FROM DETECTION TO REPORTING

Online CSAM detection: Procedures, mechanisms and limitations

Diane Leblanc-Albarel

KU Leuven · COSIC

AFSIN · 15–17 September 2026

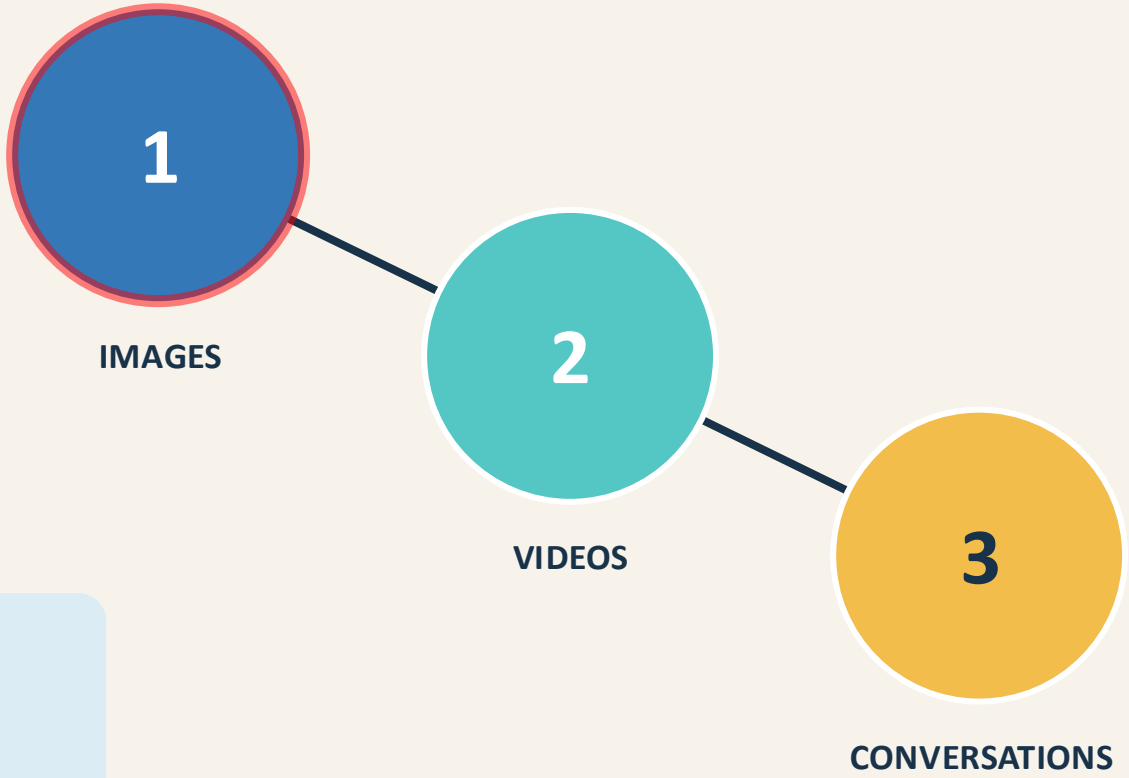
What are we talking about?

Content documenting or depicting the sexual exploitation of minors.

CSAM

Child Sexual Abuse Material

Child sexual abuse material



Two different objectives

Find previously identified content
Identify potentially new content

NCMEC: a US centre, reports from around the world

US platforms must report known cases to it, regardless of the country involved.

1984
FOUNDED
Private non-profit organisation

US LAW
LEGAL ROLE
Mandatory reporting by US platforms

NCMEC
National Center for Missing & Exploited Children

1998
REPORTING SYSTEM
Launch of the national reporting system

REGARDLESS OF THE COUNTRY
Reports to NCMEC, including from abroad

DUAL ROLE

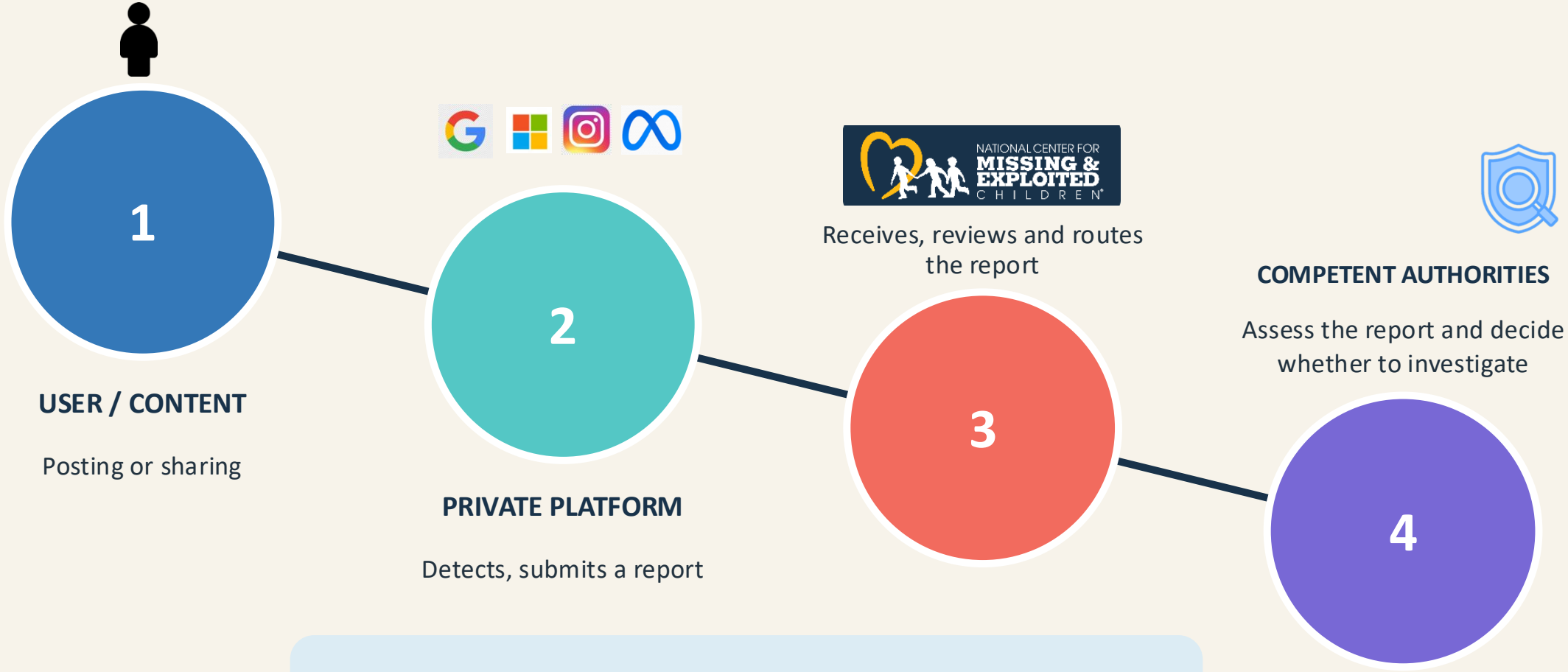
OPERATIONAL
receives, analyses and routes reports

