

Deryck · Leblanc-Albarel · Preneel

ia.cr/2026/486

pseudodna.eu

DIVULGATION RESPONSABLE

> 6 MOIS D'ÉCHANGES ET DE RÉUNIONS
avec Microsoft, dont le créateur de PhotoDNA côté Microsoft
Attaques sur un PC personnel standard
Applicabilité au système déployé confirmée · contre-mesures discutées

Du hachage cryptographique au hachage perceptuel

Deux photos presque identiques : SHA-256 change complètement, le hash perceptuel reste très proche.

Photos presque identiques

≈

SHA-256

8A1F…D04C

≠

TOTALEMENT DIFFÉRENTS

SHA-256

F93B…117E

Hash perceptuel

8B31…EA76

≈

TRÈS PROCHES

Hash perceptuel

8B32…EA76

La similarité devient une distance

Plus les empreintes sont proches, plus la distance calculée est faible.

1 NORMALISER L'IMAGE → 2 EXTRAIRE LES CARACTÉRISTIQUES → 3 CRÉER L'EMPREINTE → 4 CALCULER LA DISTANCE → 5 COMPARER AU SEUIL

SEUIL

distance faible — distance élevée

originale | peu modifiée | modifiée | image différente | image très différente

DÉTECTÉ — NON DÉTECTÉ

Pourquoi des empreintes perceptuelles ?

Reconnaître les variantes et limiter les fuites d'information

IMAGES OU HASH CRYPTOGRAPHIQUES

≠

Images différentes, hash différents

1. Comparaison exacte uniquement

Aucune reconnaissance des images proches.
TOUTES les variantes à enregistrer,
sous forme d'images ou de hash cryptographiques.

2. Images : fuite d'information

Une fuite expose directement
les contenus et les victimes

EMPREINTES PERCEPTUELLES

Liste NCMEC : 12,1 M

Safer / Thorn : ≈ 175 M

Tout signalement est envoyé à NCMEC

Reconnaissance par similarité

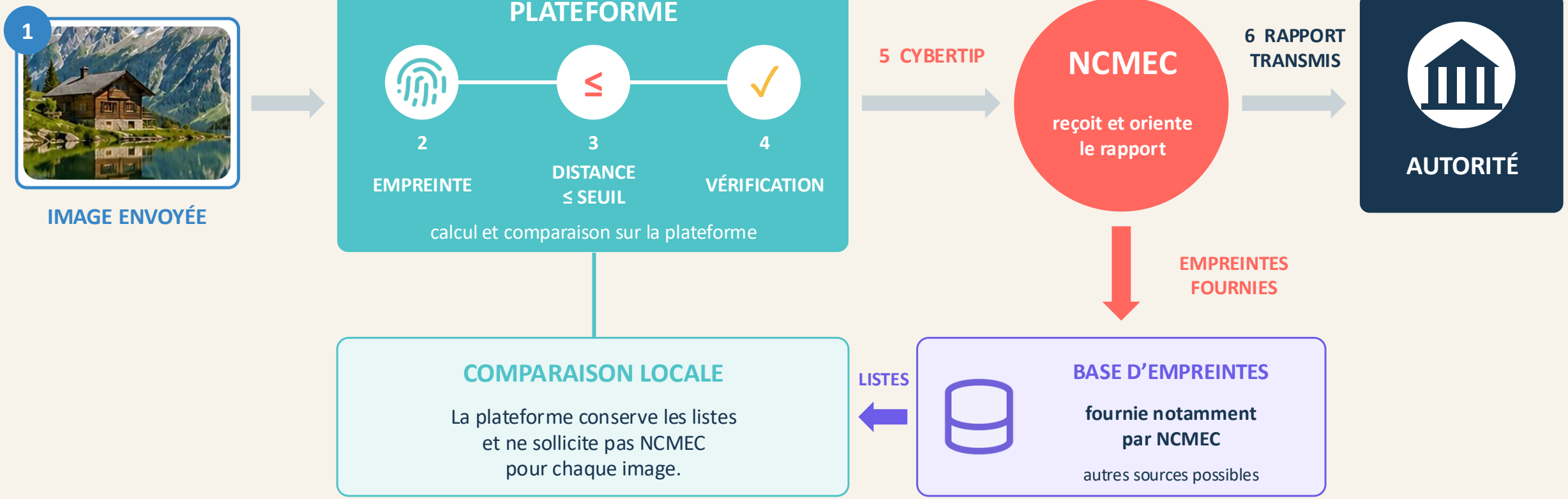
Des images proches donnent
des empreintes proches

Protection des images

La liste ne contient pas
les images originales

Le hachage perceptuel dans le circuit de signalement

La plateforme calcule et compare localement ; NCMEC reçoit un rapport seulement après une détection positive.



