Private Devices No Longer Private

The Broader Implications of Apple’s Content Scanning Push

Nadim Kobeissi, Swiss Cyber Storm 2021 — Bern, Switzerland

CYBER ATTACKS AND DEFENSE • SWISS CYBER STORM • WWW.SWISSCYBERSTORM.COM

Talk Contents

Private Devices No Longer Private

1. Apple’s Proposal
2. Arguments Against
3. Arguments For
4. Conclusion

1. Apple’s Proposal

NCMEC and CSAM

Two important acronyms

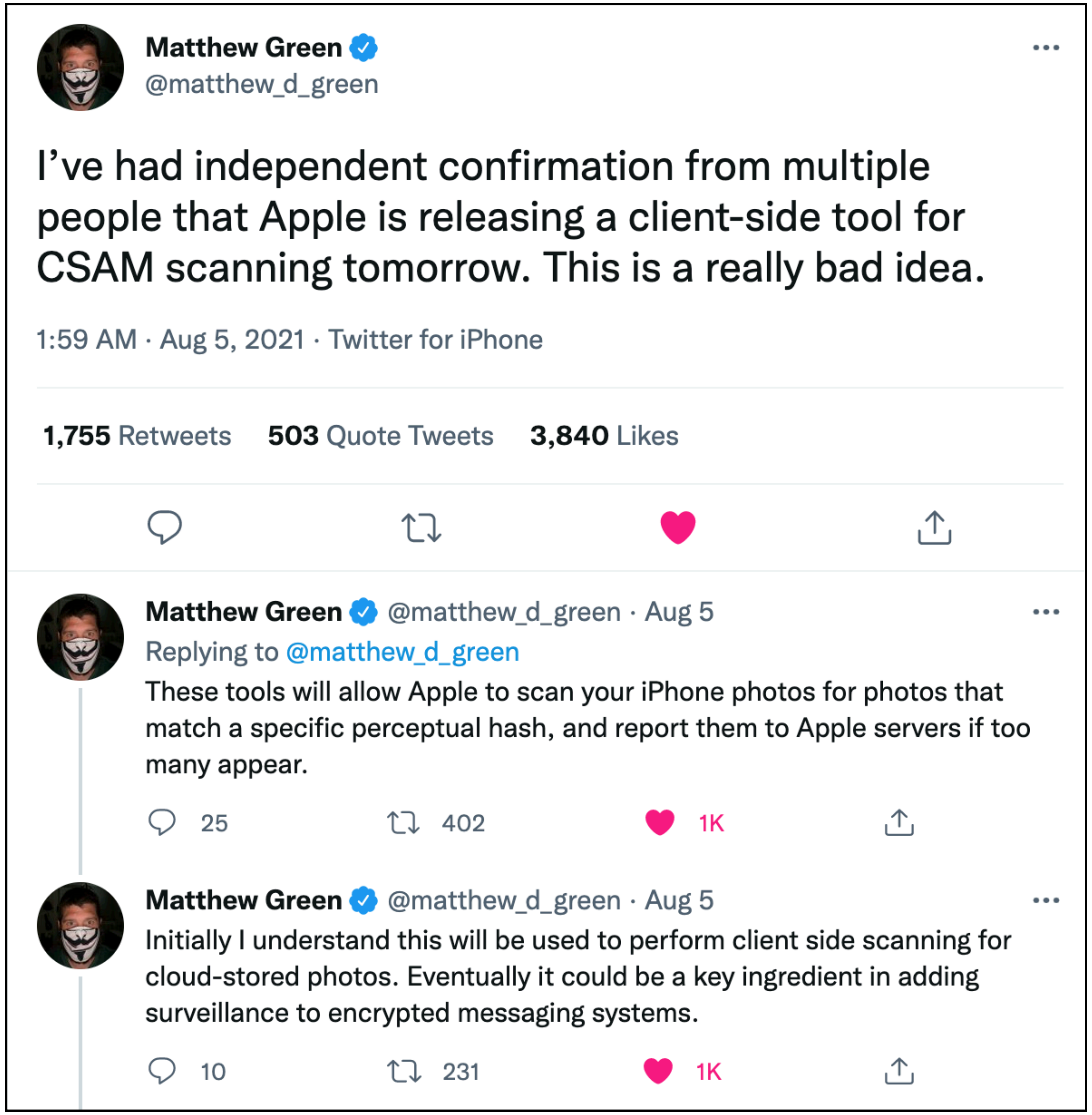
- **NCMEC:** National Center for Missing & Exploited Children.

- **CSAM:** Child Sexual Abuse Materials.

NATIONAL CENTER FOR
MISSING &
EXPLOITED
CHILDREN®

Initial Rumblings

August 4th, 2021



Matthew Green @matthew_d_green

I’ve had independent confirmation from multiple people that Apple is releasing a client-side tool for CSAM scanning tomorrow. This is a really bad idea.

1:59 AM · Aug 5, 2021 · Twitter for iPhone

1,755 Retweets 503 Quote Tweets 3,840 Likes

Matthew Green @matthew_d_green · Aug 5

Replying to @matthew_d_green

These tools will allow Apple to scan your iPhone photos for photos that match a specific perceptual hash, and report them to Apple servers if too many appear.

25 402 1K

Matthew Green @matthew_d_green · Aug 5

Initially I understand this will be used to perform client side scanning for cloud-stored photos. Eventually it could be a key ingredient in adding surveillance to encrypted messaging systems.

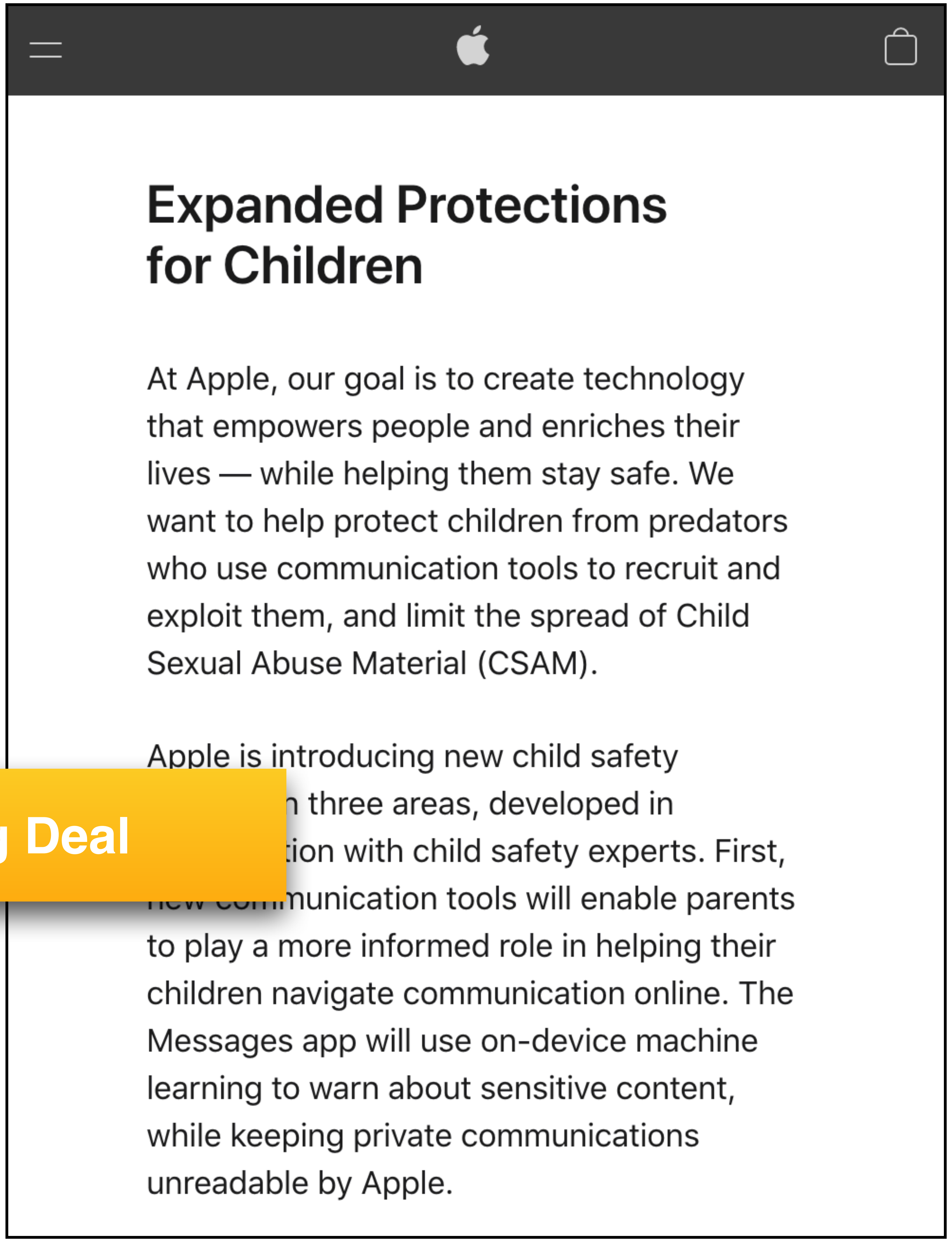
10 231 1K

Apple’s Announcement

August 5th, 2021

1. *“The Messages app will add new tools to warn children and their parents when receiving or sending sexually explicit photos.”*
2. *“New technology in iOS and iPadOS* ***will allow Apple to detect known CSAM images stored in iCloud Photos [by] performing on-device matching using a database of known CSAM image hashes.”***
3. *“Siri and Search are also being updated to intervene when users perform searches for queries related to CSAM.”*

Big Deal



Expanded Protections for Children

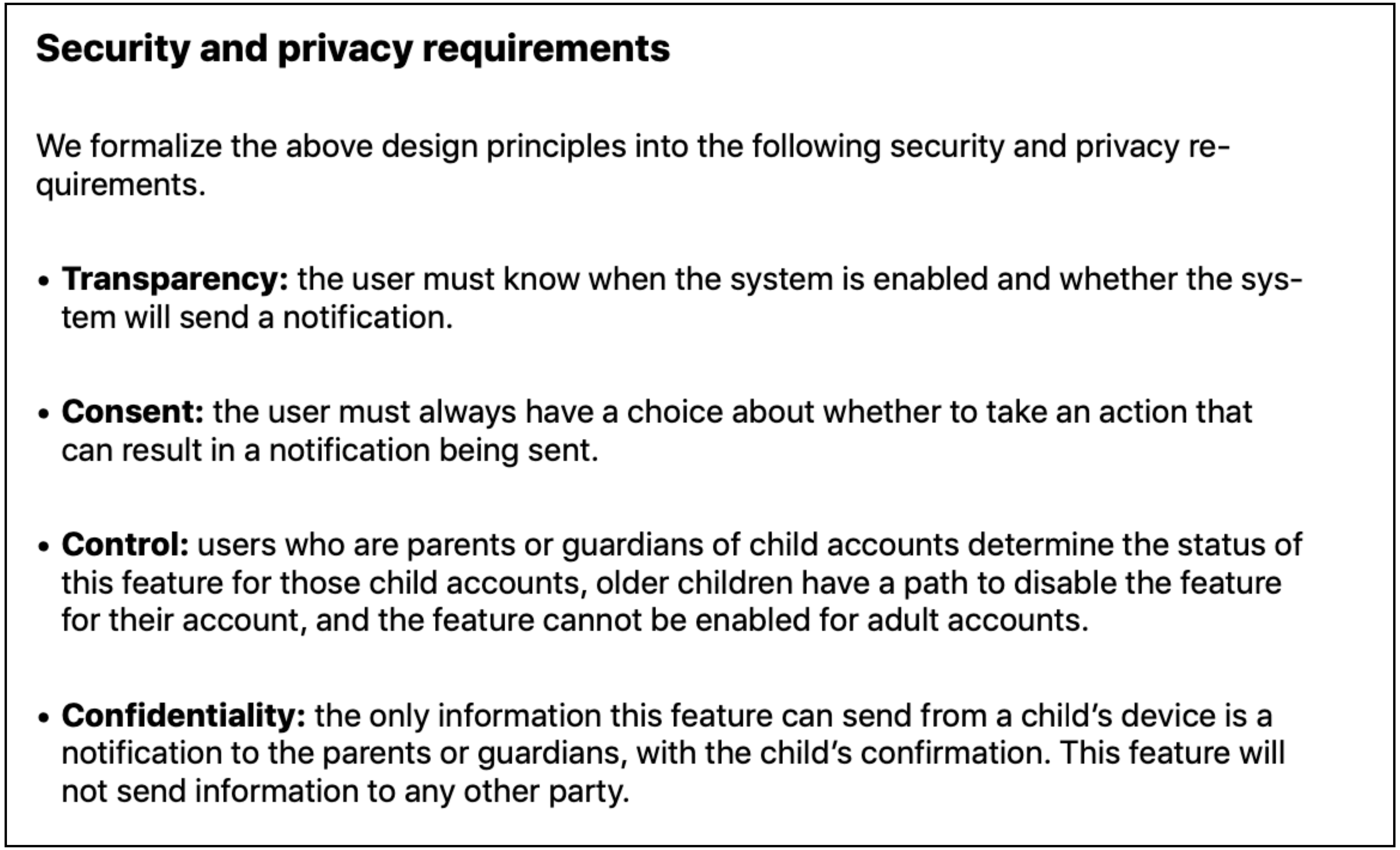
At Apple, our goal is to create technology that empowers people and enriches their lives — while helping them stay safe. We want to help protect children from predators who use communication tools to recruit and exploit them, and limit the spread of Child Sexual Abuse Material (CSAM).