INFLUENCE
engages with Congress and advocates maintaining detection capabilities

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

From the user to a possible investigation

A detection passes through several actors before any judicial assessment.



Detection can lead to a report. That report is not a judicial finding.

Where can content become visible?

End-to-end encryption does not remove all observation points, but shifts them.

1 SERVER-SIDE

E.g. Gmail: automated detection is possible on an attachment accessible to the service.

2 AFTER A REPORT

WhatsApp receives up to 5 recent messages sent by the reported account.

3 ON THE DEVICE

Detection before encryption or after decryption: potential client-side scanning.

Automatic scanning depends on the service

PRIVATE MESSAGING

Facebook

Posted or live photos/videos: automatic scanning

Messenger: automatic media scanning outside E2EE

Instagram

Photos, Reels, live videos: automatic scanning

Private messages: automatic scanning

WhatsApp

Profile/group photos: automatic scanning

E2EE messages: scanning after a report

TikTok

Public/private/live videos: automatic scanning

Private messages: automatic text + content scanning

Google

Drive, Photos, YouTube: automatic media scanning

Gmail (default setting): automatic image scanning

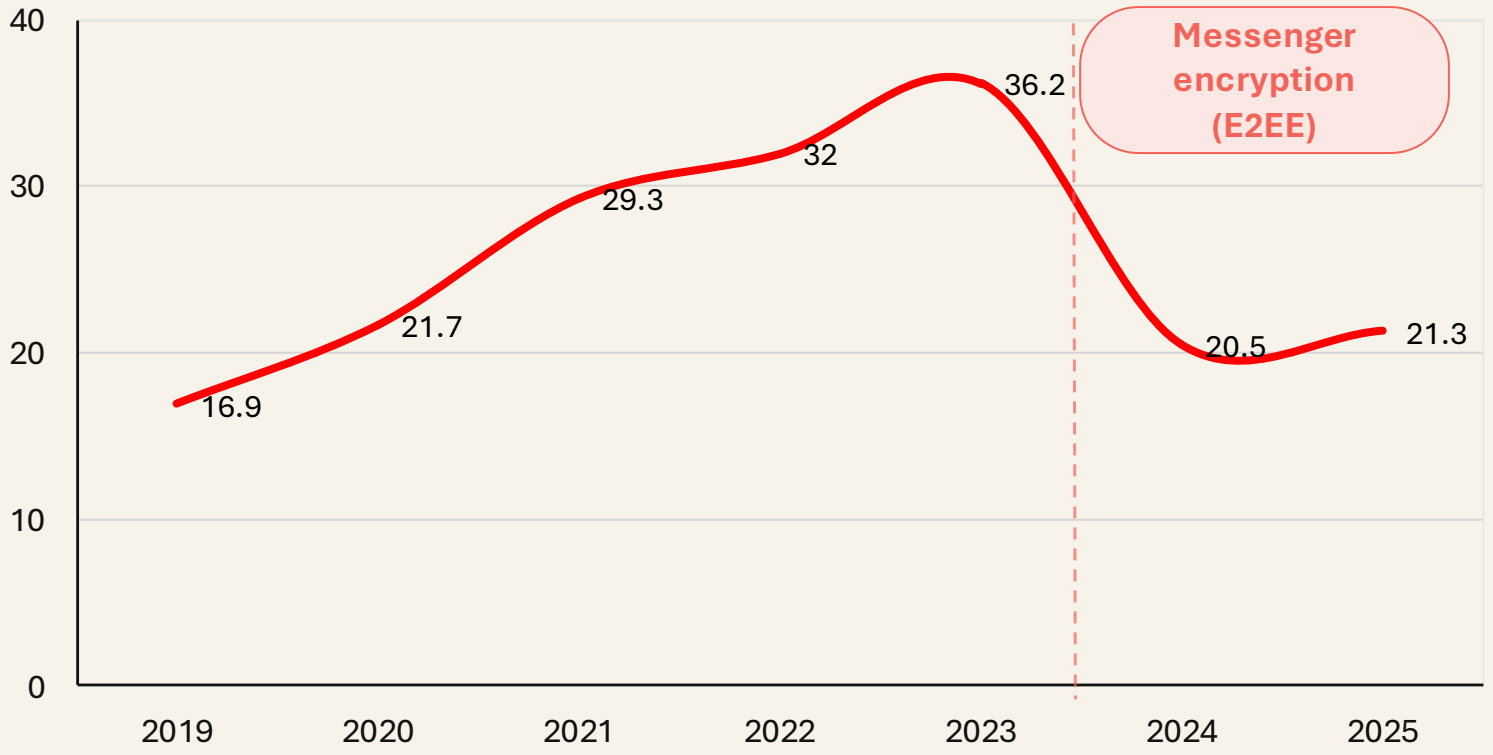
Microsoft

OneDrive, Outlook, Bing, Xbox: automatic scanning

Outlook: attachments; Teams

The scale of reporting has changed

Annual reports received by NCMEC, in millions.



NCMEC links the decline to Messenger E2EE: Facebook reports, 17.8 M → 8.6 M → 4.9 M.

< 0.5 M **in 2012**

32 M in 2022: a more than 60-fold increase in ten years.

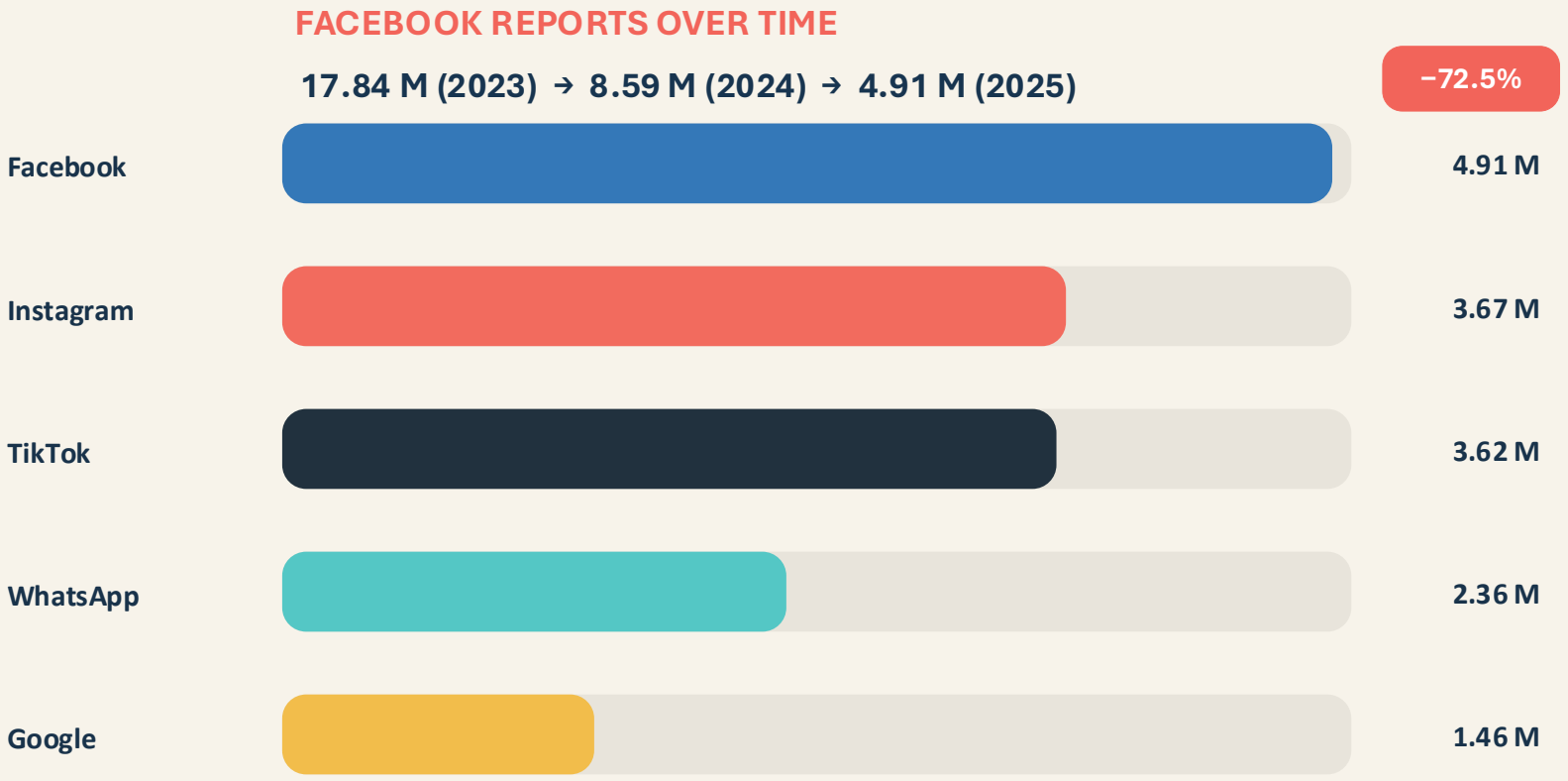
NATIONAL CENTER FOR MISSING & EXPLOITED CHILDREN®

Google Meta Microsoft SNAPCHAT

Most reports come from online platforms.