Quatre attaques contre PhotoDNA

Empreintes proches : couleurs proches. Empreintes éloignées : couleurs différentes.

Fuite d'information

INVERSION

Image reconstruite

Identique

Attaque sur PhotoDNA

Empreinte connue

Empoisonnement de la base de données

COLLISION

Deux images de départ


Attaque sur PhotoDNA


Modifiée Modifiée

Empreintes très proches ou égales

Fausse accusation

SECONDE PRÉIMAGE

Image cible

Attaque sur PhotoDNA


Intacte Construite

Cible fixée

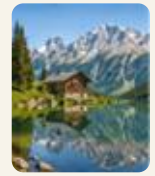
Très proche ou égale

Détection évitée

ÉVASION

Image détectée

Attaque sur PhotoDNA

Originale Modifiée

Empreintes éloignées

Schémas illustratifs sur des images bénignes

1. Inversion du hash : l'empreinte révèle des informations

Scénario opérationnel : un tiers obtient un hash sensible et en extrait une information visuelle grossière.

HASH ACCESSIBLE

base copiée, fuite ou stockage sur le terminal

→

CALCUL DE L'ATTAQUE

chercher une image qui reproduit l'empreinte

→

INFORMATION RÉVÉLÉE

catégorie, silhouette ou structure grossière

99,5 %

de réussite

11,2 s

en moyenne

Conséquence possible : Retrouver de l'information sur des images sensibles.

2. Collision : empoisonnement de base de données

Une empreinte ajoutée à la base peut faire détecter une autre image.

1 · COLLISION

L'attaquant crée X' et Y' avec la même empreinte

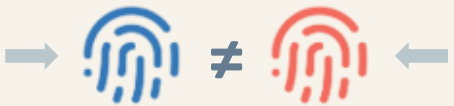
2 · BASE DE DONNÉES

Il ajoute l'empreinte de X' à la base de référence

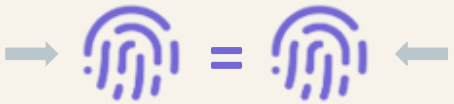
3 · FAUX POSITIF

Y' peut alors être détectée à tort

Images originales



Images modifiées



100 %

collisions exactes réussies

215 s

temps moyen

3. Seconde préimage : un faux positif ciblé

Cette fois, l'attaquant part d'une image ou d'un hash illicite déjà connu et vise une victime précise.

1 · EMPREINTE CIBLE

empreinte d'un contenu illicite connu

2 · IMAGE LICITE

modifiée pour approcher l'empreinte cible

3 · VICTIME

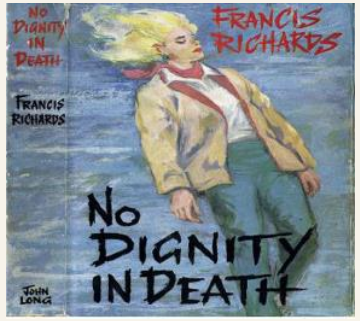
l'image arrive sur son compte ou son appareil

4 · MATCH

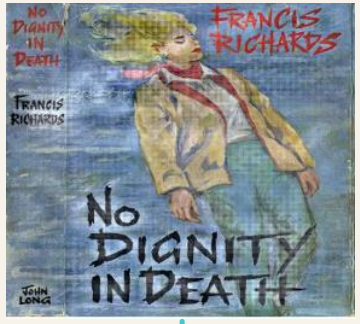
retrait, suspension ou signalement possible

EMPREINTE CIBLE

IMAGE LICITE



APRÈS ATTAQUE



apparence à préserver

≈

Correspondance

99,7 %

Réussite

≈ 118 s

Exécution

distance faible — distance élevée

seuil τ

$d \leq \tau \rightarrow$ MATCH

Conséquence possible : une image visuellement licite peut être traitée comme le contenu illicite ciblé.

4. Évasion : faire passer un contenu sous le seuil

Un criminel modifie légèrement l'image avant de la partager ; la plateforme ne retrouve plus le hash connu.