Apple is introducing new child safety n three areas, developed in tion with child safety experts. First, new communication tools will enable parents to play a more informed role in helping their children navigate communication online. The Messages app will use on-device machine learning to warn about sensitive content, while keeping private communications unreadable by Apple.

Messages

Proposed Changes

- Covered under the “Parental Controls” umbrella.
- Doesn’t send silent reports to the police.
- Not a bad design.



Security and privacy requirements

We formalize the above design principles into the following security and privacy requirements.

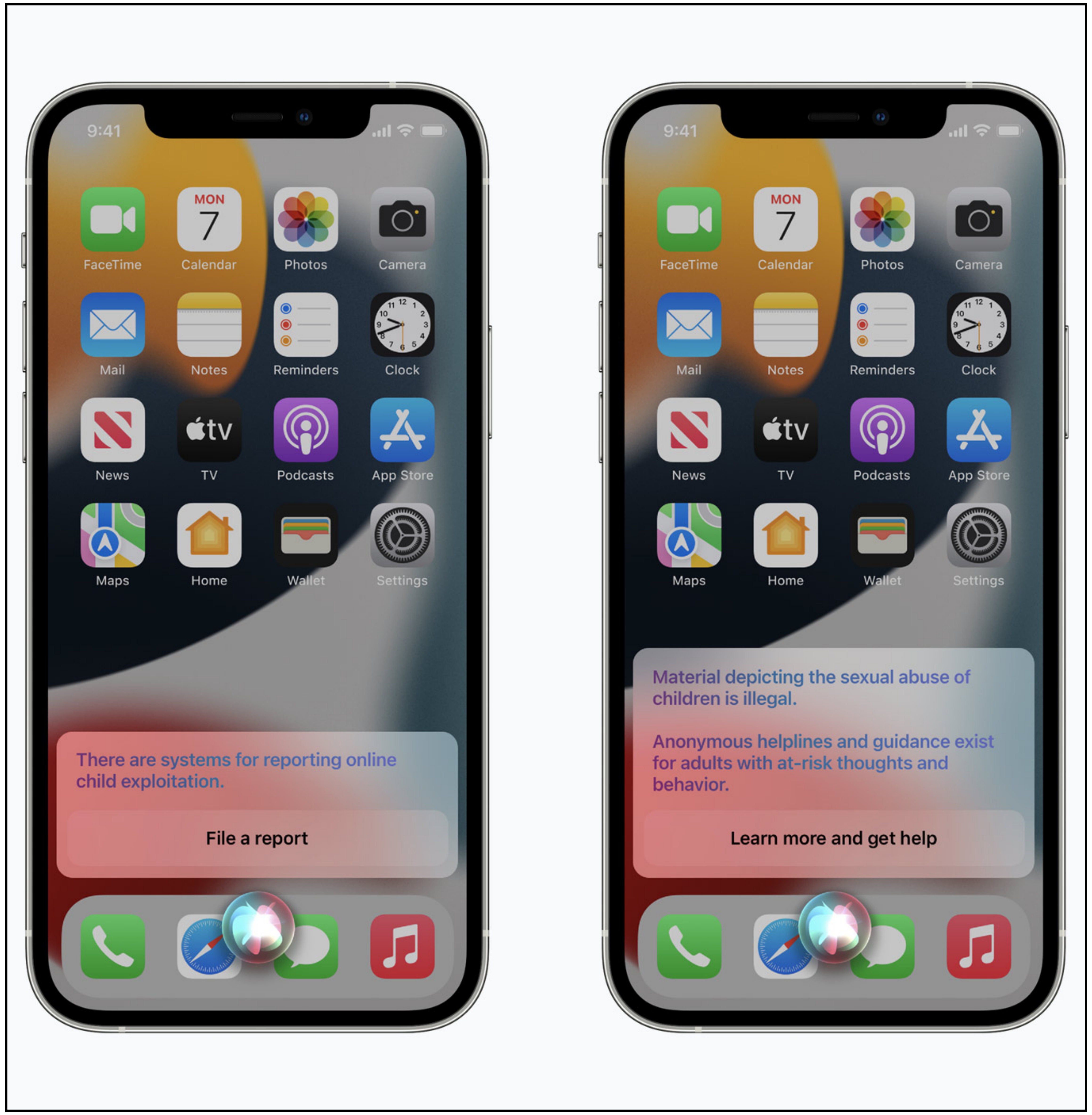
- **Transparency:** the user must know when the system is enabled and whether the system will send a notification.
- **Consent:** the user must always have a choice about whether to take an action that can result in a notification being sent.
- **Control:** users who are parents or guardians of child accounts determine the status of this feature for those child accounts, older children have a path to disable the feature for their account, and the feature cannot be enabled for adult accounts.
- **Confidentiality:** the only information this feature can send from a child’s device is a notification to the parents or guardians, with the child’s confirmation. This feature will not send information to any other party.

Siri

Proposed Changes

- Siri’s answer sets are expanded to include answers relevant to CSAM queries.

- Absolutely nothing objectionable in this proposal.



There are systems for reporting online child exploitation.

Material depicting the sexual abuse of children is illegal.

Anonymous helplines and guidance exist for adults with at-risk thoughts and behavior.

CSAM Detection

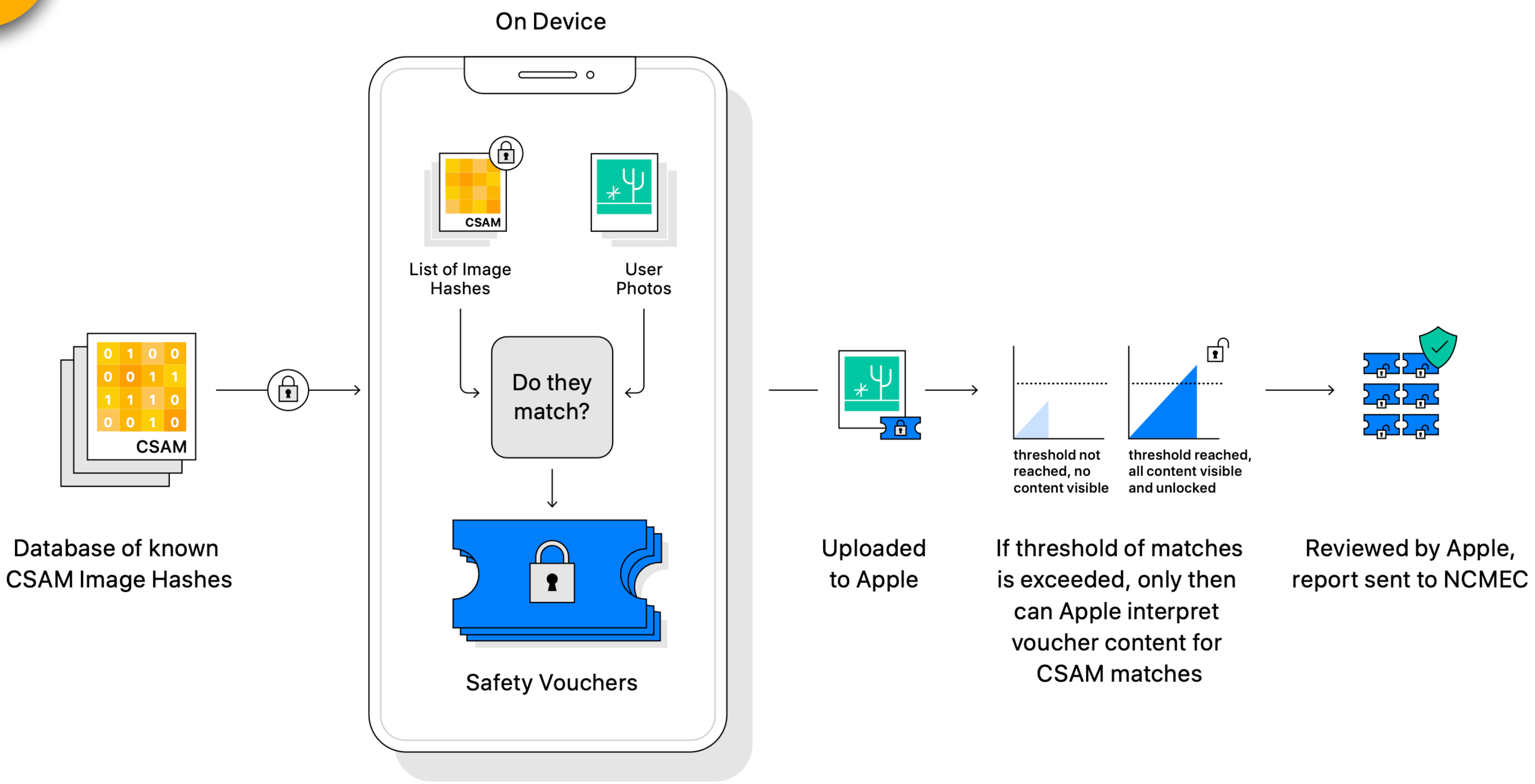
iCloud Photos

- iPhones that use iCloud Photos (pretty much all iPhones) will have *silent*, *automated*, *local* scanning of their photos to detect “objectionable content” (in this case, CSAM).

- If CSAM is found, report is sent to Apple and, potentially, the authorities.

1

How does it work?



On Device

List of Image Hashes

User Photos

Do they match?

