Rethinking the detection of child abuse on the Internet

Elie Bursztein @elie
with the help of MANY collaborators
at NCMEC, Thorn and Google

WARNING

Potentially triggering language
and difficult subject ahead

Online child sexual abuse

The production, distribution, and consumption of digital child sexual abuse material (CSAM)

Google | NATIONAL CENTER FOR MISSING & EXPLOITED CHILDREN

The online child sexual abuse ecosystem has transformed alongside technological shifts

Google | NATIONAL CENTER FOR MISSING & EXPLOITED CHILDREN

Child abusers take advantage of key technological shifts

Camera phones
easier to produce

Rich media
content diversification

Online community
easier to organize

Digital distribution
easier to share

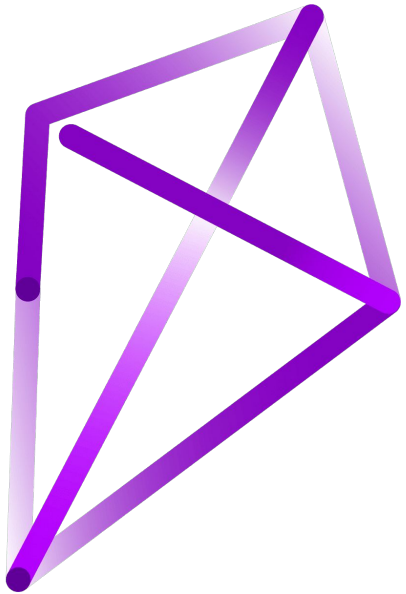
Google | NATIONAL CENTER FOR MISSING & EXPLOITED CHILDREN