CONTENU CONNU

son empreinte figure déjà dans la blocklist

→

MODIFICATION MINIMALE

avant le partage

→

AUCUN MATCH

distance au-dessus du seuil
→ pas de signalement

IMAGE CONNUE

empreinte dans la blocklist

→

APRÈS ATTAQUE

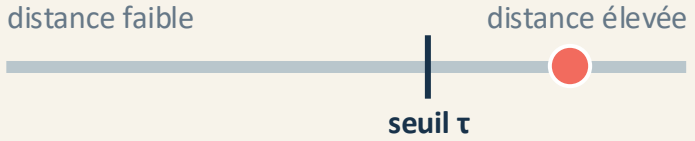
apparence proche

100 %

Réussite

9,94 s

Temps moyen



d > τ → AUCUN MATCH

Conséquence possible : le contenu circule sans être détecté.

Ce que ces attaques remettent en cause

Des conséquences concrètes pour la détection des CSAM

RÉSULTATS SUR PHOTODNA	CONSÉQUENCES OPÉRATIONNELLES POSSIBLES
Inversion Retrouver de l'information visuelle	**Fuite d'empreintes** lors d'une **cyberattaque** Extraction d'informations sur des images sensibles
Collision et seconde préimage Provoquer de faux positifs	Afflux de **signalements d'images innocentes** **Surcharge** des équipes de traitement
Évasion Éviter la détection d'images illégales	Circulation de **contenus illégaux** Absence de **signalement** ou de **poursuites**
Confirmation de Microsoft **Fonction étudiée identique à celle déployée en production**	

Quelles garanties vérifiables à grande échelle ?

31

La détection des CSAM repose sur une infrastructure mondiale

Elle combine une très grande échelle, une place centrale pour NCMEC et des mécanismes techniques peu documentés.

À L'ÉCHELLE MONDIALE

- les volumes de détection et de signalement ont fortement augmenté
- des centaines de milliards d'images sont scannées chaque année
- les grandes plateformes alimentent un circuit mondial centré sur NCMEC

DÉPENDANCE À NCMEC

- les plateformes adressent leurs rapports à NCMEC
- les CyberTips peuvent contenir des données personnelles sensibles
- plusieurs techniques restent propriétaires, sans évaluation publique reproductible
- l'Europe cherche un cadre permanent

PHOTODNA : NOS RÉSULTATS

- **ATTAQUES EN QUELQUES SECONDES**
- **IMPACT PRATIQUE CONFIRMÉ**

NOTRE CONCLUSION

- Indépendamment du débat : les mécanismes déployés ne sont pas fiables
- De la sécurité par l'obscurité à des mécanismes vérifiables et auditables par les pouvoirs publics

Questions ouvertes :
Peut-on minimiser à la fois les faux positifs et les faux négatifs ?
Comment définir « perceptuellement différent » ?

Contact / Webpage

diane.leblanc-albarel@esat.kuleuven.be

https://dianeleblancalbarel.github.io/

Références (1/18)

SLIDE 2 **De quoi parle-t-on ?**

01 **NCMEC — Child Sexual Abuse Material.**
https://www.missingkids.org/theissues/csam

02 **Règlement (UE) 2021/1232 — dérogation temporaire à la directive ePrivacy.**
https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32021R1232

SLIDE 3 **NCMEC : un centre américain, des signalements mondiaux**

01 **NCMEC — History.**
https://www.missingkids.org/footer/about/history

02 **NCMEC — 2025 CyberTipline Report et archives annuelles.**
https://www.missingkids.org/gethelpnow/cybertipline/cybertiplinedata

03 **18 U.S.C. §2258A — contenu des CyberTips et obligations des fournisseurs.**
https://uscode.house.gov/view.xhtml?edition=prelim&num=0&req=granuleid%3AUSC-prelim-title18-section2258A

04 **NCMEC — End-to-End Encryption : position publique sur le chiffrement.**
https://www.missingkids.org/theissues/end-to-end-encryption

Références (2/18)

SLIDE 4 De l’utilisateur à une éventuelle enquête

01 **18 U.S.C. §2258A — contenu des CyberTips et obligations des fournisseurs.**
https://uscode.house.gov/view.xhtml?edition=prelim&num=0&req=granuleid%3AUSC-prelim-title18-section2258A

02 **NCMEC — CyberTipline : fonctionnement et routage des rapports.**
https://www.missingkids.org/gethelpnow/cybertipline

SLIDE 5 Où le contenu peut-il devenir visible ?