Safety Vouchers

Database of known CSAM Image Hashes

Uploaded to Apple

threshold not reached, no content visible

threshold reached, all content visible and unlocked

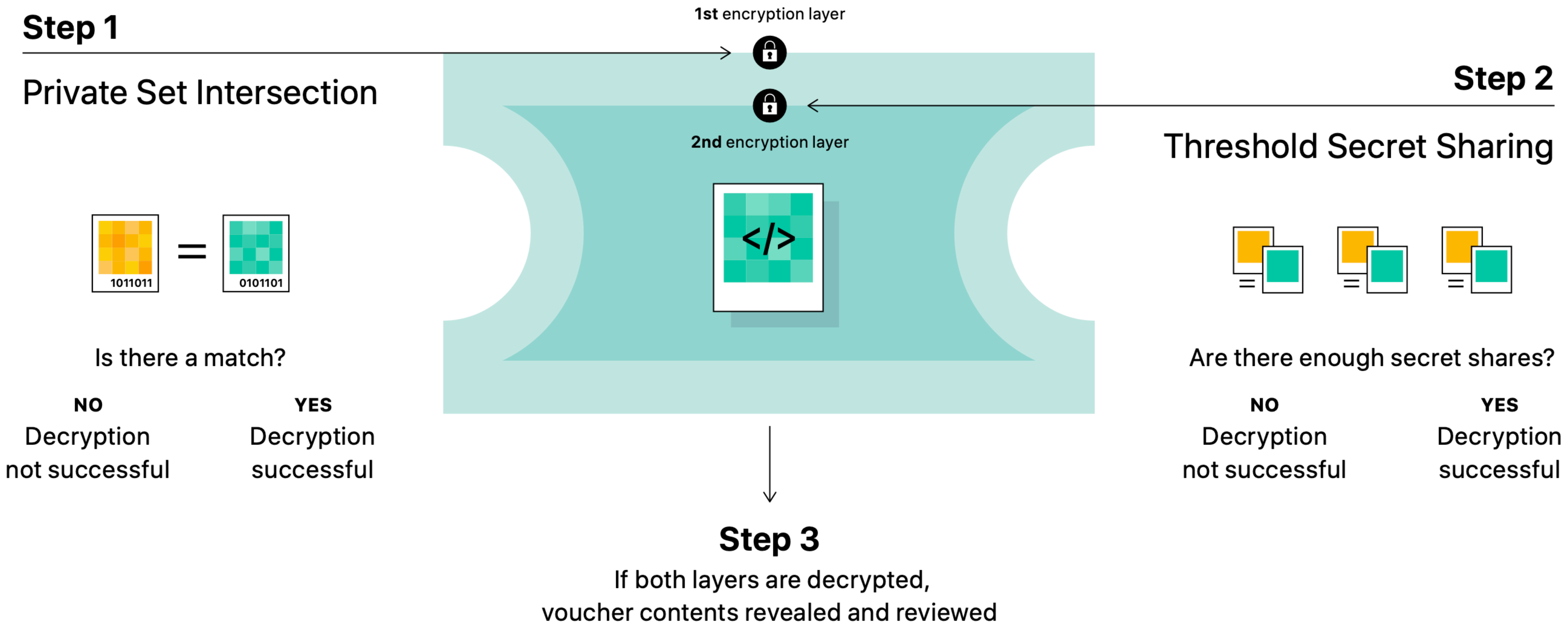
If threshold of matches is exceeded, only then can Apple interpret voucher content for CSAM matches

Reviewed by Apple, report sent to NCMEC

Source: Apple

2

How does it work?



Step 1

Private Set Intersection

Is there a match?

NO
Decryption
not successful

YES
Decryption
successful

Step 2

Threshold Secret Sharing

Are there enough secret shares?

NO
Decryption
not successful

YES
Decryption
successful

Step 3

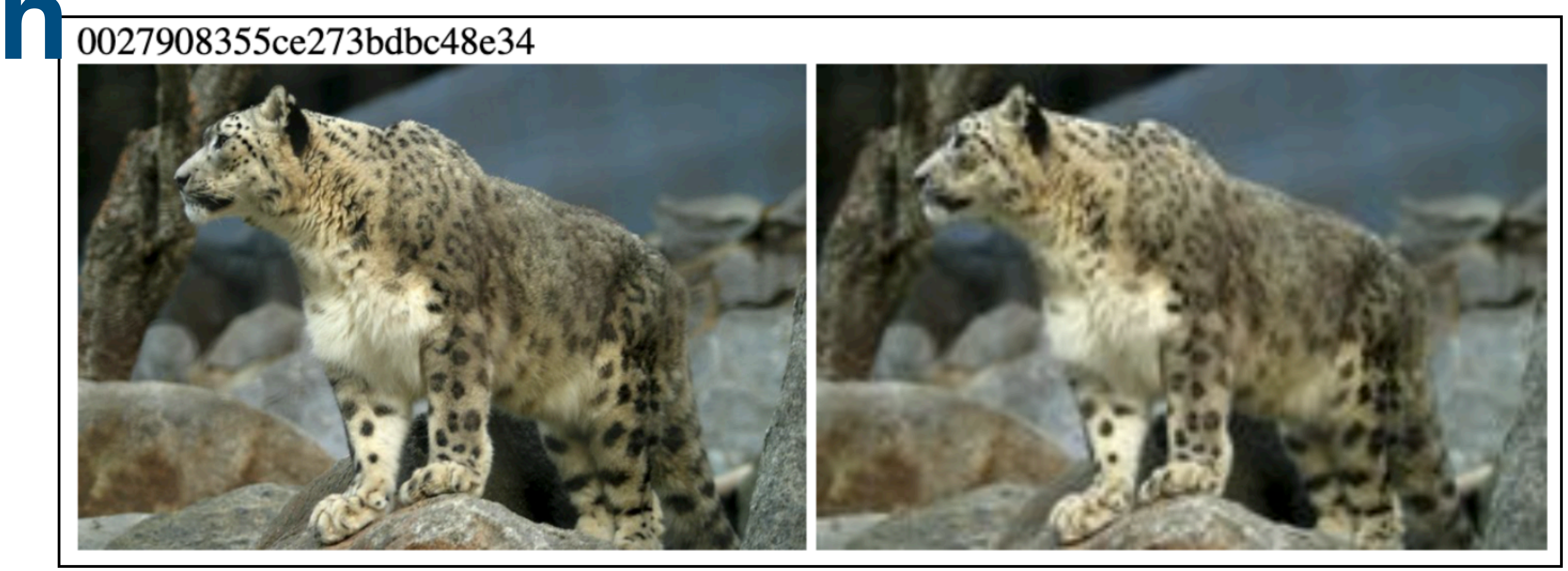
If both layers are decrypted,
voucher contents revealed and reviewed

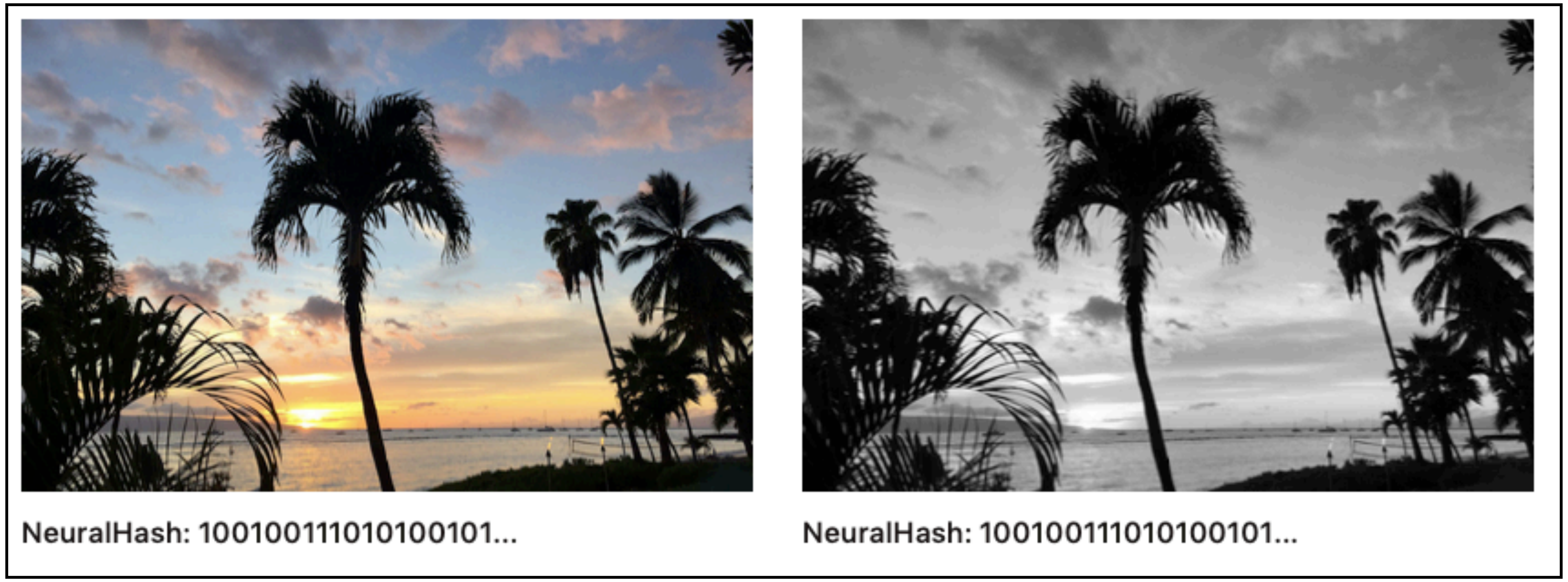
Source: Apple

CSAM Detection

NeuralHash

- “Perceptual hashing”
- Able to detect if two images are “basically the same image” (regardless of cropping, compression, color differences…)

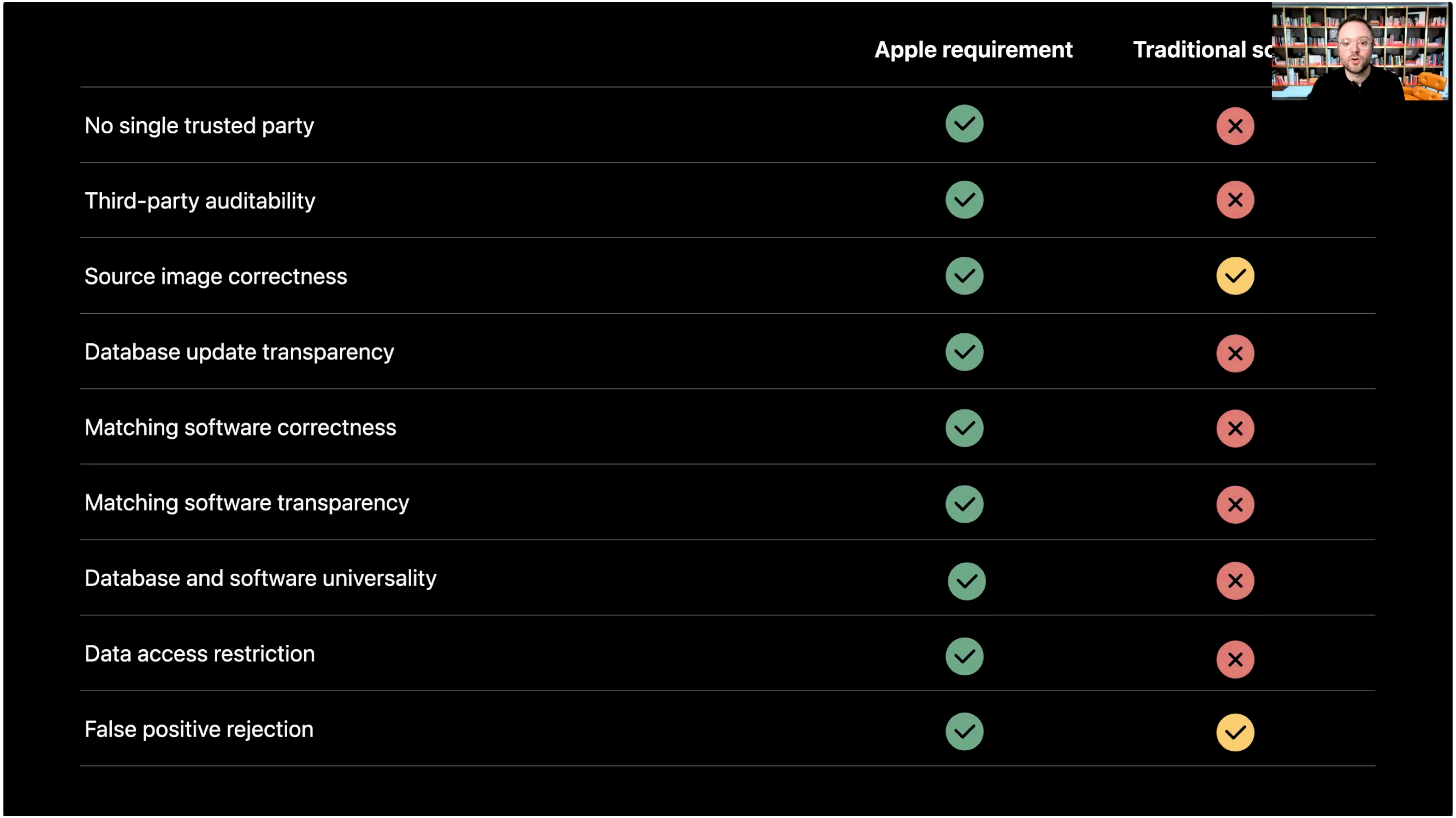



NeuralHash: 100100111010100101...
NeuralHash: 100100111010100101...