We identified 5 key challenges through a multi-year data driven study

Google | NATIONAL CENTER FOR MISSING & EXPLOITED CHILDREN

To appear at WWW 2018

Let’s build technologies to create an Internet free of child abuse

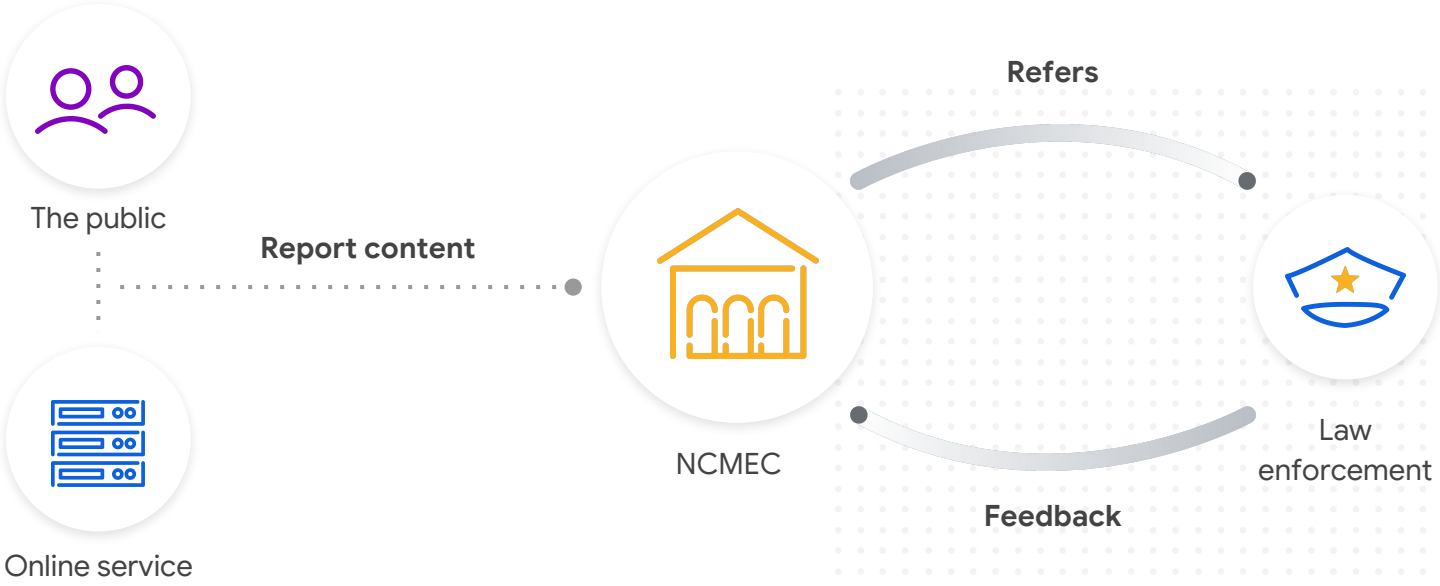


Google | NATIONAL CENTER FOR MISSING & EXPLOITED CHILDREN

NCMEC is the US clearinghouse for child abuse reports

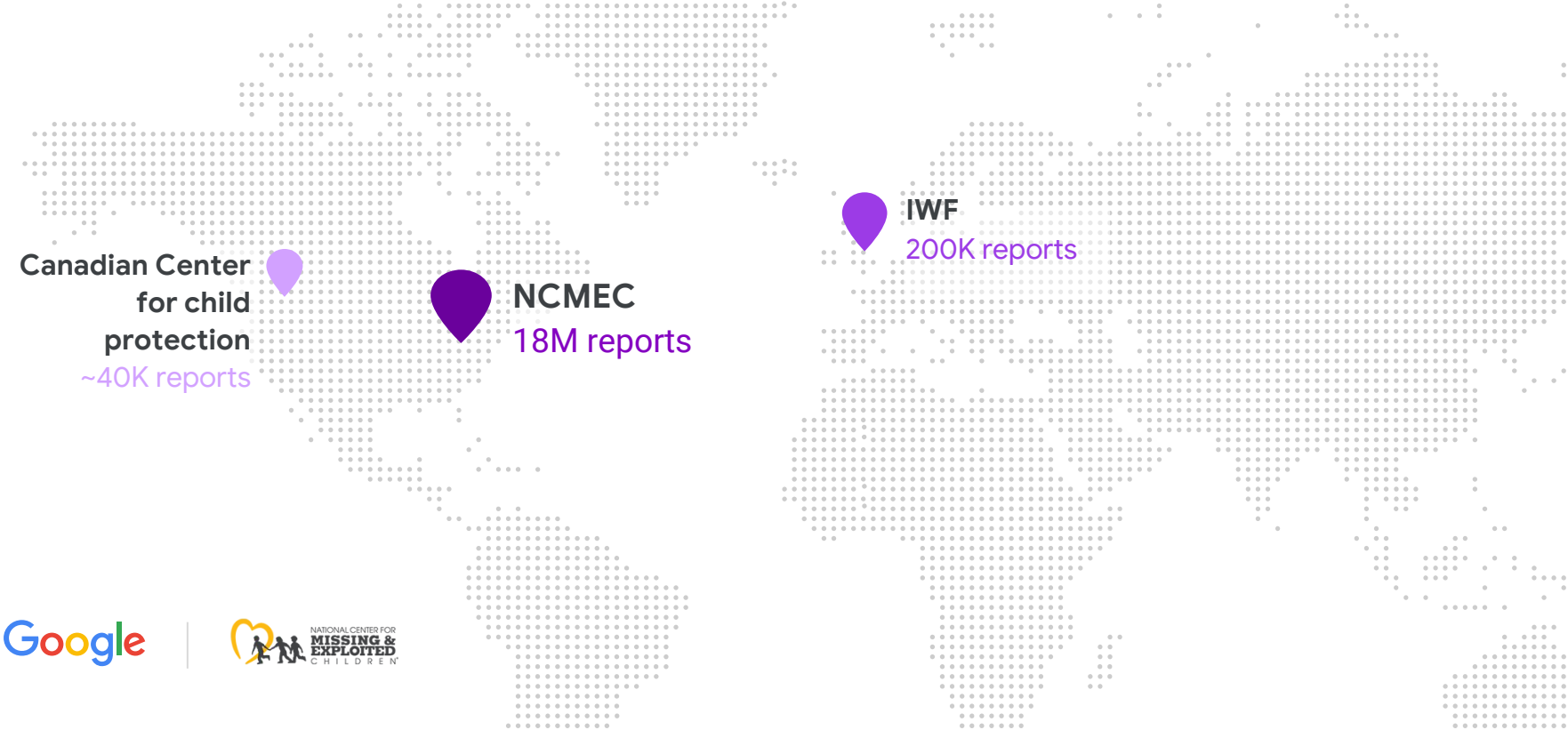
Google | NATIONAL CENTER FOR MISSING & EXPLOITED CHILDREN

