What are Hashes? What is Photo DNA?

And their application to child sexual abuse material

FACTS

Every image has a unique 'fingerprint'. Using clever mathematics - through the technology of PhotoDNA - each of these 'fingerprints' can be expressed as a unique numerical code which is commonly referred to as a 'hash'.

PhotoDNA is a technology that was first developed by Microsoft. Google is working on a tool that works similarly for videos, FMTS has already created such software.

PhotoDNA is used to identify copies of known images without a human having to look at the images again.

The technology is used by law enforcement and organisations/ companies such as Google, Twitter and Facebook.

Note
The technology cannot be used to identify a person or object in an image, nor can it be used to reverse engineer and reconstitute an image.

How does it work?

① *
Photo DNA starts with an image that has been identified as child sexual abuse material by trusted sources, such as the National Center of Missing and Exploited Children and law enforcement.

②
584030542412...

PhotoDNA transforms (or 'hashes') the image into a black-and-white format and uniform size. It then divides the image into squares and assigns a numerical value that represents the unique shading found within each square. Together these numerical values compose the hash for that image.

③
584030542412...

Hash values of known images can be compared to other images to identify copies. This process is called the matching process and can be used to: 1) identify and flag harmful content online and 2) filter out known materials from collections of images.

④
The hash represents a unique digital identifier or signature for each image. Even if the image has been altered - e.g. when the image is re-sized or when colors are altered - the hash code remains the same for that image.

* Information retrieved from Microsoft

MATCH

PhotoDNA applied to child sexual abuse material

Law Enforcement

Project Vic is an image (and video) hash-sharing initiative that is used by law enforcement and supported by the International Center for Missing and Exploited Children (ICMEC). Using a database of millions of digital hashes of known child sexual abuse material, Project Vic helps law enforcement distinguish already known images from unknown child sexual abuse material. This prevents copies of known images from having to be investigated again and enables detectives to focus on those images that are new and might involve children who still need to be identified. Thus Project Vic helps streamline investigative workflows. This is important given the increasing amount of data retrieved from offenders.

Internet Service Providers

These sets of hashes are also shared with Internet Service Providers and Social Network Websites. The hash technology helps them detect child sexual abuse material shared on their sites. This facilitates the processes of identification, removal or blocking and reporting of child sexual abuse material.

ECPAT