2. Arguments Against

Part 1: Technical Criticism

(Not the important part)



	Apple requirement	Traditional sc
No single trusted party	✓	✗
Third-party auditability	✓	✗
Source image correctness	✓	✓
Database update transparency	✓	✗
Matching software correctness	✓	✗
Matching software transparency	✓	✗
Database and software universality	✓	✗
Data access restriction	✓	✗
False positive rejection	✓	✓

Source: Apple, USENIX Security Symposium 2021 Event

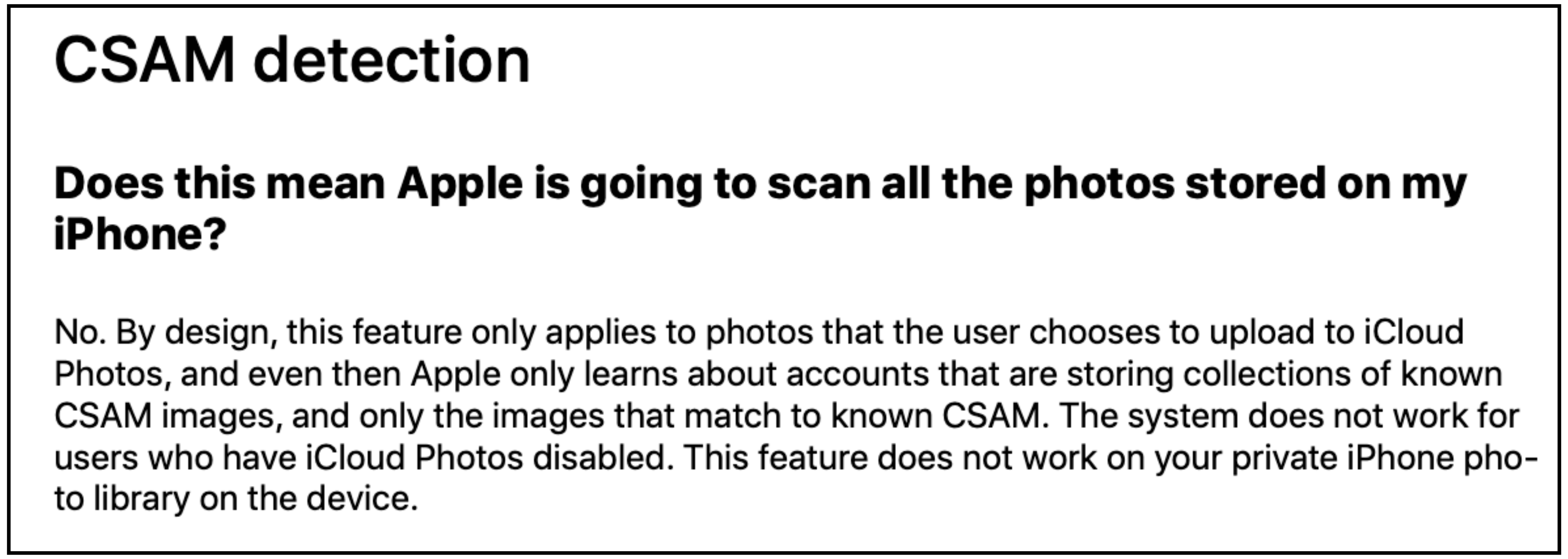
“Won’t scan all photos”

Misleading security promise 1

- Who doesn’t have iCloud Photos enabled?

- **The scanning technology does not depend on iCloud Photos: it works completely locally.**

- The same technology can be used to scan all local photos. *iCloud Photos being enabled is purely a courtesy*.



CSAM detection

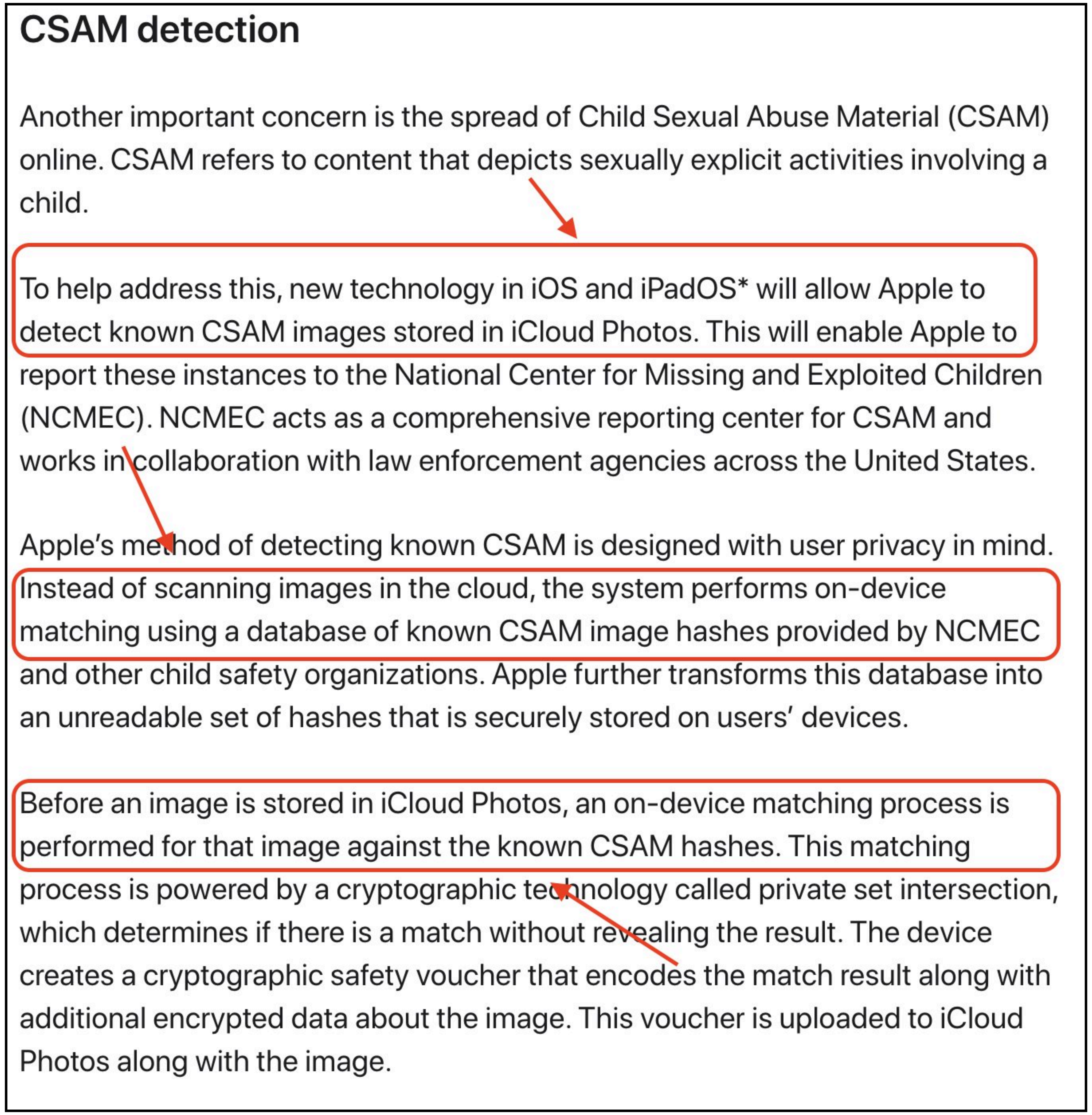
Does this mean Apple is going to scan all the photos stored on my iPhone?

No. By design, this feature only applies to photos that the user chooses to upload to iCloud Photos, and even then Apple only learns about accounts that are storing collections of known CSAM images, and only the images that match to known CSAM. The system does not work for users who have iCloud Photos disabled. This feature does not work on your private iPhone photo library on the device.

“Won’t scan all photos”

Misleading security promise 1

- *“We're not looking for CSAM on iPhones. [...] The sound-bite that got out early was "Apple is scanning my phone for images." This is not what's happening”* — **Craig Federighi**, Apple Senior Vice President of Software Engineering



CSAM detection

Another important concern is the spread of Child Sexual Abuse Material (CSAM) online. CSAM refers to content that depicts sexually explicit activities involving a child.

To help address this, new technology in iOS and iPadOS* will allow Apple to detect known CSAM images stored in iCloud Photos. This will enable Apple to report these instances to the National Center for Missing and Exploited Children (NCMEC). NCMEC acts as a comprehensive reporting center for CSAM and works in collaboration with law enforcement agencies across the United States.

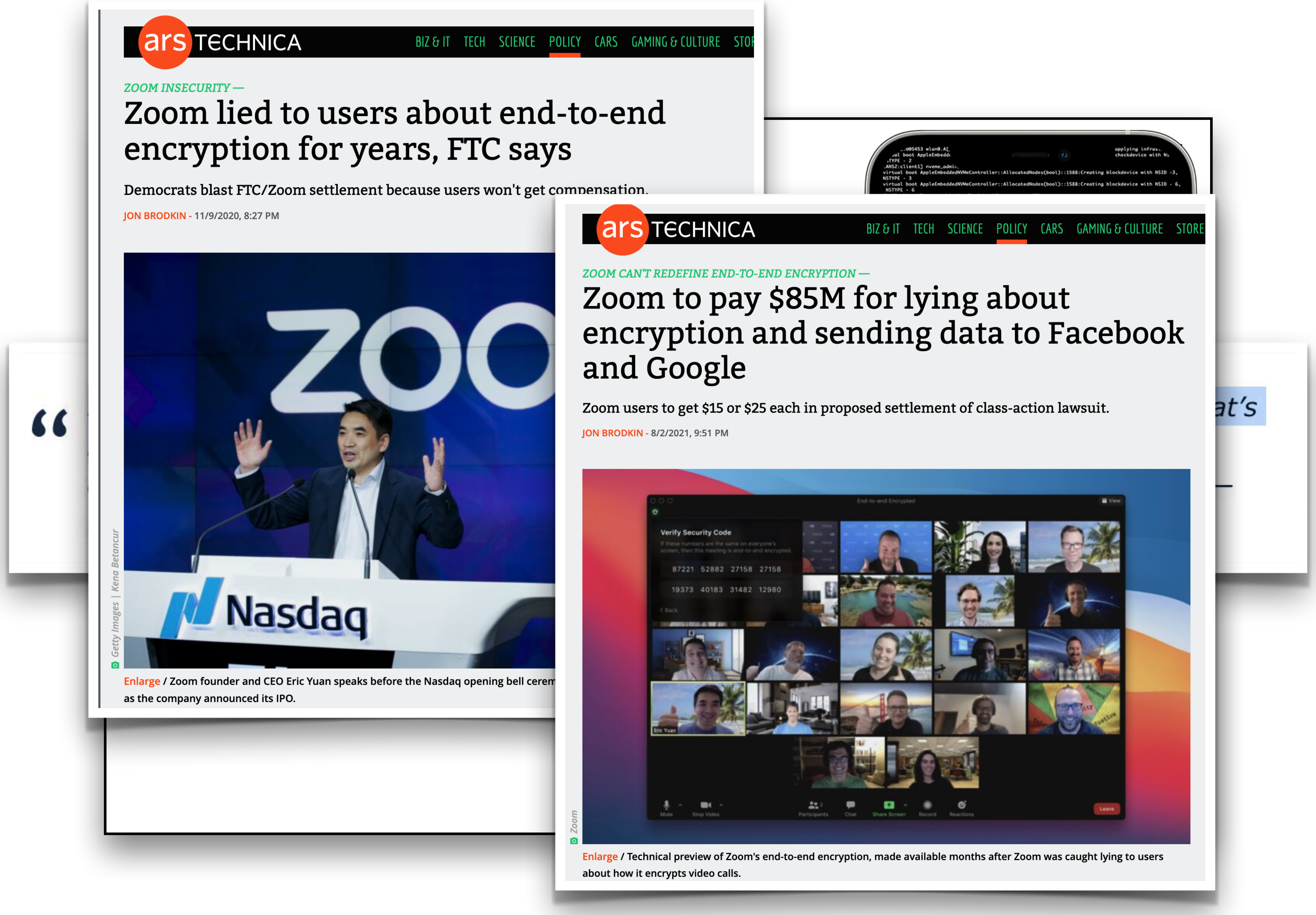
Apple’s method of detecting known CSAM is designed with user privacy in mind. Instead of scanning images in the cloud, the system performs on-device matching using a database of known CSAM image hashes provided by NCMEC and other child safety organizations. Apple further transforms this database into an unreadable set of hashes that is securely stored on users’ devices.