01 **Google — rapport UE 2023.**
https://storage.googleapis.com/transparencyreport/report-downloads/pdf-report-23_2023-1-1_2023-12-31_en_v1.pdf

02 **EDPB–EDPS — Joint Opinion 04/2022.**
https://www.edpb.europa.eu/system/files/documents/2022-07/edpb_edps_jointopinion_202204_csam_en_0.pdf

03 **WhatsApp — Signaler un contact : jusqu’à cinq messages récents.**
https://faq.whatsapp.com/414631957536067

SLIDE 6 Le scan automatique dépend du service

01 **Meta — EU CSAM Derogation Report 2025.**
https://transparency.meta.com/sr/eu-csam-derogation-report-2025/

Références (3/18)

SLIDE 6 **Le scan automatique dépend du service**

02 **eSafety Commissioner — Basic Online Safety Expectations transparency report, décembre 2022.**
https://www.esafety.gov.au/sites/default/files/2022-12/BOSE%20transparency%20report%20Dec%202022.pdf

03 **eSafety Commissioner — Basic Online Safety Expectations full transparency report, octobre 2023.**
https://www.esafety.gov.au/sites/default/files/2024-03/Basic-Online-Safety-Expectations-Full-Transparency-Report-October-2023_0.pdf

04 **eSafety Commissioner — Periodic notice Report 2, janvier–juin 2025.**
https://www.esafety.gov.au/industry/basic-online-safety-expectations/child-sexual-exploitation-and-abuse-material-and-activity/periodic-notice-report-2-snapshot

05 **WhatsApp — Information for Law Enforcement Authorities.**
https://faq.whatsapp.com/444002211197967

06 **Google — Efforts to Combat Online Child Sexual Abuse Material: FAQ.**
https://support.google.com/transparencyreport/answer/10330933

07 **Microsoft — Digital Safety Content Report 2025.**
https://www.microsoft.com/en-us/corporate-responsibility/digital-safety-content-report

08 **Google — rapport UE 2023.**
https://storage.googleapis.com/transparencyreport/report-downloads/pdf-report-23_2023-1-1_2023-12-31_en_v1.pdf

Références (4/18)

SLIDE 6 **Le scan automatique dépend du service**

09 **Microsoft — rapport UE 2022.**
https://cdn-dynmedia-1.microsoft.com/is/content/microsoftcorp/microsoft/final/en-us/microsoft-brand/documents/Microsoft-Transparency-Report%282022%29-EU-CSAM-Interim-Regulatory-Report.pdf

SLIDE 7 **Les signalements ont changé d'échelle**

01 **NCMEC — Transparency Report CY 2025, annexe A : séries annuelles et méthode de comptage.**
https://www.missingkids.org/content/dam/missingkids/pdfs/OJJDP-NCMEC-Transparency-CY-2025.pdf

02 **NCMEC — 2023 CyberTipline Reports by Electronic Service Providers.**
https://www.missingkids.org/content/dam/missingkids/pdfs/2023-reports-by-esp.pdf

03 **NCMEC — 2024 CyberTipline Reports by Electronic Service Providers.**
https://www.missingkids.org/content/dam/missingkids/pdfs/2024-reports-by-esp.pdf

04 **NCMEC — 2025 CyberTipline Reports by Electronic Service Providers.**
https://www.missingkids.org/content/dam/missingkids/pdfs/2025-reports-by-esp.pdf

05 **NCMEC — 2025 CyberTipline Report et archives annuelles.**
https://www.missingkids.org/gethelpnow/cybertipline/cybertiplinedata

06 **NCMEC — Protecting Our Nation's Children since 1984 (données 2019).**
https://www.missingkids.org/blog/2020/protecting-our-nations-children-since-1984

Références (5/18)

SLIDE 8 **Les volumes sont concentrés entre quelques plateformes**

01 **NCMEC — 2025 CyberTipline Report et archives annuelles.**
https://www.missingkids.org/gethelpnow/cybertipline/cybertiplinedata

02 **NCMEC — 2025 CyberTipline Reports by Electronic Service Providers.**
https://www.missingkids.org/content/dam/missingkids/pdfs/2025-reports-by-esp.pdf

03 **NCMEC — 2023 CyberTipline Reports by Electronic Service Providers.**
https://www.missingkids.org/content/dam/missingkids/pdfs/2023-reports-by-esp.pdf

04 **NCMEC — 2024 CyberTipline Reports by Electronic Service Providers.**
https://www.missingkids.org/content/dam/missingkids/pdfs/2024-reports-by-esp.pdf