Volumes are concentrated among a few platforms

Reports sent to NCMEC by the main providers in 2025.



75.6%

of ESP reports come from these five providers

(>2000 ESP in total)

21.18 M reports → 61.8 M files

29.4 M images

26.3 M videos

+ 6.1 M other files

A report ≠ a file ≠ an account ≠ a user ≠ an investigation ≠ a conviction.

A report is much more than an image

The report may bring together several types of data.

The report may contain sensitive personal data.

A French report also remains in the US system

Reports pass through NCMEC and U.S. agencies

Detected in Europe → reported in the US → made available to the competent authorities

The report’s lifetime remains poorly documented

The published rule for the platform does not set NCMEC’s retention schedule.

PLATFORM

AT LEAST 1 YEAR

Some items may be retained longer depending on the purposes and applicable rules.

NCMEC

PUBLIC MAXIMUM RETENTION PERIOD

NOT FOUND

No general public rule was found requiring deletion of reports ultimately left unconfirmed.

This does not establish permanent retention; it reveals a lack of public transparency about the report’s lifecycle.

Largely automated processing

Human review may often follow automated processing.

NCMEC · REPORT PROCESSING

~ 153 000 / DAY

IMAGES AND VIDEOS

AUTOMATED DETECTION

RECOGNISED FILE

AUTOMATED PROCESSING

UNRECOGNISED FILE

HUMAN REVIEW

IF NEW DATABASE ENTRY:

3 HUMAN VALIDATIONS

GOOGLE + YOUTUBE · 2025 – 8.8 M – 24 200 / DAY

99.67 %

AUTOMATIC

0.33 %

MANUAL

1 ACTION / CONTENT REMOVAL

2 REPORTING TO NCMEC

IF APPEALED

1. Automated check
2. If appeal denied, human review and final decision

Automation on both sides. Action may precede reporting.

In Europe, an exception permits message scanning

The derogation temporarily permits certain forms of voluntary detection.

RULE

ePRIVACY

Confidentiality of private communications

TEMPORARY EXCEPTION

REGULATION (EU) 2021/1232

VOLUNTARY DETECTION

Detect · remove · report
Reinstated until 3 April 2028

It allows certain processing. It does not require general scanning.

The permanent framework remains to be negotiated.

The CSAR seeks a permanent framework

The three EU institutions agree on a European Centre, but not yet on the means.

Service obligations

Assess and reduce risks
linked to their services

Report identified abuse
to the European Centre

Remove content
on orders from the authorities

CSAR

PERMANENT FRAMEWORK

European Centre's tasks

Check reports
and discard manifest errors

Forward reports
to national authorities and Europol

Help victims
get content removed

ONE POINT OF AGREEMENT · TWO QUESTIONS

EUROPEAN CENTRE

Reduce dependence on NCMEC

QUESTION 1

MANDATORY DETECTION?

QUESTION 2

AGE ASSURANCE?

France: supporting the objective, regulating the means

Two parliamentary resolutions, then an official position in the Council of the EU.

SENATE · RESOLUTION No. 77

OPINION ON THE CSAR PROPOSAL

ADOPTED ON 20 MARCH 2023

"Calls for the deletion of provisions relating to indiscriminate searches […]"

Regulate detection orders

- previously identified content
- independent audit
- protected encryption
- privacy-preserving age assurance

ASSEMBLY · RESOLUTION T.A. No. 186

OWN-INITIATIVE OPINION

ADOPTED ON 3 DECEMBER 2025

"Europe can no longer outsource its children's safety to a foreign organisation"

Reduce dependence on NCMEC

- create a European Centre
- sufficient detection capabilities
- protect privacy and data

FRANCE IN THE COUNCIL OF THE EU

TEMPORARY DEROGATION

APPROVED ON 23 JULY 2026

"Strong concern about […] the exclusion of encrypted communications"

WITHIN THE SCOPE OF VOLUNTARY DETECTION

APPROVAL

Voluntary regime until April 2028

CONCERNS

- Seeks to avoid excluding all detection in E2EE
- Regrets that investigators' difficulties accessing data remain insufficiently addressed

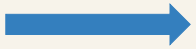
PhotoDNA: from Microsoft to global deployment

To explore the technology, let us start with PhotoDNA: a function for already known images.

PhotoDNA

Microsoft
+ Hany Farid
2009

closed function



PROVIDED TO NCMEC

CLOUD SERVICE · 2014

then deployed to qualified platforms and organisations

Find copies of
already known images

MICROSOFT ALONE

12.3 billion

items scanned in 2022 by the covered services

Full specification and code not public · no historically reproducible independent benchmark found

"PhotoDNA is generally presented as extremely reliable. Yet its public claims are not always mutually consistent."

An unverified claim becomes a reference

"1 in 50 billion": a statement to the US Congress becomes information presented as verified.

The Irish case shows the consequences after detection

A dismissed report may nevertheless leave data in the system.

ONLY

20 %

considered to constitute CSAM

> 11% false reports

≈ 69%: neither described as confirmed CSAM
nor counted as false reports

> 11% OF REPORTS

DISMISSED REPORTS

family or beach photos cited

AFTER PEOPLE WERE CLEARED

Data still retained

according to the Irish Council for Civil Liberties (ICCL)

Potential harm may arise after detection: transmission, classification and retention.

We do not know which detection mechanism triggered each of these reports.

Meta and Google: errors observed in practice

Two indicators: decisions reversed on appeal and unconfirmed detections.

