archive.today
webpage capture

Saved from https://www.wsj.com/amp/articles/instagram-vast-pedophile-network-4ab718   search   7 Jun 2023 11:07:33 UTC

history   ←prior next→

Redirected from https://www.wsj.com/articles/instagram-vast-pedophile-network-4ab7189
no other snapshots from this url

All snapshots from host www.wsj.com

| Webpage | Screenshot |

share   download .zip   report bug or abuse   Buy me a coffee

0%

THE WALL STREET JOURNAL.

LIONEL BONAVENTURE/AFP/GETTY IMAGES

Instagram Connects Vast Pedophile Network

The Meta unit's systems for fostering communities have guided users to child-sex content; company says it is improving internal controls

By Jeff Horwitz and Katherine Blunt

June 7, 2023 7:05 am ET

Instagram, the popular social-media site owned by Meta Platforms, helps connect and promote a vast network of accounts openly devoted to the commission and purchase of underage-sex content, according to investigations by The Wall Street Journal and researchers at Stanford University and the University of Massachusetts Amherst.

Pedophiles have long used the internet, but unlike the forums and file-transfer services that cater to people who have interest in illicit content, Instagram doesn't merely host these activities. Its algorithms promote them. Instagram connects pedophiles and guides them to content sellers via recommendation systems that excel at linking those who share niche interests, the Journal and the academic researchers found.

Though out of sight for most on the platform, the sexualized accounts on Instagram are brazen about their interest. The researchers found that Instagram enabled people to search explicit hashtags such as #pedowhore and #preteensex and connected them to accounts that used the terms to advertise child-sex material for sale. Such accounts often claim to be run by

the children themselves and use overtly sexual handles incorporating words such as "little slut for you."

Instagram accounts offering to sell illicit sex material generally don't publish it openly, instead posting "menus" of content. Certain accounts invite buyers to commission specific acts. Some menus include prices for videos of children harming themselves and "imagery of the minor performing sexual acts with animals," researchers at the Stanford Internet Observatory found. At the right price, children are available for in-person "meet ups."

The promotion of underage-sex content violates rules established by Meta as well as federal law.

In response to questions from the Journal, Meta acknowledged problems within its enforcement operations and said it has set up an internal task force to address the issues raised. "Child exploitation is a horrific crime," the company said, adding, "We're continuously investigating ways to actively defend against this behavior."

Meta said it has in the past two years taken down 27 pedophile networks and is planning more removals. Since receiving the Journal queries, the platform said it has blocked thousands of hashtags that sexualize children, some with millions of posts, and restricted its systems from recommending users search for terms known to be associated with sex abuse. It said it is also working on preventing its systems from recommending that potentially pedophilic adults connect with one another or interact with one another's content.

Alex Stamos, the head of the Stanford Internet Observatory and Meta's chief security officer until 2018, said that getting even obvious abuse under control would likely take a sustained effort.

"That a team of three academics with limited access could find such a huge network should set off alarms at Meta," he said, noting that the company has far more effective tools to map its pedophile network than outsiders do. "I hope the company reinvests in human investigators," he added.

Alex Stamos, director of the Stanford Internet Observatory in Palo Alto, Calif.
PHOTO: HELYNN OSPINA FOR THE WALL STREET JOURNAL

Technical and legal hurdles make determining the full scale of the network hard for anyone outside Meta to measure precisely.

Because the laws around child-sex content are extremely broad, investigating even the open promotion of it on a public platform is legally sensitive.

In its reporting, the Journal consulted with academic experts on online child safety. Stanford's Internet Observatory, a division of the university's Cyber Policy Center focused on social-media abuse, produced an independent quantitative analysis of the Instagram features that help users connect and find content.

The Journal also approached UMass's Rescue Lab, which evaluated how pedophiles on Instagram fit into the larger ecosystem of online child exploitation. Using different methods, both entities were able to quickly identify large-scale communities promoting criminal sex abuse.