SLIDE 9 **Un rapport peut partir de plusieurs types de détection**

01 **Meta — Preventing Child Exploitation on Our Apps, 23 février 2021.**
https://about.fb.com/news/2021/02/preventing-child-exploitation-on-our-apps/

02 **Meta — EU CSAM Derogation Report 2024.**
https://transparency.meta.com/sr/eu-csam-derogation-report-2024/

03 **Commission européenne — rapport d'évaluation COM(2025) 740.**
https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52025DC0740

Références (6/18)

SLIDE 10 **Un signalement français reste aussi dans le circuit américain**

01 **Google — rapport UE 2023.**
https://storage.googleapis.com/transparencyreport/report-downloads/pdf-report-23_2023-1-1_2023-12-31_en_v1.pdf

02 **Meta — coopération avec les forces de l'ordre.**
https://transparency.meta.com/en-gb/policies/improving/working-with-law-enforcement/

03 **Microsoft — rapport UE 2022.**
https://cdn-dynmedia-1.microsoft.com/is/content/microsoftcorp/microsoft/final/en-us/microsoft-brand/documents/Microsoft-Transparency-Report%282022%29-EU-CSAM-Interim-Regulatory-Report.pdf

04 **Snap — rapport UE 2024.**
https://values.snap.com/privacy/transparency/european-union-h2-2024?lang=en-GB

05 **LinkedIn — rapport UE.**
https://www.linkedin.com/help/linkedin/answer/a1347128

06 **NCMEC — France: Data Insights.**
https://www.globalchildexploitationpolicy.org/data-insights/france

07 **NCMEC — CyberTipline : fonctionnement et routage des rapports.**
https://www.missingkids.org/gethelpnow/cybertipline

08 **18 U.S.C. §2258A — contenu des CyberTips et obligations des fournisseurs.**
https://uscode.house.gov/view.xhtml?edition=prelim&num=0&req=granuleid%3AUSC-prelim-title18-section2258A

Références (7/18)

SLIDE 11 **La durée de vie du rapport reste mal documentée**

01 **18 U.S.C. §2258A — contenu des CyberTips et obligations des fournisseurs.**
https://uscode.house.gov/view.xhtml?edition=prelim&num=0&req=granuleid%3AUSC-prelim-title18-section2258A

02 **NCMEC — Privacy Policy : sécurité, conservation et transferts.**
https://www.missingkids.org/footer/privacypolicy

03 **NCMEC — REPORT Act et préservation par les plateformes.**
https://www.missingkids.org/blog/2024/first-line-of-defense-guidelines-to-help-online-platforms-detect-sexually-exploited-kids

SLIDE 12 **Des traitements largement automatisées**

01 **Google — CSAM Transparency Report : détection automatique et manuelle, 1er semestre 2025.**
https://transparencyreport.google.com/child-sexual-abuse-material/reporting?hl=en&lu=total_content_reported&total_content_reported=detectionmethod:DETECTION_METHOD_AUTOMATED;period:2025H1

02 **Google — CSAM Transparency Report : détection automatique et manuelle, 2e semestre 2025.**
https://transparencyreport.google.com/child-sexual-abuse-material/reporting?hl=en&lu=total_content_reported&total_content_reported=detectionmethod:DETECTION_METHOD_AUTOMATED;period:2025H2

Références (8/18)

SLIDE 12 **Des traitements largement automatisées**

03 **Google — Efforts to Combat Online Child Sexual Abuse Material: FAQ.**
https://support.google.com/transparencyreport/answer/10330933?hl=en-au

04 **Google — Traitement des contenus, sanctions et examen des recours.**
https://support.google.com/product-documentation/answer/15464420?hl=en

05 **NCMEC — Transparency Report CY 2025 : images, vidéos et traitement des signalements.**
https://www.missingkids.org/content/dam/missingkids/pdfs/OJJDP-NCMEC-Transparency-CY-2025.pdf

06 **NCMEC — CyberTipline 2025 : tri des fichiers et partage des empreintes.**
https://www.missingkids.org/gethelpnow/cybertipline/cybertiplinedata

SLIDE 13 **En Europe, le scan des messages est permis par une exception**

01 **Directive 2002/58/CE (ePrivacy), article 5 : confidentialité des communications.**
https://eur-lex.europa.eu/legal-content/FR/TXT/?uri=CELEX:32002L0058