META · 2023 DATA

≈ 3.6 M
items of content actioned

APPEALS
review

≈ 0.32%
content restored or account measures reversed
≈ 1 in 310

GOOGLE · Gmail · 2024

1 834
content items detected as CSAM
reported denominator

REVIEW
human

0.54 %
10 items not confirmed as CSAM
≈ 1 in 183

REPORTS CONCERNING FRANCE · 2024

165 818
Made available by NCMEC
France + listed overseas territories

158 503
Received by France from NCMEC
According to the EU report

NCMEC triages and routes reports.

Why an independent evaluation?

In security, we test a guarantee by actively trying to break the system.

PUBLIC CLAIMS

"1 in 50 billion"

PUBLIC DOCUMENTS

non-reproducible evidence
+ contradictory information

CRYPTANALYSIS

attack the
deployed system

SUCCESSFUL ATTACK

demonstrated limitation

A failed attack ≠ proof of security

White-Box Attacks on PhotoDNA
Accepted at S&P 2027 (conference)
Deryck · Leblanc-Albarel · Preneel
Full paper: *ia.cr/2026/486*

pseudodna.eu

RESPONSIBLE DISCLOSURE

> 6 MONTHS OF EXCHANGES AND MEETINGS
with Microsoft, including PhotoDNA's creator at Microsoft
Attacks on a standard personal computer
Applicability to the deployed system confirmed · countermeasures discussed

From cryptographic hashing to perceptual hashing

Two nearly identical photos: SHA-256 changes completely, the perceptual hash remains very close.

Nearly identical photos

≈

SHA-256

8A1F…D04C

≠

COMPLETELY DIFFERENT

SHA-256

F93B…117E

Perceptual hash

8B31…EA76

≈

VERY CLOSE

Perceptual hash

8B32…EA76

Similarity becomes a distance

The closer the fingerprints, the smaller the calculated distance.

1 NORMALISE THE IMAGE → 2 EXTRACT FEATURES → 3 CREATE THE FINGERPRINT → 4 CALCULATE THE DISTANCE → 5 COMPARE TO THE THRESHOLD

CUTOFF

small distance — large distance

original | slightly modified | modified | different image | very different image

DETECTED | NOT DETECTED

Why perceptual fingerprints?

Recognising variants and limiting information leakage

IMAGES OR CRYPTOGRAPHIC HASHES

≠

Different images, different hashes

1. Exact comparison only

No recognition of similar images.
ALL variants must be stored,
as images or cryptographic hashes.

2. Images: information leakage

A leak directly exposes
the content and the victims

PERCEPTUAL FINGERPRINTS

NCMEC list: 12.1 M

Safer / Thorn: ≈ 175 M

Every report is sent to NCMEC

Recognition by similarity

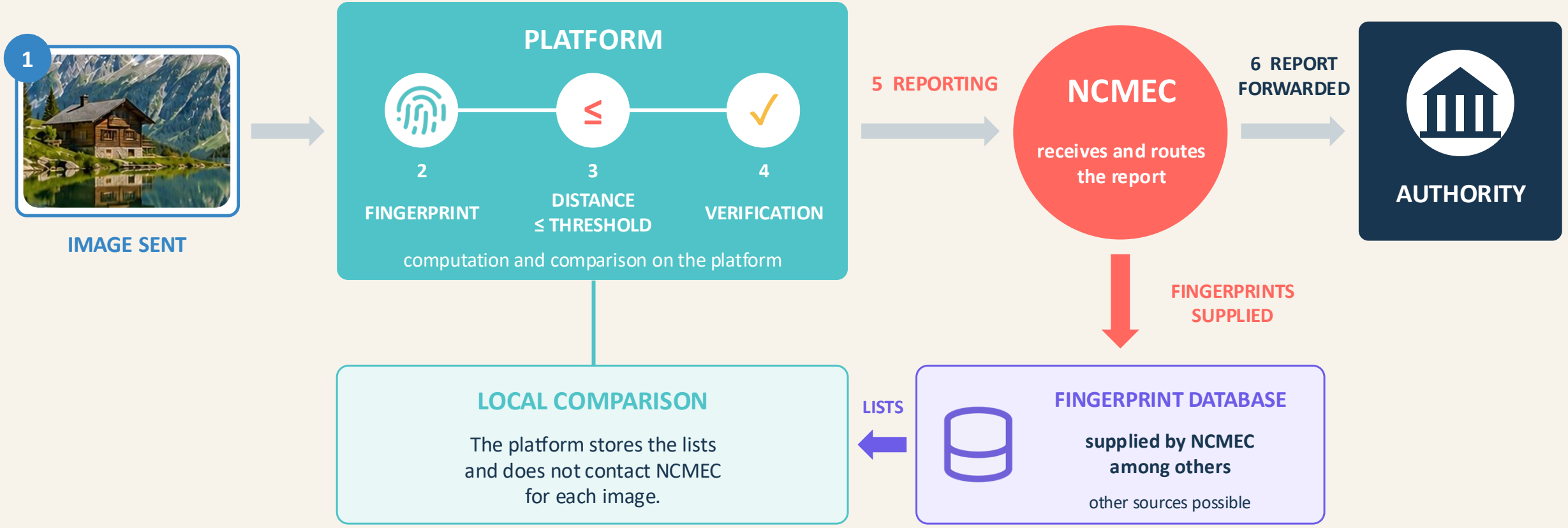
Similar images produce
similar fingerprints

Image protection

The list does not contain
the original images

Perceptual hashing in the reporting process

The platform computes and compares locally; NCMEC receives a report only after a positive detection.