Before an image is stored in iCloud Photos, an on-device matching process is performed for that image against the known CSAM hashes. This matching process is powered by a cryptographic technology called private set intersection, which determines if there is a match without revealing the result. The device creates a cryptographic safety voucher that encodes the match result along with additional encrypted data about the image. This voucher is uploaded to iCloud Photos along with the image.

“Auditability”

Misleading security promise 2

- iOS is the most closed consumer computing ecosystem in the world.
- Obtaining an iPhone you can “peer into” is subject to loaning one from Apple, approved on an individual basis.
- Incredibly closed ecosystem, programs, specs, filesystem, code, standards…



ZOOM INSECURITY —

Zoom lied to users about end-to-end encryption for years, FTC says

Democrats blast FTC/Zoom settlement because users won't get compensation.

JON BRODKIN - 11/9/2020, 8:27 PM

Enlarge / Zoom founder and CEO Eric Yuan speaks before the Nasdaq opening bell cerem as the company announced its IPO.

ZOOM CAN'T REDEFINE END-TO-END ENCRYPTION —

Zoom to pay $85M for lying about encryption and sending data to Facebook and Google

Zoom users to get $15 or $25 each in proposed settlement of class-action lawsuit.

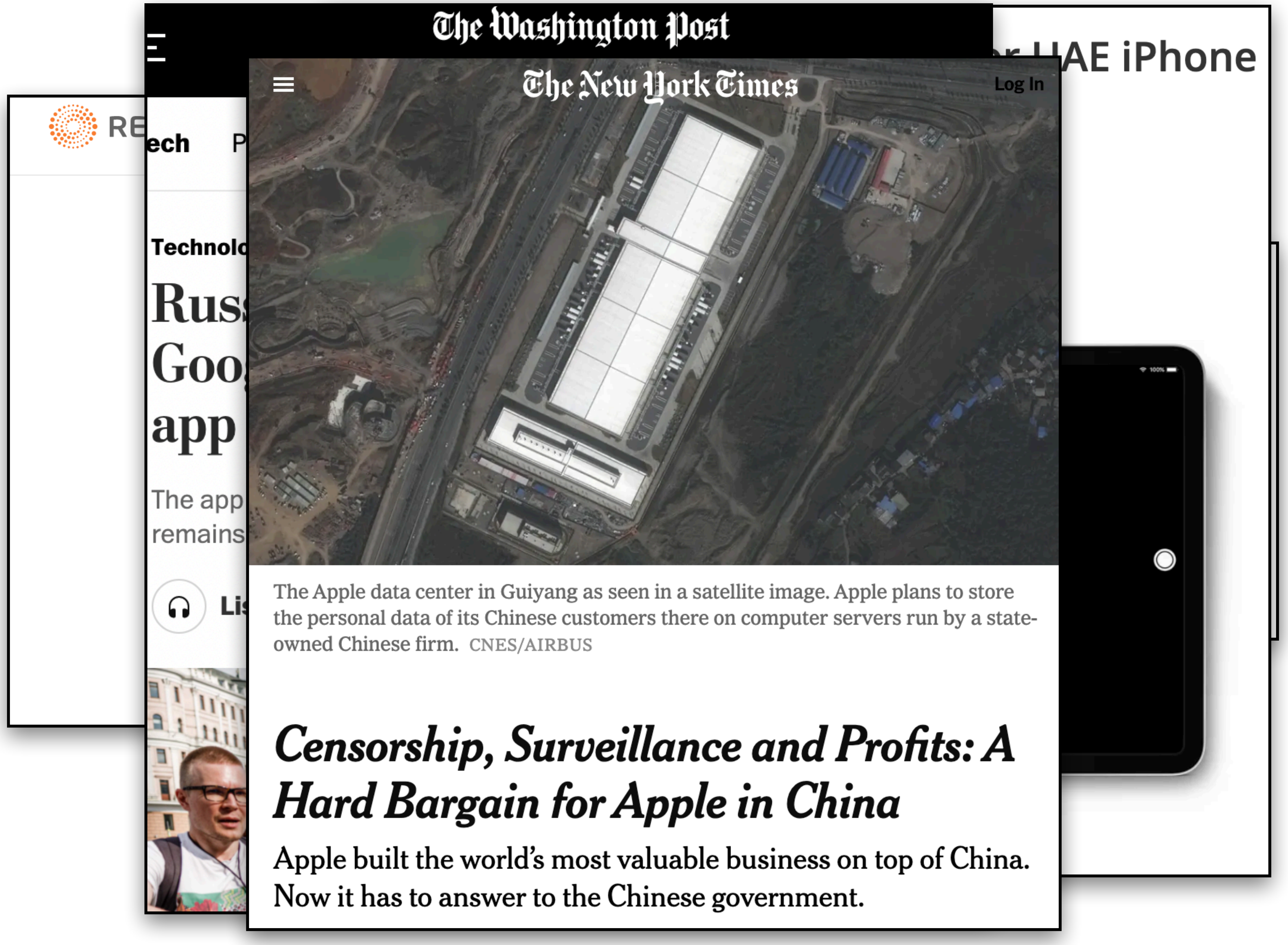
JON BRODKIN - 8/2/2021, 9:51 PM

Enlarge / Technical preview of Zoom's end-to-end encryption, made available months after Zoom was caught lying to users about how it encrypts video calls.

“Not subject to pressure”

Misleading security promise 3

- Apple has already canceled plans to encrypt iCloud backups after the FBI complained.
- Apple does not provide FaceTime in the United Arab Emirates due to government complaints about its end-to-end encryption.
- Apple pulled an app by the Russian democratic opposition that gave citizens voting information after pressure from the Kremlin (JUST A FEW WEEKS AGO!)
- Apple’s entire cloud infrastructure in China is subject to tight controls by the Chinese government.



The Apple data center in Guiyang as seen in a satellite image. Apple plans to store the personal data of its Chinese customers there on computer servers run by a state-owned Chinese firm. CNES/AIRBUS

Censorship, Surveillance and Profits: A Hard Bargain for Apple in China

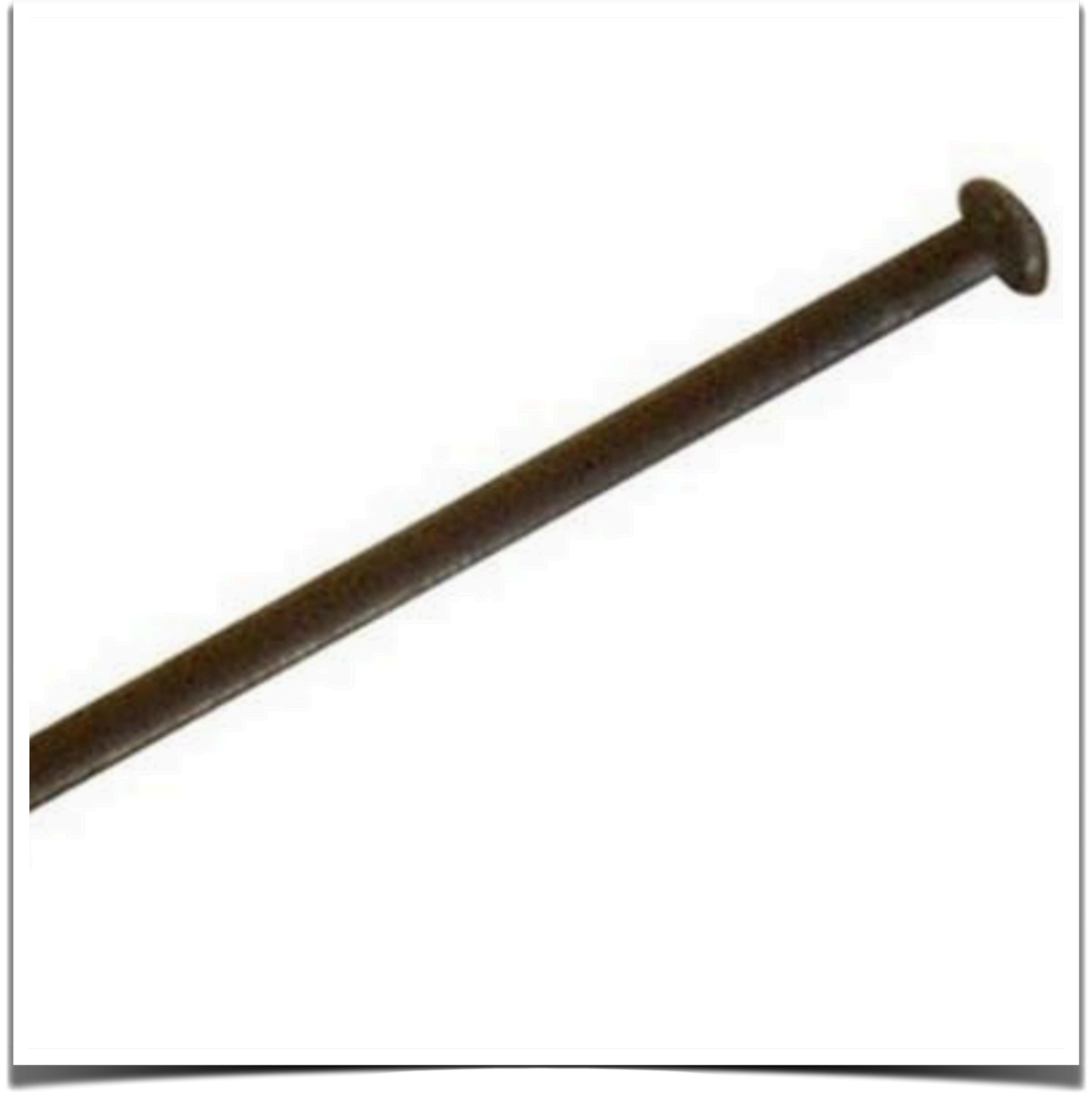
Apple built the world’s most valuable business on top of China. Now it has to answer to the Chinese government.