Life of a CSAM report



Google | NATIONAL CENTER FOR MISSING & EXPLOITED CHILDREN

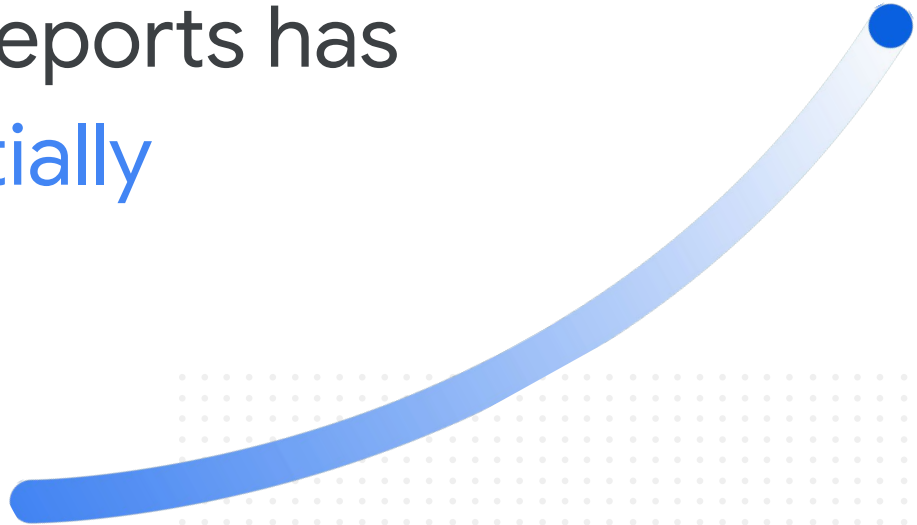
Some of the world's clearinghouses



Google | NATIONAL CENTER FOR MISSING & EXPLOITED CHILDREN

1

The number of reports has grown exponentially



Google | NATIONAL CENTER FOR MISSING & EXPLOITED CHILDREN

1998

3,000
reports

Google | NATIONAL CENTER FOR MISSING & EXPLOITED CHILDREN

1998

3,000
reports

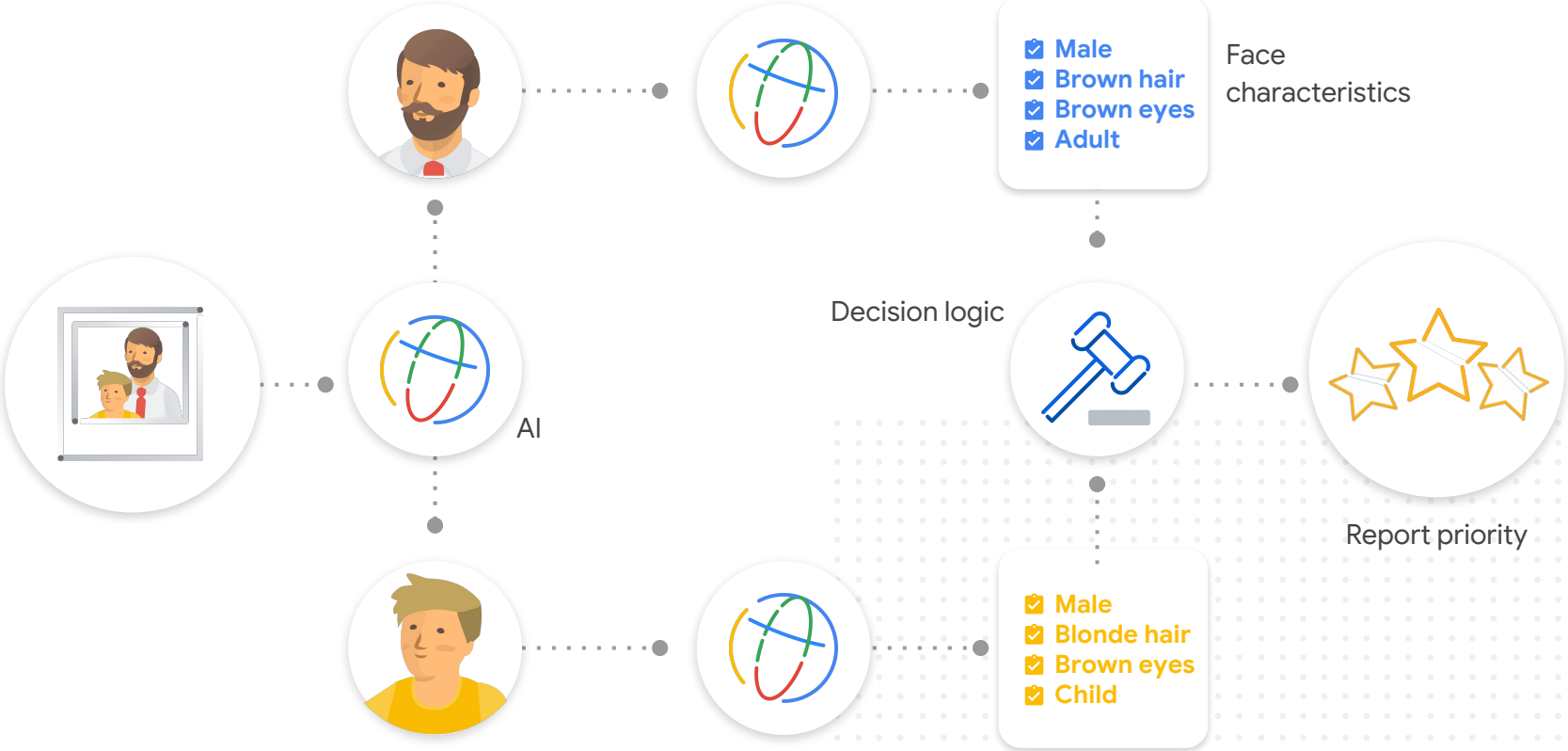
2018

18,400,000 / 6000x
reports

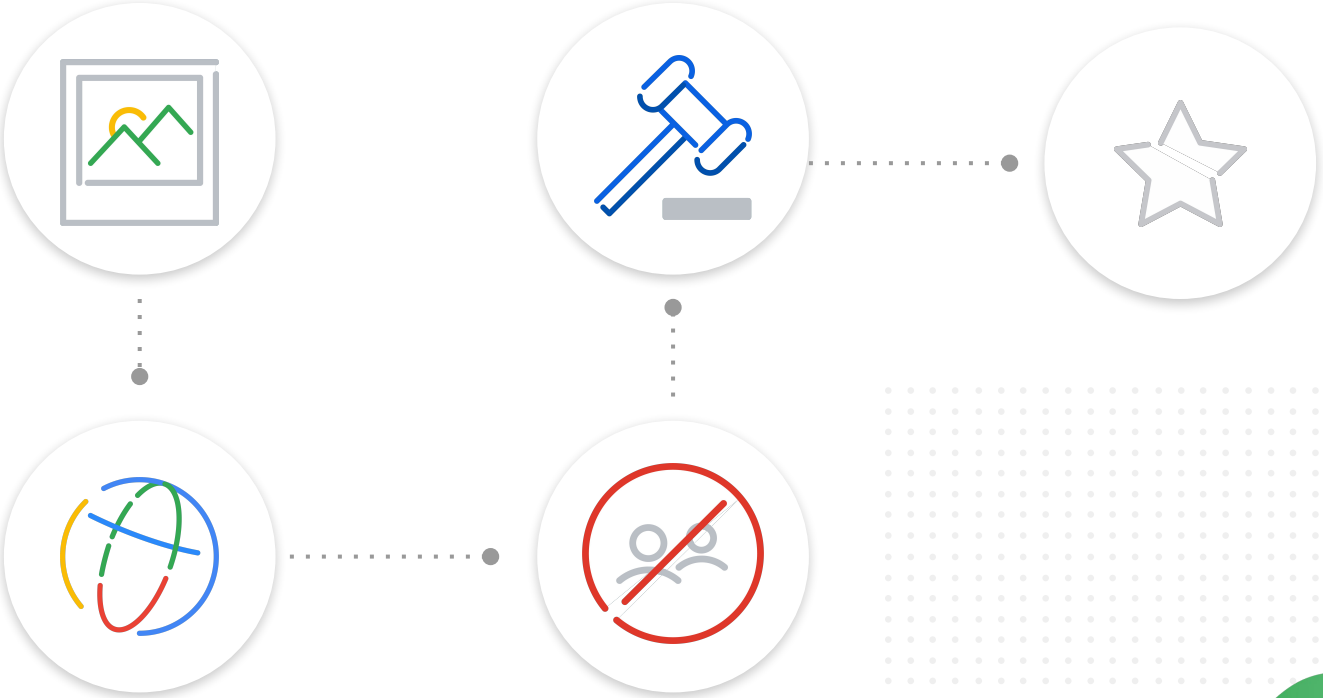
Google | NATIONAL CENTER FOR MISSING & EXPLOITED CHILDREN