Test accounts set up by researchers that viewed a single account in the network were immediately hit with "suggested for you" recommendations of purported child-sex-content sellers and buyers, as well as accounts linking to off-platform content trading sites. Following just a handful of these recommendations was enough to flood a test account with content that sexualizes children.

The Stanford Internet Observatory used hashtags associated with underage sex to find 405 sellers of what researchers labeled "self-generated" child-sex material—or accounts purportedly run by children themselves, some saying they were as young as 12. According to data gathered via Maltego, a network mapping software, 112 of those seller accounts collectively had 22,000 unique followers.

Underage-sex-content creators and buyers are just a corner of a larger ecosystem devoted to sexualized child content. Other accounts in the pedophile community on Instagram aggregate pro-pedophilia memes, or discuss their access to children. Current and former Meta employees who have worked on Instagram child-safety initiatives estimate the number of accounts that exist primarily to follow such content is in the high hundreds of thousands, if not millions.

A Meta spokesman said the company actively seeks to remove such users, taking down 490,000 accounts for violating its child safety policies in January alone.

"Instagram is an on ramp to places on the internet where there's more explicit child sexual abuse," said Brian Levine, director of the UMass Rescue Lab, which researches online child victimization and builds forensic tools to combat it. Levine is an author of a 2022 report for the National Institute of Justice, the Justice Department's research arm, on internet child exploitation.

Instagram, estimated to have more than 1.3 billion users, is especially popular with teens. The Stanford researchers found some similar sexually exploitative activity on other, smaller social platforms, but said they found that the problem on Instagram is particularly severe. "The most important platform for these networks of buyers and sellers seems to be Instagram," they wrote in a report slated for release on June 7.

The Palo Alto campus of Meta Platforms.
PHOTO: HELYNN OSPINA FOR THE WALL STREET JOURNAL

Instagram said that its internal statistics show that users see child exploitation in less than one in 10 thousand posts viewed.

The effort by social-media platforms and law enforcement to fight the spread of child pornography online centers largely on hunting for confirmed images and videos, known as child sexual abuse material, or CSAM, which already are known to be in circulation. The National Center for Missing & Exploited Children, a U.S. nonprofit organization that works with law enforcement, maintains a database of digital fingerprints for such images and videos and a platform for sharing such data among internet companies.

Internet company algorithms check the digital fingerprints of images posted on their platforms against that list, and report back to the center when they detect them, as U.S. federal law requires. In 2022, the center received 31.9 million reports of child pornography, mostly from internet companies—up 47% from two years earlier.

Meta, with more than 3 billion users across its apps, which include Instagram, Facebook and WhatsApp, is able to detect these types of known images if they aren't encrypted. Meta accounted for 85% of the child pornography reports filed to the center, including some 5 million from Instagram.

Meta's automated screening for existing child exploitation content can't detect new images or efforts to advertise their sale. Preventing and detecting such activity requires not just reviewing user reports but tracking and disrupting pedophile networks, say current and former staffers as well as the Stanford researchers. The goal is to make it difficult for such users to connect with each other, find content and recruit victims.

Such work is vital because law-enforcement agencies lack the resources to investigate more than a tiny fraction of the tips NCMEC receives, said Levine of UMass. That means the platforms have primary responsibility to prevent a community from forming and normalizing child sexual abuse.

Meta has struggled with these efforts more than other platforms both because of weak enforcement and design features that promote content discovery of legal as well as illicit material, Stanford found.

The Stanford team found 128 accounts offering to sell child-sex-abuse material on Twitter, less than a third the number they found on Instagram, which has a far larger overall user base than Twitter. Twitter didn't recommend such accounts to the same degree as Instagram, and it took them down far more quickly, the team found.

Among other platforms popular with young people, Snapchat is used mainly for its direct messaging, so it doesn't help create networks. And TikTok's platform is one where "this type of content does not appear to proliferate," the Stanford report said.

Twitter didn't respond to requests for comment. TikTok and Snapchat declined to comment.