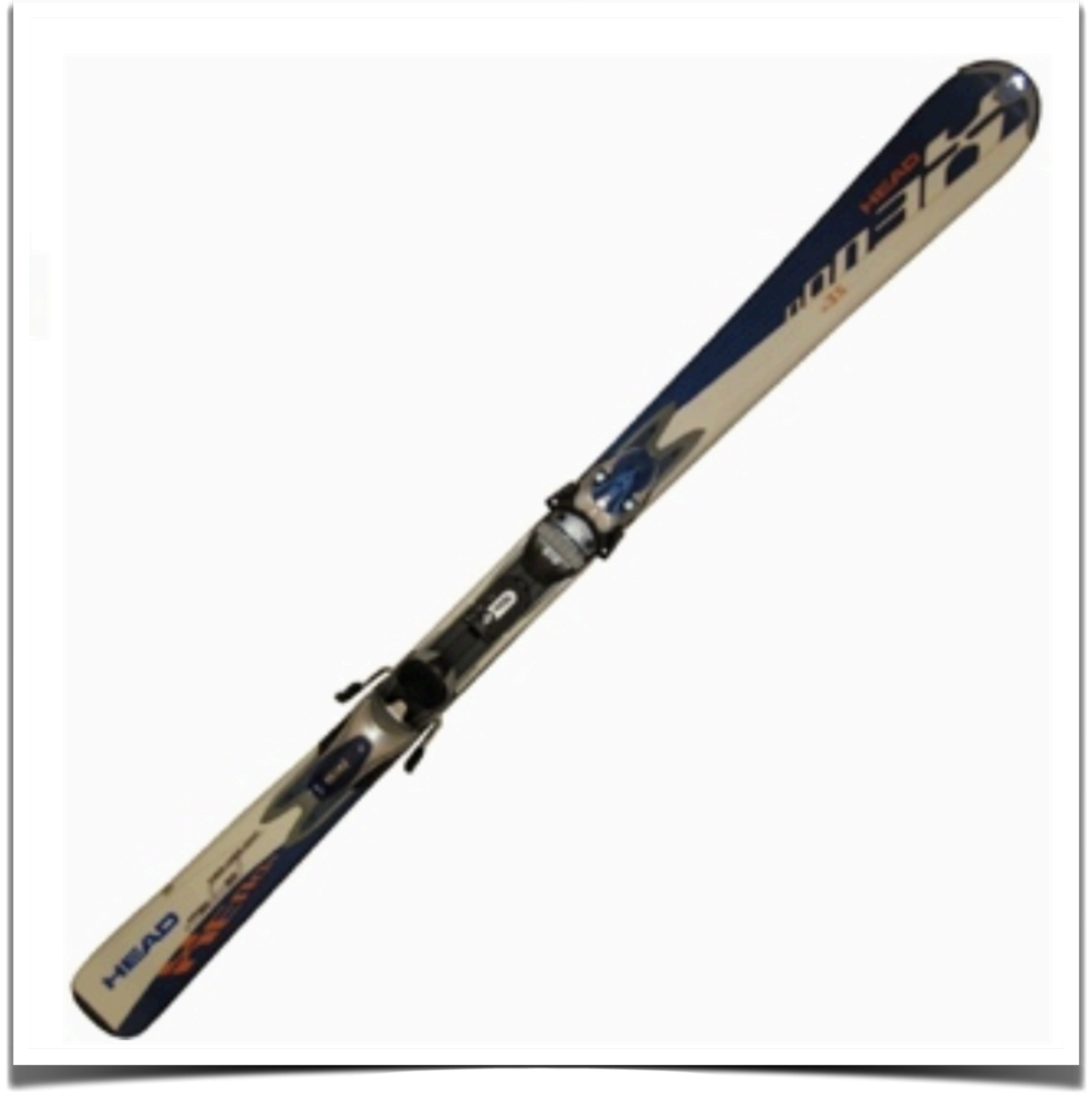
> “Sometimes, Philip Shoemaker, who ran Apple’s App Store from 2009 to 2016, would be awakened in the middle of the night with demands from the Chinese government to remove an app.”

“Censorship, Surveillance and Profits: A Hard Bargain for Apple in China”— The New York Times

> “Chinese intelligence has physical control over your hardware — that’s basically a threat level you can’t let it get to.”

Prof. Matthew D. Green, Johns Hopkins University The New York Times





These are the same photo

(According to NeuralHash).

NeuralHash Forgeries

Do they matter?

- **Potentially!**
  - WhatsApp saves received images to iCloud Photos by default.
  - Attacker sends you a dozen “puppy photos” that secretly match NeuralHash hashes.

- **Apple’s countermeasures:**
  1. Hash database is secret.
  2. Two types of perceptual hashing (?)
  3. Manual human review.

2. Arguments Against

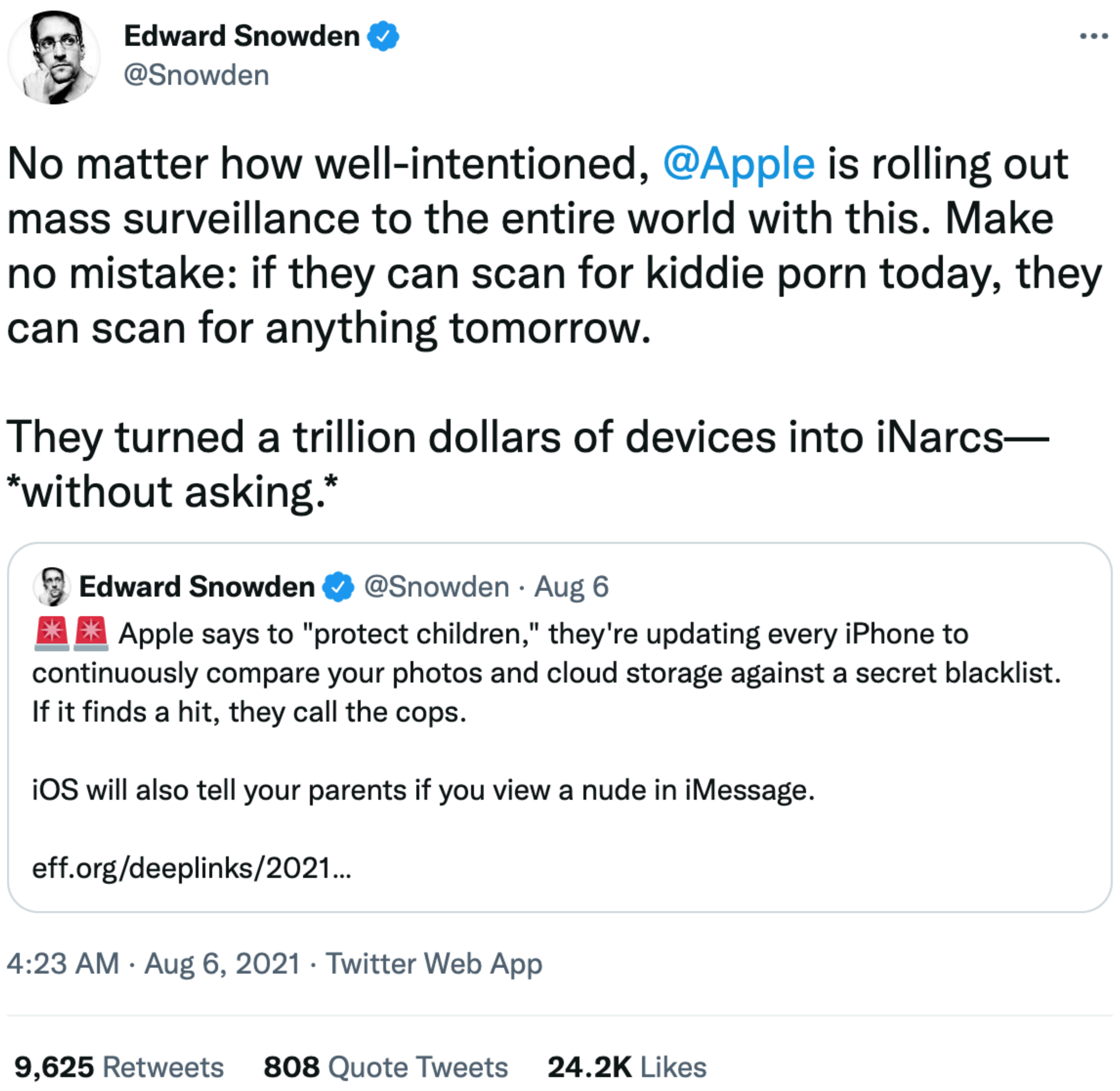
Part 2: Civic Criticism

(The important part)

Local scanning for bad content

Questions regarding personal privacy

- Nobody expects their photo albums to secretly watch their content and snitch to the police if objectionable content is found.
- Unlike Google Photos etc., Apple's scanning is totally local.
  - In Google's case, scanning happens on their servers.
  - It's not your local phone in your house that's watching itself for objectionable content and quietly informing the authorities.
  - *iCloud Photos being enabled is purely a courtesy.*



Edward Snowden
@Snowden

No matter how well-intentioned, @Apple is rolling out mass surveillance to the entire world with this. Make no mistake: if they can scan for kiddie porn today, they can scan for anything tomorrow.

They turned a trillion dollars of devices into iNarcs—*without asking.*

> Edward Snowden @Snowden · Aug 6
>
> 🚨🚨 Apple says to "protect children," they're updating every iPhone to continuously compare your photos and cloud storage against a secret blacklist. If it finds a hit, they call the cops.
>
> iOS will also tell your parents if you view a nude in iMessage.
>
> eff.org/deeplinks/2021...