Content-based report prioritization to focus on the most actionable ones first

Google | NATIONAL CENTER FOR MISSING & EXPLOITED CHILDREN



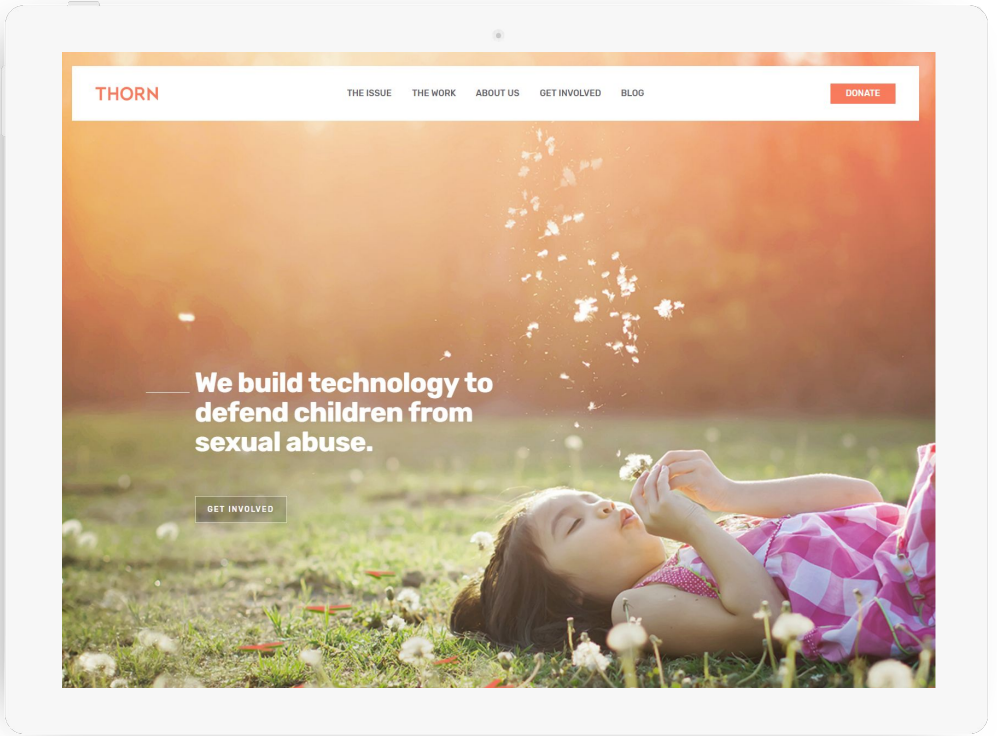
Google | NATIONAL CENTER FOR MISSING & EXPLOITED CHILDREN



Google | NATIONAL CENTER FOR MISSING & EXPLOITED CHILDREN

Open-source

Tensorflow based system available through Thorn



Google | NATIONAL CENTER FOR MISSING & EXPLOITED CHILDREN

An informative image

gender	male	97.3%
age group	young adult	86.3%
hair	black	72.1%
head cover	cap	62.1%

gender	female	78.3%
age group	young adult	63.3%
hair	black	87.1%
head cover	none	97.7%