David Thiel, chief technologist at the Stanford Internet Observatory, said, "Instagram's problem comes down to content-discovery features, the ways topics are recommended and how much the platform relies on search and links between accounts." Thiel, who previously worked at Meta on security and safety issues, added, "You have to put guardrails in place for something that growth-intensive to still be nominally safe, and Instagram hasn't."

The platform has struggled to oversee a basic technology: keywords. Hashtags are a central part of content discovery on Instagram, allowing users to tag and find posts of interest to a particular community—from broad topics such as #fashion or #nba to narrower ones such as #embroidery or #spelunking.
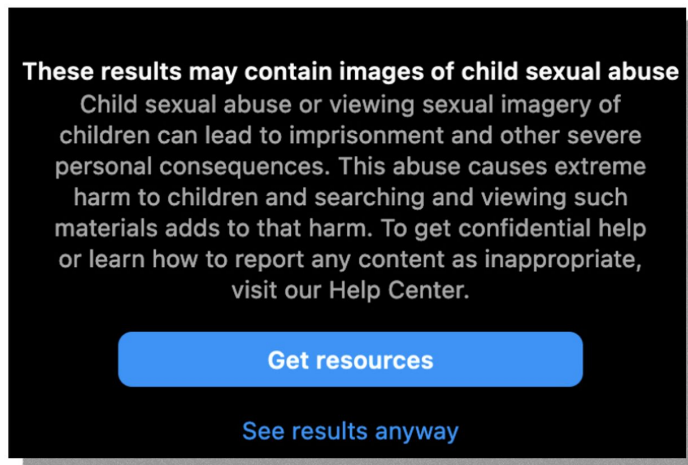
A screenshot taken by the Stanford Internet Observatory shows the warning and clickthrough option when searching for a pedophilia-related hashtag on Instagram.
PHOTO: STANFORD INTERNET OBSERVATORY

Pedophiles have their chosen hashtags, too. Search terms such as #pedobait and variations on #mnsfw ("minor not safe for work") had been used to tag thousands of posts dedicated to advertising sex content featuring children, rendering them easily findable by buyers, the academic researchers found. Following queries from the Journal, Meta said it is in the process of banning such terms.

In many cases, Instagram has permitted users to search for terms that its own algorithms know may be associated with illegal material. In such cases, a pop-up screen for users warned that "These results may contain images of child sexual abuse," and noted that production and consumption of such material causes "extreme harm" to children. The screen offered two options for users: "Get resources" and "See results anyway."

In response to questions from the Journal, Instagram removed the option for users to view search results for terms likely to produce illegal images. The company declined to say why it had offered the option.

The pedophilic accounts on Instagram mix brazenness with superficial efforts to veil their activity, researchers found. Certain emojis function as a kind of code, such as an image of a map—shorthand for "minor-attracted person"—or one of "cheese pizza," which shares its initials with "child

pornography," according to Levine of UMass. Many declare themselves "lovers of the little things in life."

Accounts identify themselves as "seller" or "s3ller," and many state their preferred form of payment in their bios. These seller accounts often convey the child's purported age by saying they are "on chapter 14," or "age 31" followed by an emoji of a reverse arrow.

Some of the accounts bore indications of sex trafficking, said Levine of UMass, such as one displaying a teenager with the word WHORE scrawled across her face.

Some users claiming to sell self-produced sex content say they are "faceless"—offering images only from the neck down—because of past experiences in which customers have stalked or blackmailed them. Others take the risk, charging a premium for images and videos that could reveal their identity by showing their face.

Many of the accounts show users with cutting scars on the inside of their arms or thighs, and a number of them cite past sexual abuse.

Even glancing contact with an account in Instagram's pedophile community can trigger the platform to begin recommending that users join it.

Sarah Adams, a Canadian mother of two, has built an Instagram audience combatting child exploitation.
PHOTO: ALANA PATERSON FOR THE WALL STREET JOURNAL

Sarah Adams, a Canadian mother of two, has built an Instagram audience discussing child exploitation and the dangers of oversharing on social media. Given her focus, Adams' followers sometimes send her disturbing things they've encountered on the platform. In February, she said, one messaged her with an account branded with the term "incest toddlers."