Four attacks on PhotoDNA

Close fingerprints: similar colours. Distant fingerprints: different colours.

Information leakage

INVERSION

Reconstructed image

Identical

Attack on PhotoDNA

Known fingerprint

Database poisoning

COLLISION

Two starting images

≠

Attack on PhotoDNA

Modified

≠

Modified

≈

Very close or equal fingerprints

False accusation

SECOND PREIMAGE

Target image

Attack on PhotoDNA

Unchanged

≠

Constructed

≈

Fixed target

Very close or equal

Detection avoided

EVASION

Detected image

Attack on PhotoDNA

Original

≈

Modified

≠

Distant fingerprints

Illustrative diagrams using benign images

1. The fingerprint reveals information

Operational scenario: a third party obtains a sensitive fingerprint and extracts coarse visual information.

ACCESSIBLE HASH

copied database, leak or storage on the device

→

ATTACK COMPUTATION

search for an image that reproduces the fingerprint

→

INFORMATION REVEALED

category, silhouette or coarse structure

99.5%

success rate

11.2 s

on average

Possible consequence: Recovering information about sensitive images.

2. Collision: database poisoning

A fingerprint added to the database can trigger detection of another image.

1 · COLLISION

The attacker creates X’ and Y’ with the same fingerprint

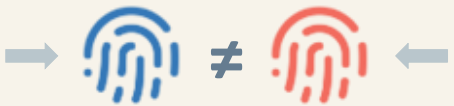
2 · DATABASE

They add X”s fingerprint to the reference database

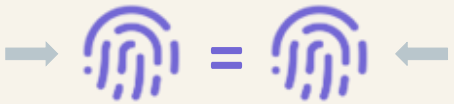
3 · FALSE POSITIVE

Y’ can then be wrongly detected

Original images



Modified images



100 %

successful exact collisions

215 s

average time

3. Second preimage: False accusation

This time, the attacker starts with a known illicit image or fingerprint and targets a specific victim.

1 · TARGET FINGERPRINT

fingerprint of known illicit content

2 · LAWFUL IMAGE

modified to approach the target fingerprint

3 · VICTIM

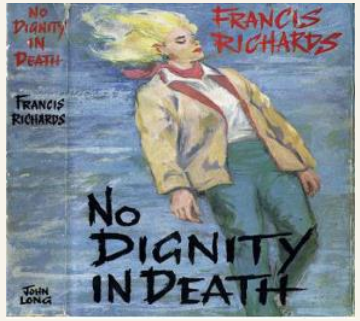
the image reaches their account or device

4 · MATCH

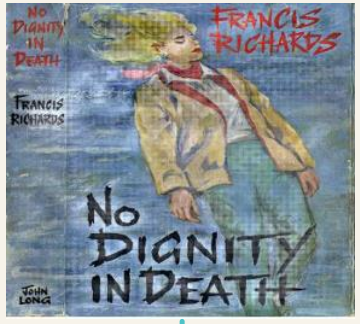
possible removal, suspension or reporting

TARGET FINGERPRINT

LAWFUL IMAGE



AFTER ATTACK



appearance to preserve

≈

Match

99.7%

Success rate

≈ 118 s

Runtime

small distance — large distance

threshold τ

$d \leq \tau \rightarrow$ MATCH

Possible consequence: a visually lawful image may be treated as the targeted illicit content.

4. Evasion: spreading illegal content

A criminal slightly modifies the image before sharing it; the platform no longer finds the known fingerprint.

KNOWN CONTENT

its fingerprint is already in the blocklist

→

MINIMAL MODIFICATION

before sharing

→

NO MATCH

distance above the threshold → no report

KNOWN IMAGE

Known illegal image

≠

AFTER ATTACK

similar appearance

100 %

Success rate

9.94 s

Average time

d > τ → NO MATCH

Possible consequence: the content circulates undetected.

What these attacks call into question

Practical consequences for CSAM detection

RESULTS ON PHOTODNA	POSSIBLE OPERATIONAL CONSEQUENCES
Inversion Recover visual information	**Fingerprints leaked** in a **cyberattack** Extracting information about sensitive images
Collision and second preimage Trigger false positives	Influx of **reports of innocent images** **Increased workload** for review teams
Evasion Avoid detection of illegal images	Circulation of **illegal content** No **reporting** or **prosecution**
Confirmation from Microsoft **Studied function identical to the one deployed in production**	

What verifiable guarantees at scale?

CSAM detection relies on a global infrastructure

It combines a very large scale, a central role for NCMEC and poorly documented technical mechanisms.

AT A GLOBAL SCALE

- detection and reporting volumes have increased sharply
- hundreds of billions of images are scanned every year
- major platforms feed a global process centred on NCMEC

DEPENDENCE ON NCMEC

- platforms send their reports to NCMEC
- reports may contain sensitive personal data
- several techniques remain proprietary, without reproducible public evaluation
- Europe seeks a permanent framework

PHOTODNA: OUR RESULTS

- **ATTACKS IN SECONDS**
- **CONFIRMED PRACTICAL IMPACT**

OUR CONCLUSION

- Regardless of the debate: deployed mechanisms are not reliable
- From security through obscurity to mechanisms public authorities can verify and audit

Open question:
Can we minimise false accusations while maintaining sufficient effectiveness?

Contact / Webpage

diane.leblanc-albarel@esat.kuleuven.be

https://dianeleblancalbarel.github.io/

References (1/18)

SLIDE 2 What are we talking about?