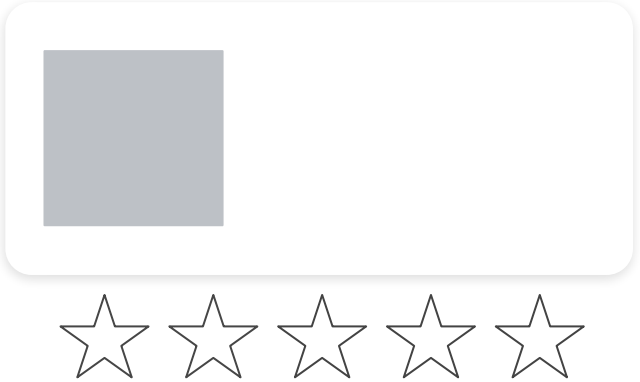
Google | NATIONAL CENTER FOR MISSING & EXPLOITED CHILDREN

A less informative image

gender	male	100%
age group	young adult	92.3%
hair	unknown	74.4%
head cover	cap	97.7%
facial hair	beard	60.7%
eyewear	sunglasses	78.4%

Google | NATIONAL CENTER FOR MISSING & EXPLOITED CHILDREN

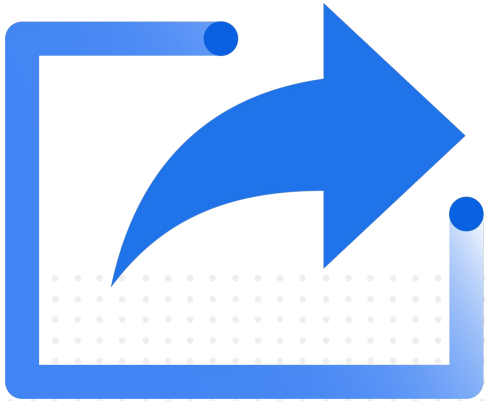
A non informative image



Google | NATIONAL CENTER FOR MISSING & EXPLOITED CHILDREN

2

Most content is only shared once



Google | NATIONAL CENTER FOR MISSING & EXPLOITED CHILDREN

Key media types

84%

Images reported once

91%

Videos reported once

Google | NATIONAL CENTER FOR MISSING & EXPLOITED CHILDREN

content very long tail makes traditional hash blacklists ineffective

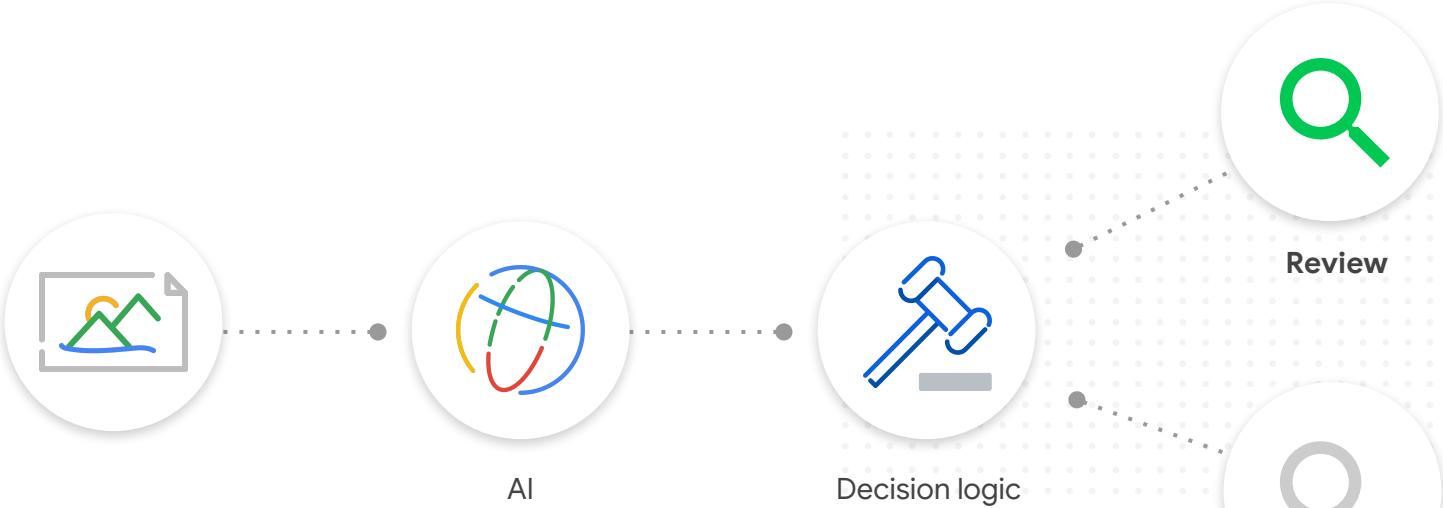
Google | NATIONAL CENTER FOR MISSING & EXPLOITED CHILDREN