02 **Règlement (UE) 2021/1232 — dérogation temporaire à la directive ePrivacy.**
https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32021R1232

03 **Règlement (UE) 2024/1307.**
https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32024R1307

04 **Conseil de l'UE — 23 juillet 2026.**
https://www.consilium.europa.eu/en/press/press-releases/2026/07/23/fighting-child-sexual-abuse-online-interim-measure-protecting-children-now-reinstated/

Références (9/18)

SLIDE 14 **Le CSAR cherche un cadre permanent**

01 **Commission européenne — proposition CSAR COM(2022) 209.**
https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52022PC0209

02 **Conseil de l'UE — position du 26 novembre 2025 sur la proposition CSAR.**
https://www.consilium.europa.eu/en/press/press-releases/2025/11/26/child-sexual-abuse-council-reaches-position-on-law-protecting-children-from-online-abuse/

03 **Conseil de l'UE — mandat ST 15318/2025.**
https://data.consilium.europa.eu/doc/document/ST-15318-2025-INIT/en/pdf

04 **Commission européenne — cadre juridique.**
https://home-affairs.ec.europa.eu/policies/internal-security/protecting-children-sexual-abuse/legal-framework-protect-children_en

05 **Parlement européen — mandat de novembre 2023.**
https://www.europarl.europa.eu/news/en/press-room/20231110IPR10118/child-sexual-abuse-online-effective-measures-no-mass-surveillance

06 **EDPB–EDPS — Joint Opinion 04/2022.**
https://www.edpb.europa.eu/system/files/documents/2022-07/edpb_edps_jointopinion_202204_csam_en_0.pdf

07 **Parlement européen — Legislative Train : état du dossier CSAR.**
https://www.europarl.europa.eu/legislative-train/spotlight-JD22/file-combating-child-sexual-abuse-online

Références (10/18)

SLIDE 15 En France : soutenir l'objectif, encadrer les moyens

01 **Sénat — Résolution européenne n° 77, 20 mars 2023.**
https://www.senat.fr/leg/tas22-077.html

02 **Assemblée nationale — Résolution européenne T.A. n° 186, 3 décembre 2025.**
https://www.assemblee-nationale.fr/dyn/17/textes/l17t0186_texte-adopte-seance

03 **Conseil de l'UE — Vote et déclaration française, CM 3756/1/26 REV 1, 23 juillet 2026.**
https://data.consilium.europa.eu/doc/document/CM-3756-2026-REV-1/fr/pdf

04 **Sénat — Commission des affaires européennes, 15 février 2023.**
https://www.senat.fr/compte-rendu-commissions/20230213/europ.html

05 **Assemblée nationale — Commission des affaires européennes, 15 octobre 2025.**
https://www.assemblee-nationale.fr/dyn/17/comptes-rendus/due/l17due25260042_compte-rendu.pdf

06 **Sénat — Règlement, article 73 quinquies C : procédure des résolutions européennes.**
https://www.senat.fr/reglement/reglement52.html

07 **Assemblée nationale — Règlement, articles 151-5 à 151-7.**
https://www.assemblee-nationale.fr/dyn/17/divers/texte_reference/02_reglement_assemblee_nationale.pdf

SLIDE 16 PhotoDNA : de Microsoft à un déploiement mondial

01 **Microsoft — PhotoDNA : présentation et affirmations publiques.**
https://www.microsoft.com/en-us/photodna

Références (11/18)

SLIDE 16 PhotoDNA : de Microsoft à un déploiement mondial

02 **Microsoft — PhotoDNA Cloud Service : annonce du 15 juillet 2015.**
https://news.microsoft.com/features/microsofts-photodna-protecting-children-and-businesses-in-the-cloud/

03 **Microsoft — rapport UE 2022.**
https://cdn-dynmedia-1.microsoft.com/is/content/microsoftcorp/microsoft/final/en-us/microsoft-brand/documents/Microsoft-Transparency-Report%282022%29-EU-CSAM-Interim-Regulatory-Report.pdf

04 **Deryck, Leblanc-Albarel & Preneel — White-Box Attacks on PhotoDNA (2026).**
https://eprint.iacr.org/2026/486

SLIDE 17 Une affirmation non vérifiée devient une référence

01 **U.S. Government Publishing Office — hearing transcript, 16 October 2019.**
https://www.govinfo.gov/content/pkg/CHRG-116hhrg43533/html/CHRG-116hhrg43533.htm