Adams said she accessed the account—a collection of pro-incest memes with more than 10,000 followers—for only the few seconds that it took to report to Instagram, then tried to forget about it. But over the course of the next few days, she began hearing from horrified parents. When they looked at Adams' Instagram profile, she said they were being recommended "incest toddlers" as a result of Adams' contact with the account.

A Meta spokesman said that "incest toddlers" violated its rules and that Instagram had erred on enforcement. The company said it plans to address such inappropriate recommendations as part of its newly formed child safety task force.

As with most social-media platforms, the core of Instagram's recommendations are based on behavioral patterns, not by matching a user's interests to specific subjects. This approach is efficient in increasing the relevance of recommendations, and it works most reliably for communities that share a narrow set of interests.

In theory, this same tightness of the pedophile community on Instagram should make it easier for Instagram to map out the network and take steps to combat it. Documents previously reviewed by the Journal show that Meta has done this sort of work in the past to suppress account networks it deems harmful, such as with accounts promoting election delegitimization in the U.S. after the Jan. 6 Capitol riot.

Like other platforms, Instagram says it enlists its users to help detect accounts that are breaking rules. But those efforts haven't always been effective.

Sometimes user reports of nudity involving a child went unanswered for months, according to a review of scores of reports filed over the last year by numerous child-safety advocates.

Earlier this year, an anti-pedophile activist discovered an Instagram account claiming to belong to a girl selling underage-sex content, including a post declaring, "This teen is ready for you pervs." When the activist reported the account, Instagram responded with an automated message saying: "Because of the high volume of reports we receive, our team hasn't been able to review this post."

After the same activist reported another post, this one of a scantily clad young girl with a graphically sexual caption, Instagram responded, "Our review team has found that [the account's] post does not go against our Community Guidelines." The response suggested that the user hide the account to avoid seeing its content.

A Meta spokesman acknowledged that Meta had received the reports and failed to act on them. A review of how the company handled reports of child sex abuse found that a software glitch was preventing a substantial portion of user reports from being processed, and that the company's moderation staff wasn't properly enforcing the platform's rules, the spokesman said. The company said it has since fixed the bug in its reporting system and is providing new training to its content moderators.

Even when Instagram does take down accounts selling underage-sex content, they don't always stay gone.

Under the platform's internal guidelines, penalties for violating its community standards are generally levied on accounts, not users or devices. Because Instagram allows users to run multiple linked accounts, the system makes it easy to evade meaningful enforcement. Users regularly list the handles of "backup" accounts in their bios, allowing them to simply resume posting to the same set of followers if Instagram removes them.

Sarah Adams' followers

In some instances, Instagram's recommendations systems directly undercut efforts by its own safety staff. After the company decided to crack down on links from a specific encrypted file transfer service notorious for transmitting child sex content, Instagram blocked searches for its name.

Instagram's AI-driven hashtag suggestions didn't get the message. Despite refusing to show results for the service's name, the platform's autofill feature recommended that users try variations on the name with the words "boys" and "CP" added to the end.

The company tried to disable those hashtags amid its response to the queries by the Journal. But within a few days Instagram was again recommending new variations of the service's name that also led to accounts selling purported underage-sex content.

Following the company's initial sweep of accounts brought to its attention by Stanford and the Journal, UMass's Levine checked in on some of the remaining underage seller accounts on Instagram. As before, viewing even one of them led Instagram to recommend new ones. Instagram's suggestions were helping to rebuild the network that the platform's own safety staff was in the middle of trying to dismantle.

A Meta spokesman said its systems to prevent such recommendations are currently being built. Levine called Instagram's role in promoting pedophilic content and accounts unacceptable.

"Pull the emergency brake," he said. "Are the economic benefits worth the harms to these children?"

Write to Jeff Horwitz at jeff.horwitz@wsj.com and Katherine Blunt at katherine.blunt@wsj.com