Content aware detection

to ensure new content is quickly identified and reported

Google | NATIONAL CENTER FOR MISSING & EXPLOITED CHILDREN

Prioritization API

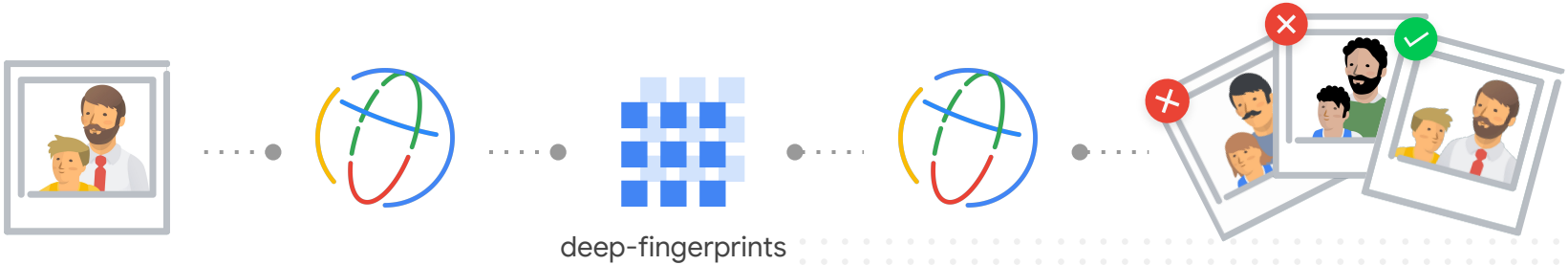


Google | NATIONAL CENTER FOR MISSING & EXPLOITED CHILDREN

https://www.blog.google/technology/safety-security/continuing-fight-against-child-sexual-abuse-online/

Highly experimental

Toward applying AI to related images detection



Google | NATIONAL CENTER FOR MISSING & EXPLOITED CHILDREN

3

Media types are rapidly diversifying

Google | NATIONAL CENTER FOR MISSING & EXPLOITED CHILDREN

Video is the fastest growing medium

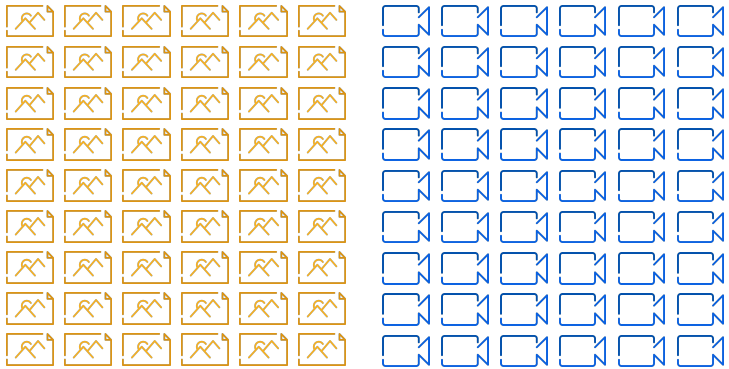
2016

Google | NATIONAL CENTER FOR MISSING & EXPLOITED CHILDREN

Video is the fastest growing medium

2016

2018

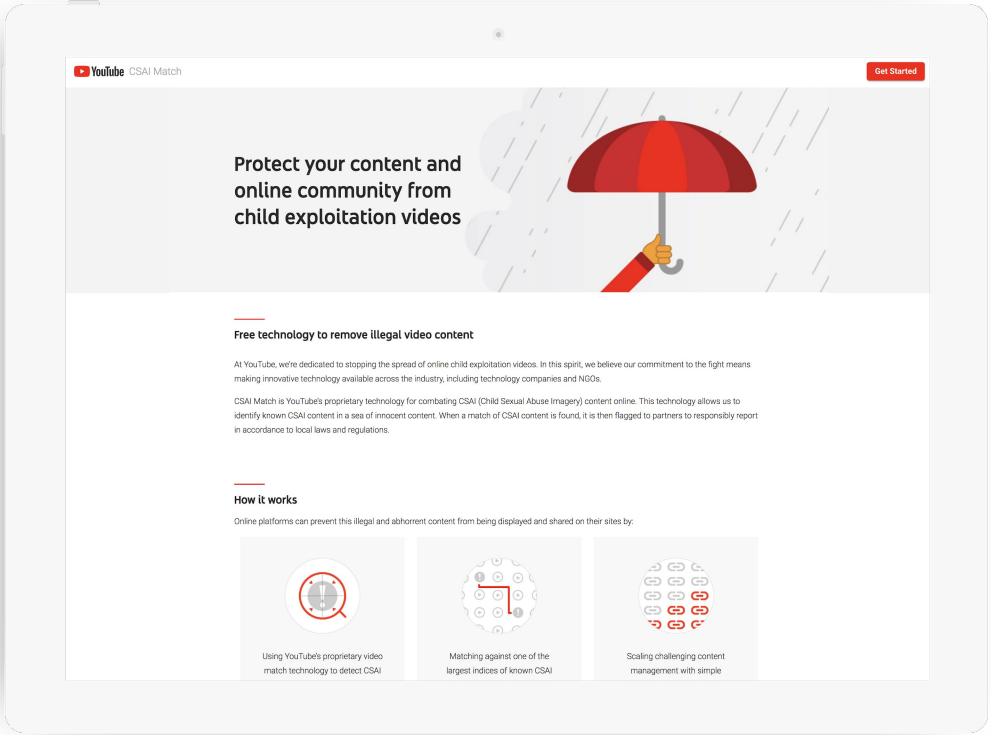


Google | NATIONAL CENTER FOR MISSING & EXPLOITED CHILDREN

Reliable detection across all media to keep pace with medium diversification

Google | NATIONAL CENTER FOR MISSING & EXPLOITED CHILDREN

YouTube video fingerprint



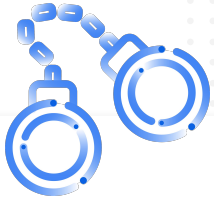
Google | NATIONAL CENTER FOR MISSING & EXPLOITED CHILDREN