02 **EPRS — Targeted impact assessment 2021.**
https://www.europarl.europa.eu/RegData/etudes/STUD/2021/662598/EPRS_STU%282021%29662598_EN.pdf

03 **Commission européenne — SWD(2022) 209.**
https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52022SC0209

04 **Commission européenne — rapport COM(2023) 797 sur le règlement intérimaire.**
https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52023DC0797

05 **Commission européenne — rapport d'évaluation COM(2025) 740.**
https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52025DC0740

Références (12/18)

SLIDE 18 **Le cas irlandais montre les conséquences après la détection**

01 **ICCL — An Garda Síochána unlawfully retains files on innocent people, 2022.**
https://www.iccl.ie/news/an-garda-siochana-unlawfully-retains-files-on-innocent-people-who-it-has-already-cleared-of-producing-or-sharing-of-child-sex-abuse-material/

02 **EDRi — Irish story on online scanning, 2022.**
https://edri.org/our-work/breaking-irish-story-shows-that-eus-csam-proposal-can-never-work/

03 **Irish Legal News — Gardaí retaining data of people cleared.**
https://www.irishlegal.com/articles/gardai-retaining-data-of-people-cleared-of-child-sex-abuse-suspicion-without-legal-basis

SLIDE 19 **Meta et Google : des erreurs constatées en pratique**

01 **Commission européenne — COM(2025) 740, § 2.1.6, tableau 5 : Meta 2023 et Google 2024.**
https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52025DC0740

02 **Meta — EU CSAM Derogation Report 2024.**
https://transparency.meta.com/sr/eu-csam-derogation-report-2024/

SLIDE 20 **Pourquoi une évaluation indépendante ?**

01 **Deryck, Leblanc-Albarel & Preneel — White-Box Attacks on PhotoDNA (2026).**
https://eprint.iacr.org/2026/486

Références (13/18)

SLIDE 20 **Pourquoi une évaluation indépendante ?**

02 **PseudoDNA — démonstrations et implications pratiques des attaques.**
https://www.pseudodna.eu/

03 **Échanges de divulgation coordonnée et réunions avec Microsoft — archives des auteurs.**
Document non public — archives des auteurs.

SLIDE 21 **Du hachage cryptographique au hachage perceptuel**

01 **NIST — FIPS 180-4, Secure Hash Standard (SHA-256).**
https://csrc.nist.gov/pubs/fips/180-4/upd1/final

02 **Commission européenne — rapport d'évaluation COM(2025) 740.**
https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52025DC0740

03 **Microsoft — PhotoDNA : présentation et affirmations publiques.**
https://www.microsoft.com/en-us/photodna

04 **Deryck, Leblanc-Albarel & Preneel — White-Box Attacks on PhotoDNA (2026).**
https://eprint.iacr.org/2026/486

SLIDE 22 **La similarité devient une distance**

01 **Microsoft — PhotoDNA : présentation et affirmations publiques.**
https://www.microsoft.com/en-us/photodna

Références (14/18)

SLIDE 22 La similarité devient une distance

02 **Commission européenne — rapport d'évaluation COM(2025) 740.**
https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52025DC0740

03 **Deryck, Leblanc-Albarel & Preneel — White-Box Attacks on PhotoDNA (2026).**
https://eprint.iacr.org/2026/486

SLIDE 23 Pourquoi des empreintes perceptuelles ?

01 **Microsoft — PhotoDNA : présentation et affirmations publiques.**
https://www.microsoft.com/en-us/photodna

02 **NCMEC — CyberTipline 2025 : plus de 12,1 millions d'empreintes partagées.**
https://www.missingkids.org/gethelpnow/cybertipline/cybertiplinedata

03 **Thorn — Safer Solutions : base de plus de 172,5 millions d'empreintes vérifiées.**
https://safer.io/solutions/

04 **Deryck, Leblanc-Albarel & Preneel — White-Box Attacks on PhotoDNA (2026).**
https://eprint.iacr.org/2026/486

05 **18 U.S.C. §2258A — signalement par les fournisseurs soumis à la loi américaine.**
https://uscode.house.gov/view.xhtml?edition=prelim&num=0&req=granuleid%3AUSC-prelim-title18-section2258A

Références (15/18)

SLIDE 24 **Le hachage perceptuel dans le circuit de signalement**

01 **Commission européenne — rapport d'évaluation COM(2025) 740.**
https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52025DC0740