4:23 AM · Aug 6, 2021 · Twitter Web App

9,625 Retweets **808** Quote Tweets **24.2K** Likes

Local scanning for bad content

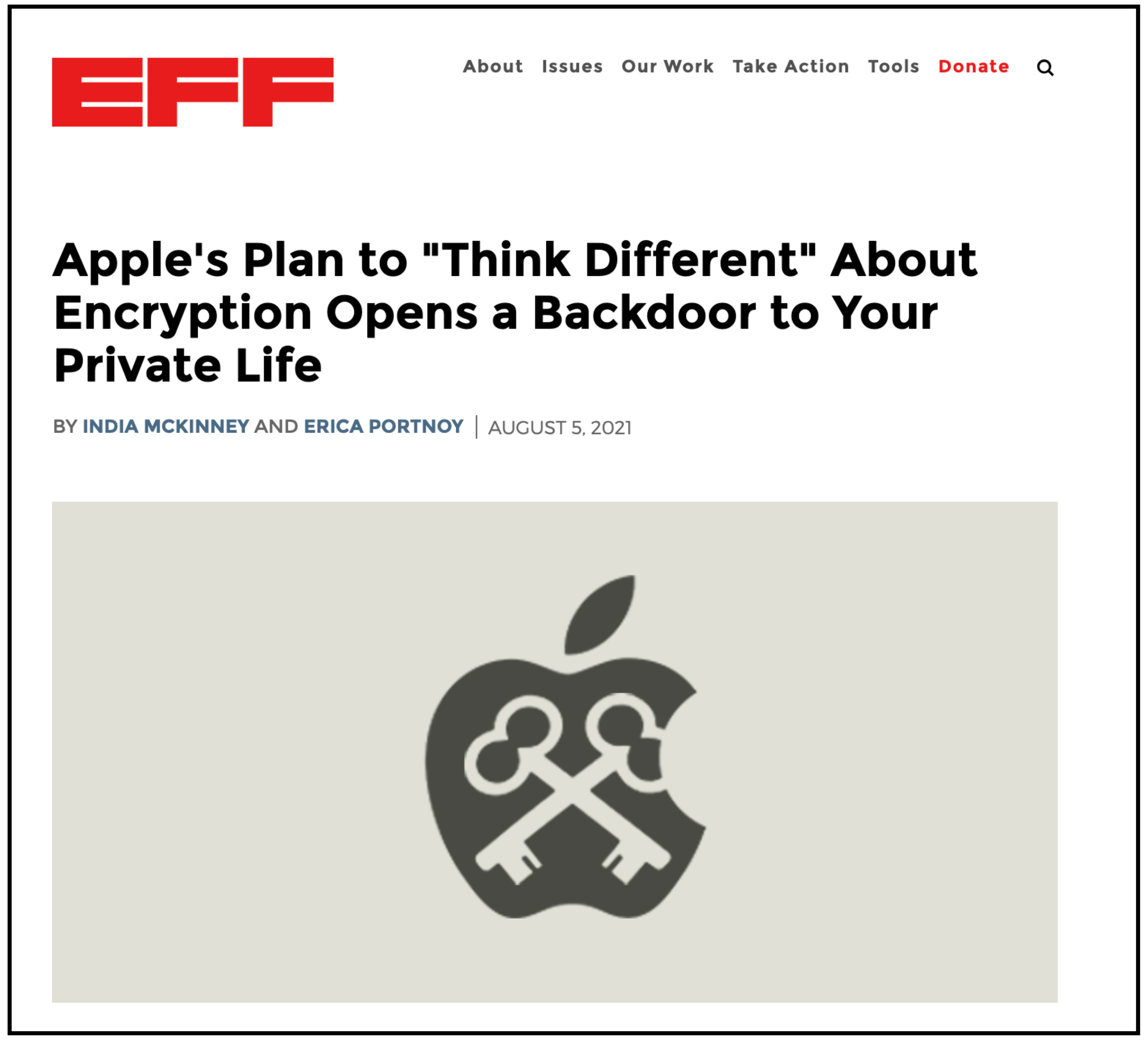
Questions regarding personal privacy

- Would you buy a safe if the manufacturer insisted on passing by once a month when you’re not home to check that you’re not putting drugs inside?
- Is it OK for Alexa to secretly call the cops if you ask for instructions on how to commit a crime?

Mission Creep

CSAM Databases used outside of their scope

*“One of the technologies originally built to scan and hash child sexual abuse imagery has been repurposed to create a database of “terrorist” content that companies can contribute to and access for the purpose of banning such content. The database, managed by the Global Internet Forum to Counter Terrorism (**GIFCT**), is troublingly without external oversight, despite calls from civil society. While it’s therefore impossible to know whether the database has overreached, we do know that platforms regularly flag critical content as “terrorism,” including documentation of violence and repression, counterspeech, art, and satire.”*



Apple's Plan to "Think Different" About Encryption Opens a Backdoor to Your Private Life

BY INDIA MCKINNEY AND ERICA PORTNOY | AUGUST 5, 2021

GIFCT
Global Internet Forum
to Counter Terrorism

“Stop using encryption so we can can check your messages for criminal activity” becomes “Allow us to scan all the files on your computer for criminal activity”.

Sarah Jamie Lewis, Director, OpenPrivacy.ca

“How long do you think it will be before the database is expanded to include “terrorist” content? “harmful-but-legal” content? **state-specific censorship**?”

Reminder: homosexuality and insulting the King are both punishable in Saudi Arabia.

Sarah Jamie Lewis, Director, OpenPrivacy.ca

“The pressure is going to come from the UK, from the US, from India, from *China*. I’m terrified about what that’s going to look like.”

Prof. Matthew D. Green, Johns Hopkins University

What happens when the bad guys simply disable iCloud Photos?

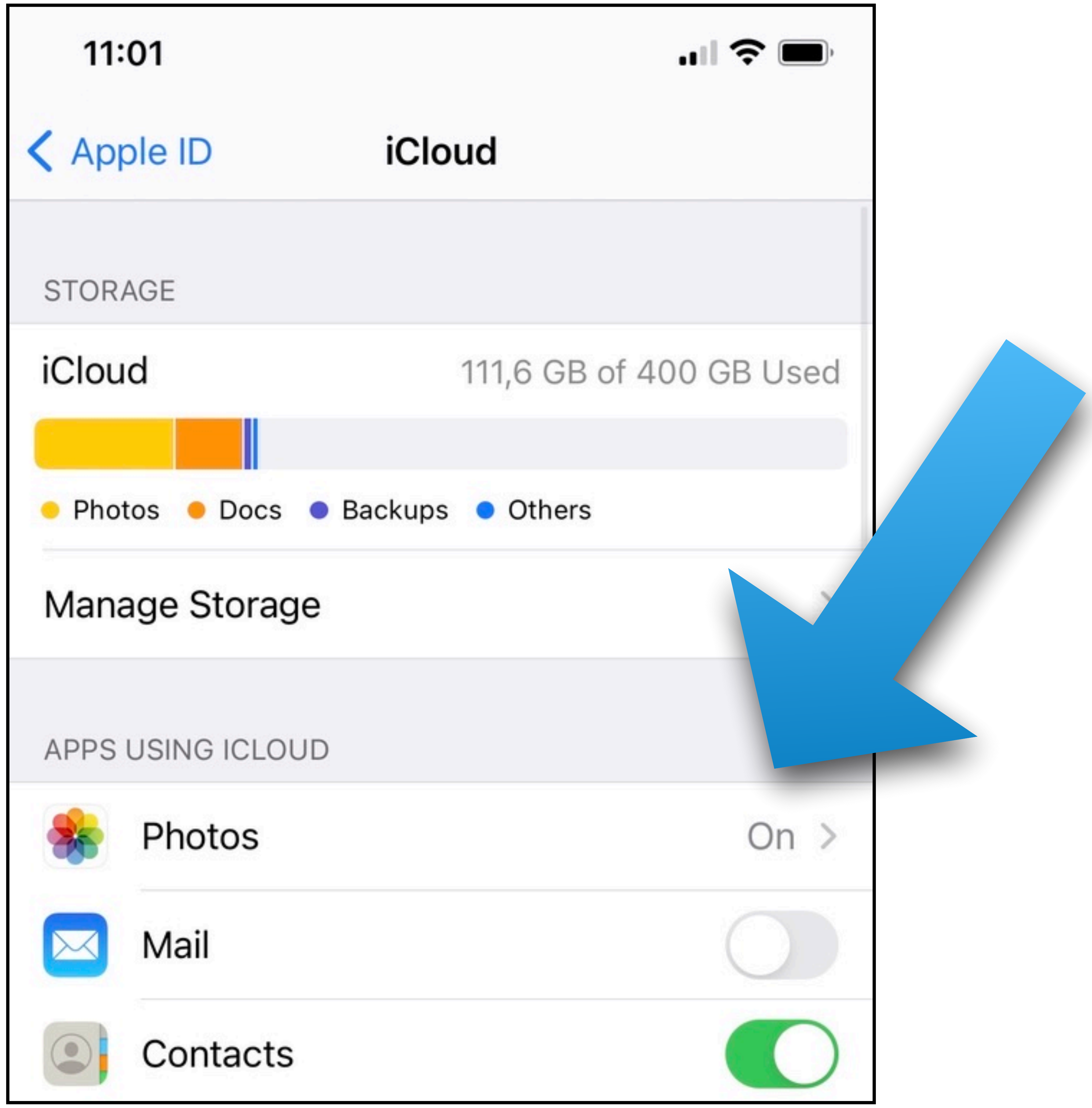
1. *“This was working great until the bad guys caught up!”*

2. *“Oh, the content scanning technology is local! It doesn’t depend on iCloud Photos being enabled in the first place!”*

3. **Local content scanning regardless of cloud service now becomes the norm.**

4. **Mission creep broadens the scope.**

5. We’re screwed.



11:01

Apple ID

iCloud

STORAGE

iCloud 111,6 GB of 400 GB Used

● Photos ● Docs ● Backups ● Others

Manage Storage

APPS USING ICLOUD

Photos On

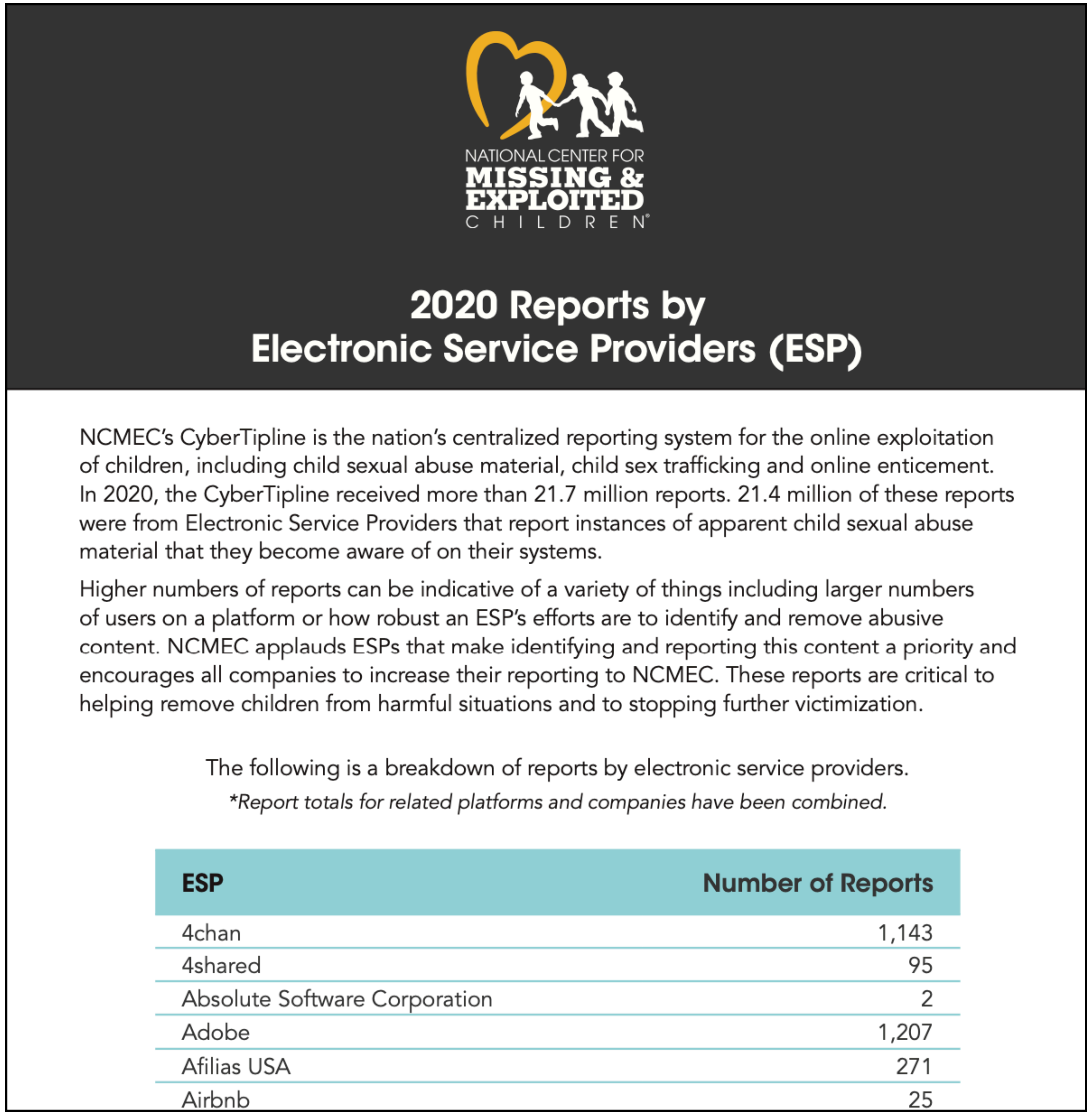
Mail ○

Contacts ●

3. Arguments For

Comparing Apples to oranges

- NCMEC’s 2020 “Reports by Electronic Service Providers” claims Facebook gave 20,307,216 reports, Dropbox gave 20,928 reports, while Apple only gave 265 reports.
- Not a valid comparison. Facebook is used to publish. Dropbox is used to share files. Apple’s services are almost entirely private!