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

02 **Regulation (EU) 2021/1232 — temporary derogation from the ePrivacy Directive.**
https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32021R1232

SLIDE 3 NCMEC: a US centre, reports from around the world

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

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

03 **18 U.S.C. §2258A — CyberTip contents and providers' obligations.**
https://www.govinfo.gov/link/uscode/18/2258A

04 **NCMEC — End-to-End Encryption: public position on encryption.**
https://www.missingkids.org/theissues/end-to-end-encryption

SLIDE 4 From the user to a possible investigation

01 **18 U.S.C. §2258A — CyberTip contents and providers' obligations.**
https://www.govinfo.gov/link/uscode/18/2258A

References (2/18)

SLIDE 4 **From the user to a possible investigation**

02 **NCMEC — CyberTipline: how reports are processed and routed.**
https://www.missingkids.org/gethelpnow/cybertipline

SLIDE 5 **Where can content become visible?**

01 **Google — EU report 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 — Reporting a contact: up to five recent messages.**
https://faq.whatsapp.com/414631957536067

SLIDE 6 **Automatic scanning depends on the service**

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

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

References (3/18)

SLIDE 6 **Automatic scanning depends on the service**

03 **eSafety Commissioner — Basic Online Safety Expectations full transparency report, October 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, January–June 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 — EU report 2023.**
https://storage.googleapis.com/transparencyreport/report-downloads/pdf-report-23_2023-1-1_2023-12-31_en_v1.pdf

09 **Microsoft — EU report 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

References (4/18)

SLIDE 7 The scale of reporting has changed

01 **NCMEC — Transparency Report CY 2025, Appendix A: annual series and counting methodology.**
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 and annual archives.**
https://www.missingkids.org/gethelpnow/cybertipline/cybertiplinedata

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

SLIDE 8 Volumes are concentrated among a few platforms

01 **NCMEC — 2025 CyberTipline Report and annual archives.**
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

References (5/18)

SLIDE 8 **Volumes are concentrated among a few platforms**

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 **A report is much more than an image**

01 **18 U.S.C. §2258A — CyberTip contents and providers' obligations.**
https://www.govinfo.gov/link/uscode/18/2258A

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

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

04 **Microsoft — EU report 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

References (6/18)

SLIDE 10 **A French report also remains in the US system**

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

02 **Meta — working with law enforcement.**
https://transparency.meta.com/en-gb/policies/improving/working-with-law-enforcement/

03 **Microsoft — EU report 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 — EU report 2024.**
https://values.snap.com/privacy/transparency/european-union-h2-2024?lang=en-GB

05 **LinkedIn — EU report.**
https://www.linkedin.com/help/linkedin/answer/a1347128

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

07 **NCMEC — CyberTipline: how reports are processed and routed.**
https://www.missingkids.org/gethelpnow/cybertipline

08 **18 U.S.C. §2258A — CyberTip contents and providers' obligations.**
https://www.govinfo.gov/link/uscode/18/2258A

References (7/18)

SLIDE 11 The report's lifetime remains poorly documented

01 **18 U.S.C. §2258A — CyberTip contents and providers' obligations.**
https://www.govinfo.gov/link/uscode/18/2258A

02 **NCMEC — Privacy Policy: security, retention and transfers.**
https://www.missingkids.org/footer/privacypolicy

03 **NCMEC — REPORT Act and preservation by platforms.**
https://www.missingkids.org/blog/2024/first-line-of-defense-guidelines-to-help-online-platforms-detect-sexually-exploited-kids

SLIDE 12 Largely automated processing

01 **Google — CSAM Transparency Report: automatic and manual detection, first half of 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: automatic and manual detection, second half of 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

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

References (8/18)

SLIDE 12 **Largely automated processing**

04 **Google — Content handling, enforcement and appeal reviews.**
https://support.google.com/product-documentation/answer/15464420?hl=en

05 **NCMEC — CY 2025 Transparency Report: images, videos and report processing.**
https://www.missingkids.org/content/dam/missingkids/pdfs/OJJDP-NCMEC-Transparency-CY-2025.pdf

06 **NCMEC — CyberTipline 2025: file triage and fingerprint sharing.**
https://www.missingkids.org/gethelpnow/cybertipline/cybertiplinedata

SLIDE 13 **In Europe, an exception permits message scanning**

01 **Directive 2002/58/EC (ePrivacy), Article 5: confidentiality of communications.**
https://eur-lex.europa.eu/legal-content/FR/TXT/?uri=CELEX:32002L0058

02 **Regulation (EU) 2021/1232 — temporary derogation from the ePrivacy Directive.**
https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32021R1232

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

04 **Council of the EU — 23 July 2026.**
https://www.consilium.europa.eu/en/press/press-releases/2026/07/23/fighting-child-sexual-abuse-online-interim-measure-protecting-children-now-reinstated/

References (9/18)

SLIDE 14 The CSAR seeks a permanent framework

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

02 **Council of the EU — position of 26 November 2025 on the CSAR proposal.**
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 **Council of the EU — mandate ST 15318/2025.**
https://data.consilium.europa.eu/doc/document/ST-15318-2025-INIT/en/pdf

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

05 **European Parliament — November 2023 mandate.**
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 **European Parliament — Legislative Train: status of the CSAR file.**
https://www.europarl.europa.eu/legislative-train/spotlight-JD22/file-combating-child-sexual-abuse-online

References (10/18)