Child abuse cross borders

Google | NATIONAL CENTER FOR MISSING & EXPLOITED CHILDREN

PlayPen dark web forum takedown

150,000
Forum users from around the world

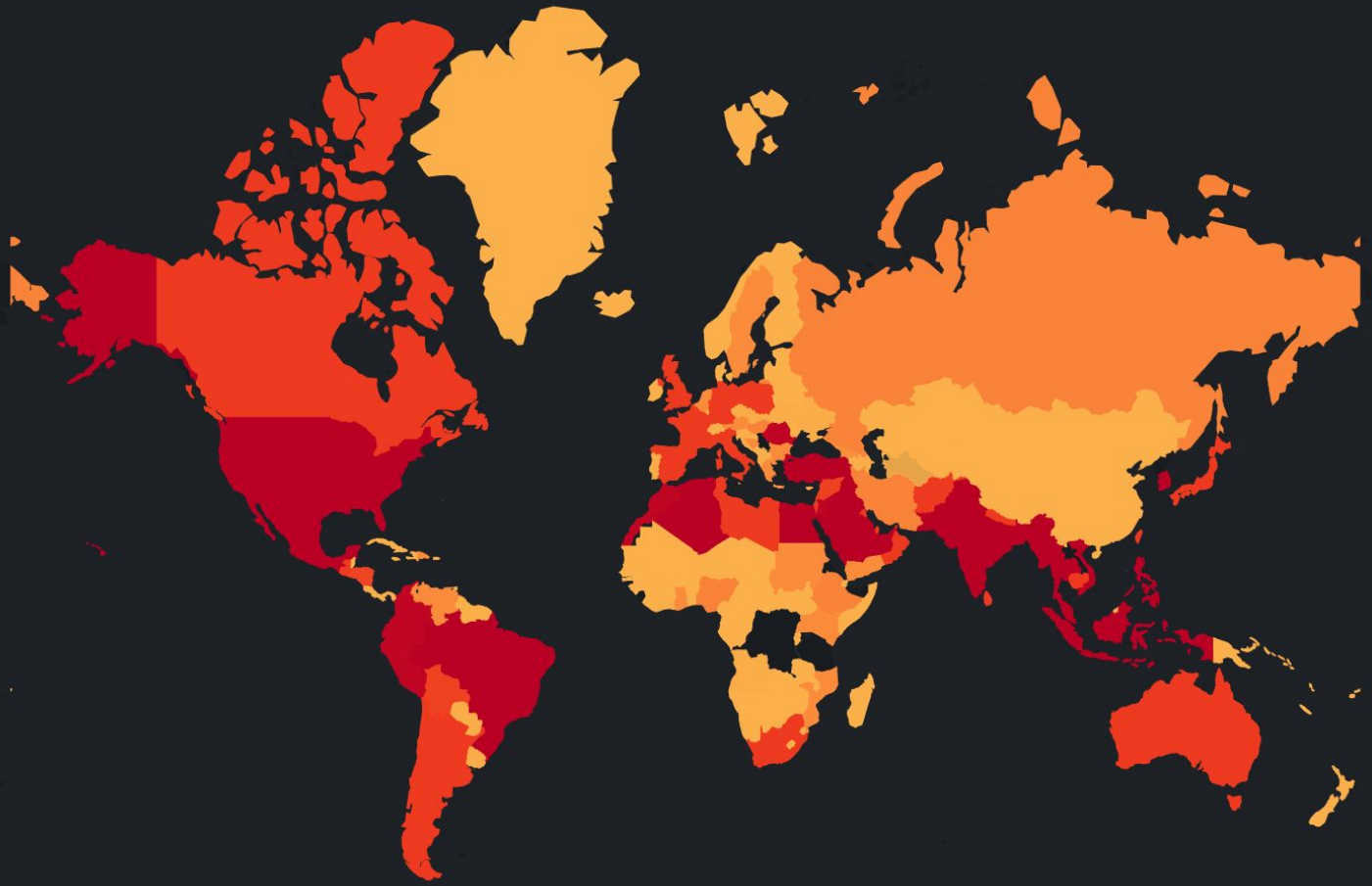


548
International arrests

296
Abused children rescued

Google | NATIONAL CENTER FOR MISSING & EXPLOITED CHILDREN

https://www.fbi.gov/news/stories/playpen-creator-sentenced-to-30-years



Streamline collaboration

to increase efficiency and improve transparency

Google | NATIONAL CENTER FOR MISSING & EXPLOITED CHILDREN

The emotional toll of manual review

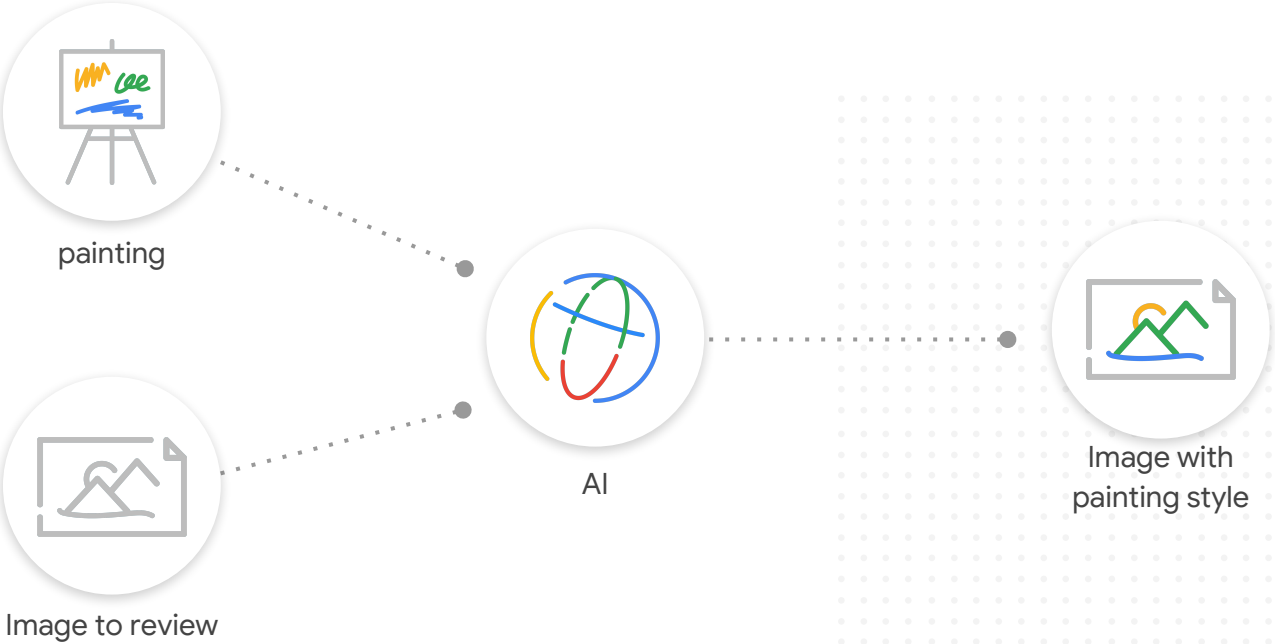
Manual reviews are needed to check the validity of the reports and find potential clues that can help rescue childrens

Google | NATIONAL CENTER FOR MISSING & EXPLOITED CHILDREN

Reduce negative emotional impact via AI based image transformations

Google | NATIONAL CENTER FOR MISSING & EXPLOITED CHILDREN

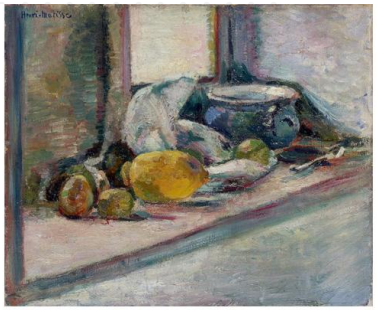
Style transfer explained