02 **Google — rapport UE 2023.**
https://storage.googleapis.com/transparencyreport/report-downloads/pdf-report-23_2023-1-1_2023-12-31_en_v1.pdf

03 **Microsoft — rapport UE 2022.**
https://cdn-dynmedia-1.microsoft.com/is/content/microsoftcorp/microsoft/final/en-us/microsoft-brand/documents/Microsoft-Transparency-Report%282022%29-EU-CSAM-Interim-Regulatory-Report.pdf

04 **NCMEC — CyberTipline : fonctionnement et routage des rapports.**
https://www.missingkids.org/gethelpnow/cybertipline

SLIDE 25 **Le point de départ détermine ce que l'attaque permet**

01 **Deryck, Leblanc-Albarel & Preneel — White-Box Attacks on PhotoDNA (2026).**
https://eprint.iacr.org/2026/486

02 **PseudoDNA — démonstrations et implications pratiques des attaques.**
https://www.pseudodna.eu/

Références (16/18)

SLIDE 26 **Quatre attaques contre PhotoDNA**

01 **Deryck, Leblanc-Albarel & Preneel — White-Box Attacks on PhotoDNA (2026).**
https://eprint.iacr.org/2026/486

02 **PseudoDNA — démonstrations et implications pratiques des attaques.**
https://www.pseudodna.eu/

SLIDE 27 **1. Inversion du hash : l'empreinte révèle des informations**

01 **Deryck, Leblanc-Albarel & Preneel — White-Box Attacks on PhotoDNA (2026).**
https://eprint.iacr.org/2026/486

02 **Microsoft — PhotoDNA : présentation et affirmations publiques.**
https://www.microsoft.com/en-us/photodna

SLIDE 28 **2. Collision : empoisonnement de base de données**

01 **Deryck, Leblanc-Albarel & Preneel — White-Box Attacks on PhotoDNA (2026).**
https://eprint.iacr.org/2026/486

SLIDE 29 **3. Seconde préimage : un faux positif ciblé**

01 **Deryck, Leblanc-Albarel & Preneel — White-Box Attacks on PhotoDNA (2026).**
https://eprint.iacr.org/2026/486

Références (17/18)

SLIDE 29 **3. Seconde préimage : un faux positif ciblé**

02 **PseudoDNA — démonstrations et implications pratiques des attaques.**
https://www.pseudodna.eu/

SLIDE 30 **4. Évasion : faire passer un contenu sous le seuil**

01 **Deryck, Leblanc-Albarel & Preneel — White-Box Attacks on PhotoDNA (2026).**
https://eprint.iacr.org/2026/486

02 **PseudoDNA — démonstrations et implications pratiques des attaques.**
https://www.pseudodna.eu/

SLIDE 31 **Ce que ces attaques remettent en cause**

01 **Deryck, Leblanc-Albarel & Preneel — White-Box Attacks on PhotoDNA (2026).**
https://eprint.iacr.org/2026/486

02 **PseudoDNA — démonstrations et implications pratiques des attaques.**
https://www.pseudodna.eu/

03 **Confirmation de Microsoft sur la fonction déployée — échanges de divulgation coordonnée.**
Document non public — archives des auteurs.

Références (18/18)

SLIDE 32 **La détection des CSAM repose sur une infrastructure mondiale**

01 **Commission européenne — cadre juridique.**
https://home-affairs.ec.europa.eu/policies/internal-security/protecting-children-sexual-abuse/legal-framework-protect-children_en

02 **Commission européenne — rapport COM(2023) 797 sur le règlement intérimaire.**
https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52023DC0797

03 **NCMEC — 2025 CyberTipline Report et archives annuelles.**
https://www.missingkids.org/gethelpnow/cybertipline/cybertiplinedata

04 **PseudoDNA — démonstrations et implications pratiques des attaques.**
https://www.pseudodna.eu/

05 **Deryck, Leblanc-Albarel & Preneel — White-Box Attacks on PhotoDNA (2026).**
https://eprint.iacr.org/2026/486

06 **18 U.S.C. §2258A — contenu des CyberTips et obligations des fournisseurs.**
https://uscode.house.gov/view.xhtml?edition=prelim&num=0&req=granuleid%3AUSC-prelim-title18-section2258A

07 **NCMEC — France: Data Insights.**
https://www.globalchildexploitationpolicy.org/data-insights/france

08 **Échanges de divulgation coordonnée et réunions avec Microsoft — archives des auteurs.**
Document non public — archives des auteurs.