SLIDE 15 **France: supporting the objective, regulating the means**

01 **Senate — European resolution No. 77, 20 March 2023.**
https://www.senat.fr/leg/tas22-077.html

02 **National Assembly — European resolution T.A. No. 186, 3 December 2025.**
https://www.assemblee-nationale.fr/dyn/17/textes/l17t0186_texte-adopte-seance

03 **Council of the EU — vote and French statement, CM 3756/1/26 REV 1, 23 July 2026.**
https://data.consilium.europa.eu/doc/document/CM-3756-2026-REV-1/fr/pdf

04 **Senate — European Affairs Committee, 15 February 2023.**
https://www.senat.fr/compte-rendu-commissions/20230213/europ.html

05 **National Assembly — European Affairs Committee, 15 October 2025.**
https://www.assemblee-nationale.fr/dyn/17/comptes-rendus/due/l17due25260042_compte-rendu.pdf

06 **Senate — Rules of Procedure, Article 73 quinquies C: European resolution procedure.**
https://www.senat.fr/reglement/reglement52.html

07 **National Assembly — Rules of Procedure, Articles 151-5 to 151-7.**
https://www.assemblee-nationale.fr/dyn/17/divers/texte_reference/02_reglement_assemblee_nationale.pdf

SLIDE 16 **PhotoDNA: from Microsoft to global deployment**

01 **Microsoft — PhotoDNA: overview and public claims.**
https://www.microsoft.com/en-us/photodna

References (11/18)

SLIDE 16 PhotoDNA: from Microsoft to global deployment

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

03 **Microsoft — EU report 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 An unverified claim becomes a reference

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

04 **European Commission — report COM(2023) 797 on the interim regulation.**
https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52023DC0797

05 **European Commission — evaluation report COM(2025) 740.**
https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52025DC0740

References (12/18)

SLIDE 18 The Irish case shows the consequences after detection

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 and Google: errors observed in practice

01 **European Commission — COM(2025) 740, Tables 5 (Meta / Google) and 6 (France).**
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/

03 **NCMEC — 2024 CyberTipline Reports by Country: France and overseas territories.**
https://www.missingkids.org/content/dam/missingkids/pdfs/2024-reports-by-country.pdf

References (13/18)

SLIDE 20 **Why an independent evaluation?**

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

02 **PseudoDNA — demonstrations and practical implications of the attacks.**
https://www.pseudodna.eu/

03 **Coordinated disclosure exchanges and meetings with Microsoft — authors' archives.**
Non-public document — authors' records.

SLIDE 21 **From cryptographic hashing to perceptual hashing**

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

02 **European Commission — evaluation report COM(2025) 740.**
https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52025DC0740

03 **Microsoft — PhotoDNA: overview and public claims.**
https://www.microsoft.com/en-us/photodna

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

References (14/18)

SLIDE 22 Similarity becomes a distance

01 **Microsoft — PhotoDNA: overview and public claims.**
https://www.microsoft.com/en-us/photodna

02 **European Commission — evaluation report 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 Why perceptual fingerprints?

01 **Microsoft — PhotoDNA: overview and public claims.**
https://www.microsoft.com/en-us/photodna

02 **NCMEC — CyberTipline 2025: over 12.1 million fingerprints shared.**
https://www.missingkids.org/gethelpnow/cybertipline/cybertiplinedata

03 **Thorn — Safer Solutions: database of over 172.5 million verified fingerprints.**
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 — reporting by providers subject to US law.**
https://www.govinfo.gov/link/uscode/18/2258A

References (15/18)

SLIDE 24 **Perceptual hashing in the reporting process**

01 **European Commission — evaluation report COM(2025) 740.**
https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52025DC0740

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

03 **Microsoft — EU report 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: how reports are processed and routed.**
https://www.missingkids.org/gethelpnow/cybertipline

SLIDE 25 **Four attacks on PhotoDNA**

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

02 **PseudoDNA — demonstrations and practical implications of the attacks.**
https://www.pseudodna.eu/

References (16/18)

SLIDE 26 **1. The fingerprint reveals information**

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

02 **Microsoft — PhotoDNA: overview and public claims.**
https://www.microsoft.com/en-us/photodna

SLIDE 27 **2. Collision: database poisoning**

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

SLIDE 28 **3. Second preimage: False accusation**

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

02 **PseudoDNA — demonstrations and practical implications of the attacks.**
https://www.pseudodna.eu/

SLIDE 29 **4. Evasion: spreading illegal content**

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

References (17/18)

SLIDE 29 **4. Evasion: spreading illegal content**

02 **PseudoDNA — demonstrations and practical implications of the attacks.**
https://www.pseudodna.eu/

SLIDE 30 **What these attacks call into question**

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

02 **PseudoDNA — demonstrations and practical implications of the attacks.**
https://www.pseudodna.eu/

03 **Microsoft's confirmation regarding the deployed function — coordinated disclosure exchanges.**
Non-public document — authors' records.

SLIDE 31 **CSAM detection relies on a global infrastructure**

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

02 **European Commission — report COM(2023) 797 on the interim regulation.**
https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52023DC0797

References (18/18)

SLIDE 31 **CSAM detection relies on a global infrastructure**

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

04 **PseudoDNA — demonstrations and practical implications of the attacks.**
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 — CyberTip contents and providers' obligations.**
https://www.govinfo.gov/link/uscode/18/2258A

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

08 **Coordinated disclosure exchanges and meetings with Microsoft — authors' archives.**
Non-public document — authors' records.