Google | NATIONAL CENTER FOR MISSING & EXPLOITED CHILDREN

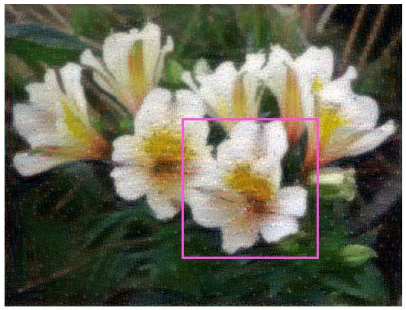
original image

+



style reference

=

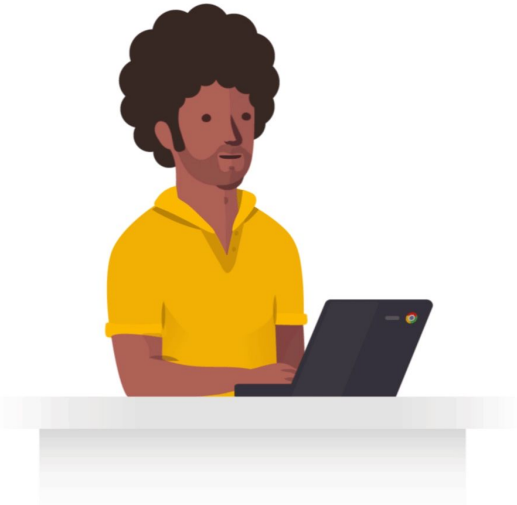


stylized image

Artistic style transfer preserve image semantic while changing its appearance

Google | NATIONAL CENTER FOR MISSING & EXPLOITED CHILDREN

Highly experimental



Reduce
negative emotion by 16%
on explicit and violent
images

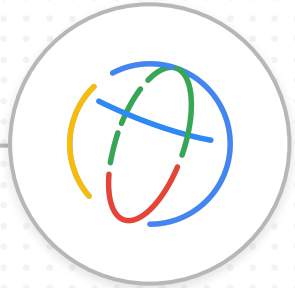
Google | NATIONAL CENTER FOR MISSING & EXPLOITED CHILDREN

Initial work under submission

Takeaway

Technological shifts weaponization

Explosion of child sexual abuse content



Build transformative technologies to curb online child abuse

Google | NATIONAL CENTER FOR MISSING & EXPLOITED CHILDREN

Questions?

Google | NATIONAL CENTER FOR MISSING & EXPLOITED CHILDREN

Let’s build together the technologies needed to make the Internet free of child abuse

Google | NATIONAL CENTER FOR MISSING & EXPLOITED CHILDREN

contact: @elie / elieb@google.com