NATIONAL CENTER FOR MISSING & EXPLOITED CHILDREN®

2020 Reports by Electronic Service Providers (ESP)

NCMEC’s CyberTipline is the nation’s centralized reporting system for the online exploitation of children, including child sexual abuse material, child sex trafficking and online enticement. In 2020, the CyberTipline received more than 21.7 million reports. 21.4 million of these reports were from Electronic Service Providers that report instances of apparent child sexual abuse material that they become aware of on their systems.

Higher numbers of reports can be indicative of a variety of things including larger numbers of users on a platform or how robust an ESP’s efforts are to identify and remove abusive content. NCMEC applauds ESPs that make identifying and reporting this content a priority and encourages all companies to increase their reporting to NCMEC. These reports are critical to helping remove children from harmful situations and to stopping further victimization.

The following is a breakdown of reports by electronic service providers.

**Report totals for related platforms and companies have been combined.*

ESP	Number of Reports
4chan	1,143
4shared	95
Absolute Software Corporation	2
Adobe	1,207
Afilias USA	271
Airbnb	25

Child Abuse is a serious problem

- Everyone wants children to be safe.
- Measures and technologies adopted focus on *short-term* child protection while ignoring *long-term* consequences on personal rights and privacy.

4. Conclusion

What happens when the bad guys simply disable iCloud Photos?

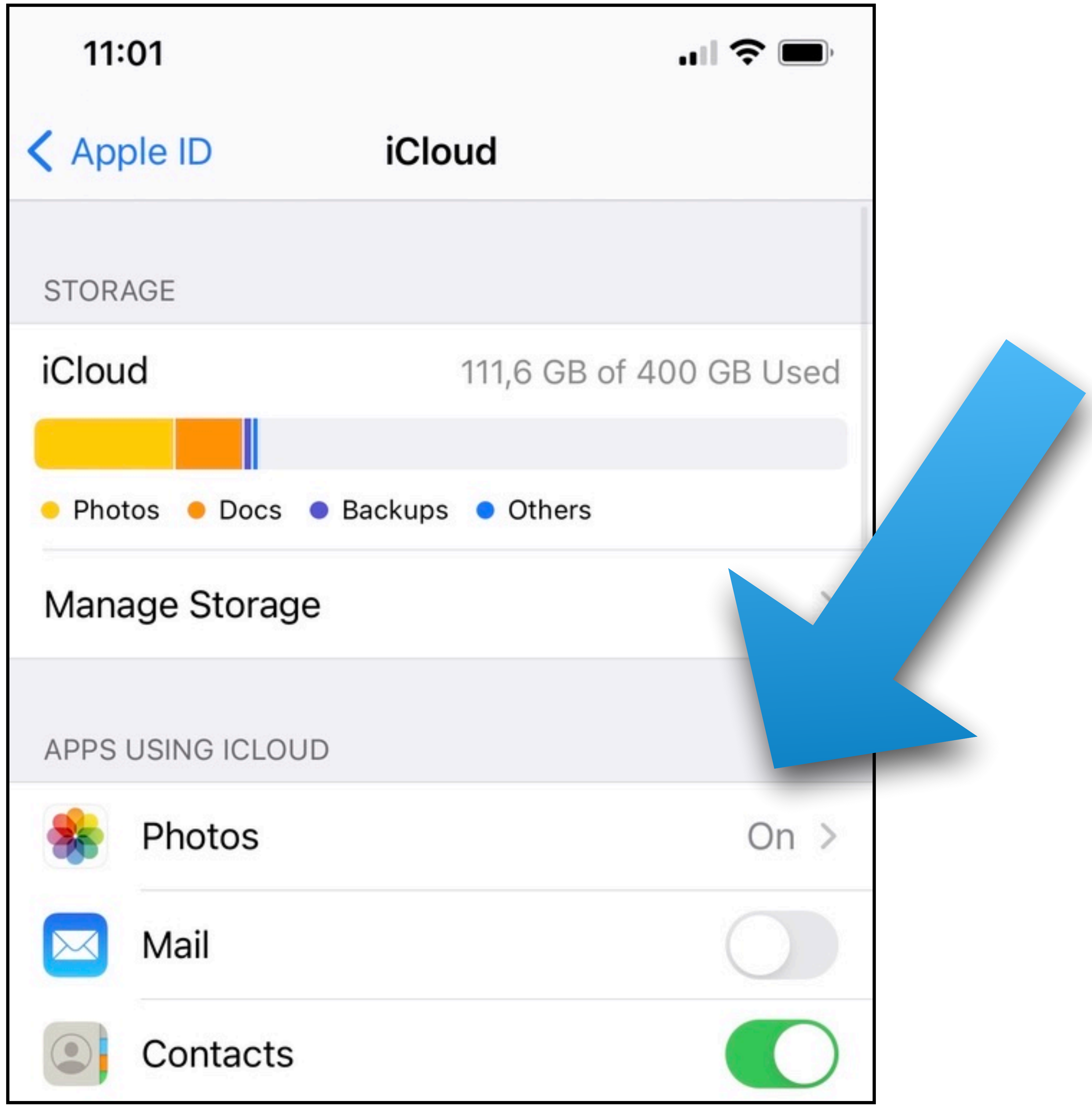
1. *“This was working great until the bad guys caught up!”*

2. *“Oh, the content scanning technology is local! It doesn’t depend on iCloud Photos being enabled in the first place!”*

3. **Local content scanning regardless of cloud service now becomes the norm.**

4. **Mission creep broadens the scope.**

5. We’re screwed.



11:01

Apple ID

iCloud

STORAGE

iCloud 111,6 GB of 400 GB Used

● Photos ● Docs ● Backups ● Others

Manage Storage

APPS USING ICLOUD

Photos On

Mail ○

Contacts ●

“I know it’s been a long day and that many of you probably haven’t slept in 24 hours. We know that the days to come will be filled with **the screeching voices of the minority**.

Our voices will be louder.”

Marita Rodriguez, Executive Director of Strategic Partnerships, National Center for Missing and Exploited Children, In a memo shared with Apple employees by Sebastien Marineau-Mes, Vice President of Software at Apple

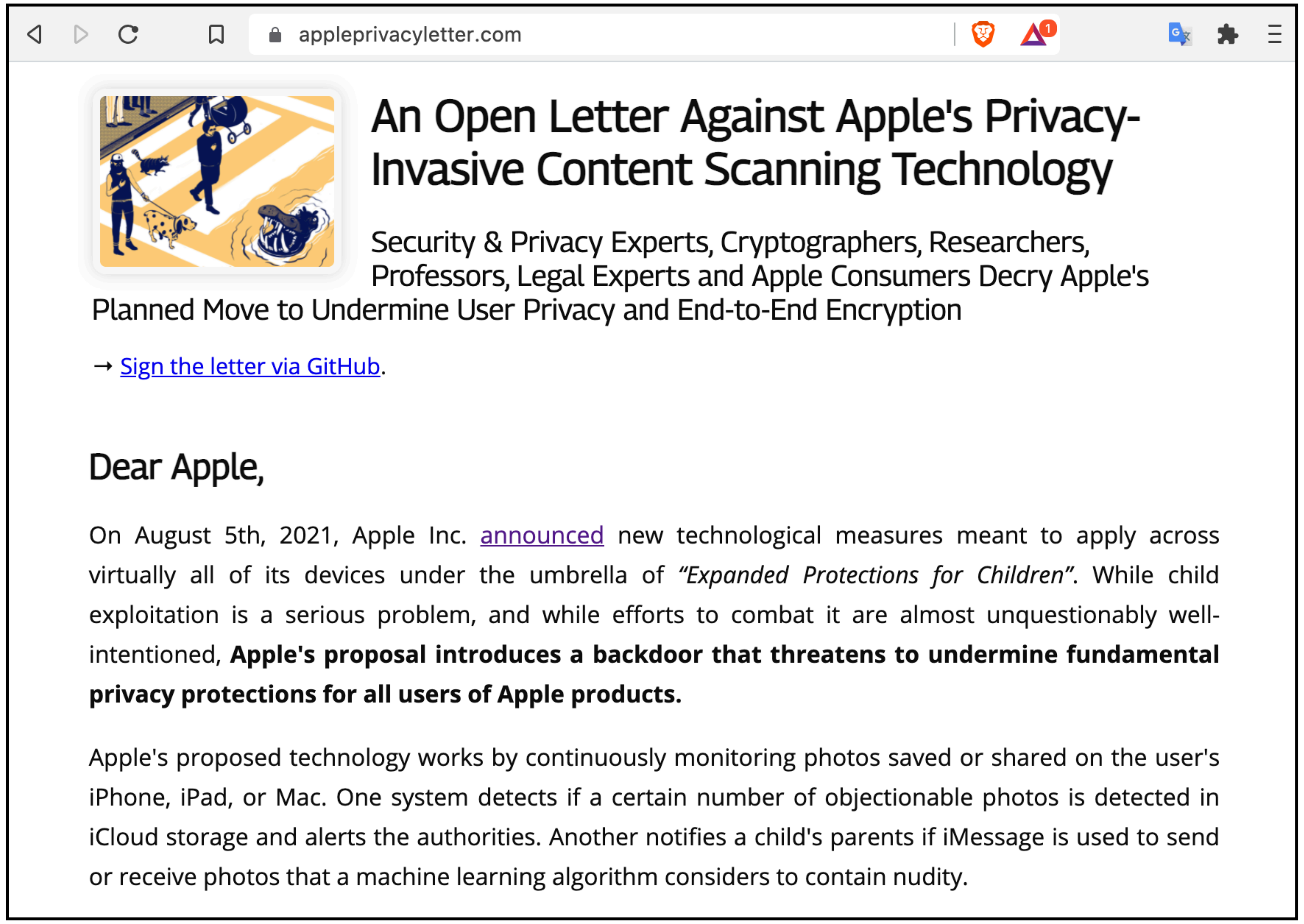
“**Update as of September 3, 2021**: Previously we announced plans for features intended to help protect children from predators who use communication tools to recruit and exploit them and to help limit the spread of Child Sexual Abuse Material. Based on feedback from customers, advocacy groups, researchers, and others, we have decided to take additional time over the coming months to collect input and make improvements before releasing these critically important child safety features.”

Apple

Civil society had an impact

ApplePrivacyLetter.com

- I wrote and organized an open letter, almost 9,000 signatures in a couple of weeks.
- EFF launched an open letter and a global coalition.
- Many experts spoke out.



An Open Letter Against Apple's Privacy-Invasive Content Scanning Technology

Security & Privacy Experts, Cryptographers, Researchers, Professors, Legal Experts and Apple Consumers Decry Apple's Planned Move to Undermine User Privacy and End-to-End Encryption

→ Sign the letter via GitHub.

Dear Apple,

On August 5th, 2021, Apple Inc. announced new technological measures meant to apply across virtually all of its devices under the umbrella of *"Expanded Protections for Children"*. While child exploitation is a serious problem, and while efforts to combat it are almost unquestionably well-intentioned, **Apple's proposal introduces a backdoor that threatens to undermine fundamental privacy protections for all users of Apple products.**

Apple's proposed technology works by continuously monitoring photos saved or shared on the user's iPhone, iPad, or Mac. One system detects if a certain number of objectionable photos is detected in iCloud storage and alerts the authorities. Another notifies a child's parents if iMessage is used to send or receive photos that a machine learning algorithm considers to contain nudity.

Private devices should remain private.

- Local content scanning can and will be abused to turn private property into constant automated policing in your pocket of your private information and files.
- Scanning something on a server vs. your devices suspecting you locally: fundamentally different.

- **Thank you for listening.**
  - Website: nadim.computer
  - Twitter: @kaepora