

FILED UNDER SEAL

EXHIBIT 31

**FIRST JUDICIAL DISTRICT COURT
COUNTY OF SANTA FE
STATE OF NEW MEXICO**

STATE OF NEW MEXICO, EX REL.
RAÚL TORREZ, ATTORNEY GENERAL,

Plaintiff,

v.

NO. D-101-CV-2023-02838

META PLATFORMS, INC.;
INSTAGRAM, LLC;
META PAYMENTS, INC.; and
META PLATFORMS TECHNOLOGIES, LLC,

Defendants.

**DEFENDANTS' FIRST SUPPLEMENTAL AND AMENDED OBJECTIONS AND
RESPONSES TO PLAINTIFF'S FIRST SET OF INTERROGATORIES TO
DEFENDANTS**

Defendants Meta Platforms, Inc., Instagram, LLC, Meta Payments, Inc., and Meta Platforms Technologies, LLC (collectively "Meta"), by and through their attorneys, Holland & Hart LLP, and Covington & Burling LLP, hereby provide Supplemental Objections and Responses to Plaintiff's First Set of Interrogatories as follows. Meta's responses to the Interrogatories are made to the best of its current knowledge, information, and belief. Meta reserves the right to supplement or amend any of its responses and objections as appropriate as discovery and its investigation proceed.

RECURRING SPECIFIC OBJECTIONS

1. Meta objects to the Interrogatories to the extent they conflict with the Rules of Civil Procedure for the District Courts of New Mexico or the local rules of this Court ("Discovery Rules"). See Rule 1-026(B) NMRA (defining the scope of permissible

HIGHLY CONFIDENTIAL

this Interrogatory on the ground that it is vague and ambiguous in its use of the undefined phrase “absolute or total number of offending or violative content.” See Rule 1-026(B)(1) (requirement to provide “responsive information” implies a clearly stated request). Meta further objects to this Interrogatory as seeking discovery that is not relevant to the claims and defenses in this action; not limited to New Mexico; and not limited to a particular time period. Based on a reasonable investigation, Meta’s Community Standards Enforcement Reports (“CSERs”) do not measure an “absolute or total number of offending or violative content.” Meta is continuing to investigate whether and the extent to which CSERs measure “(b) the absolute or total number of “views” of that offending or violative content that [may be] identified in [a] sample;” or “(c) the overall number of “views” that [may be] used to calculate the ‘prevalence’ metric.”

Subject to and without waiving the foregoing objections, Meta states that its investigation remains ongoing as to whether (1) responsive, non-privileged information exists that has been created and is maintained in the ordinary course of business and (2) the burden of obtaining and producing such information is proportional to the needs of the case.

Interrogatory 15. Identify all steps, procedures, and technology You employ to monitor for, identify, remove and report Inappropriate Content, CSAM or CSEC on Meta’s Platforms, and, for each step, procedure, or technology, identify the time frame(s) it was in use.

AMENDED RESPONSE: Meta incorporates its Responses and Objections Applicable to All Requests by reference as though fully set forth herein. Meta further objects to this Interrogatory on the ground that it is vague and ambiguous in its use of the undefined terms “steps, procedures and technology.” See Rule 1-026(B)(1) (requirement to provide “responsive information” implies a clearly stated request). Meta construes “steps, procedures and technology” to mean its content moderation processes and technologies. See Rule 1-033(C)(1).

HIGHLY CONFIDENTIAL

Meta further objects to this Interrogatory on the ground that it is vague and ambiguous in its use of the undefined phrase “monitor for, identify, remove and report.” See Rule 1-026(B)(1) (requirement to provide “responsive information” implies a clearly stated request). Meta construes “monitor for, identify, remove and report” to refer to its content moderation efforts. See Rule 1-033(C)(1). Additionally, Meta interprets the phrase “monitor for” to refer to Meta’s efforts to detect the categories of harm listed in Interrogatory 15, rather than Meta’s efforts to prevent such harms. Meta interprets “Inappropriate Content” to refer to content that violates Meta’s policies prohibiting child sexual exploitation.

Meta further objects to this Interrogatory to the extent it seeks “the time frame(s)” that “each step, procedure or technology” “was in use,” to the extent that it asks for a specific time frame during which any process or technology was in use at Meta. See Rule 1-026(B)(1) (requirement to provide information “then known to the party” implies information should be reasonably ascertainable). Each content moderation process and each classifier was developed through an iterative process and through a collaborative effort between Meta’s Product, Operations, Content Policy, Product Policy, and Safety Policy teams, such that any particular process or classifier could have multiple historic iterations.

Meta further objects to this Interrogatory as overbroad, unduly burdensome, and not proportional to the needs of the case because it seeks “all” steps, procedures, and technology. . See Rule 1-026(B)(1), (2); *see also Lara*, 1999-NMCA-012, ¶ 10 (discovery must be narrowly tailored to avoid irrelevant information); *Ruiz*, 1981-NMCA-094, ¶ 3 (interrogatories that are “exceedingly broad” are objectionable); *Archuleta*, 2005-NMSC-006, ¶ 23 (“a party will not be required to respond to an overly broad discovery request unless adequate guidance exists as to what extent the request is not objectionable.”) (internal quotation marks and citation omitted).

Meta further objects to this Interrogatory on the grounds that it is unduly burdensome and overly broad insofar as it contains numerous subparts, and on that basis constitutes more than a single interrogatory. *See* Rule LR1-202(E) (“Subparts of an interrogatory shall relate directly to the subject matter of the interrogatory”).

Subject to and without waiving the foregoing objections, Meta provides the following response to Interrogatory No. 15:

Meta’s industry-leading efforts to combat child exploitation focus on preventing abuse, detecting and reporting content that violates Meta’s policies, and working with experts and authorities to help keep children safe. These efforts and the company’s focus on safety and integrity have been and will continue to be a priority. Meta has invested over \$20 billion since 2016 on overall safety and security—including around \$5 billion in 2023 alone—and has approximately 40,000 people dedicated to working on these efforts. This work includes thousands of individuals trained to spot and enforce against potentially violating content, and the development of industry-leading technologies to help Meta proactively identify content, prioritize the most critical content to be reviewed, and act on content that violates Meta’s policies. Meta continues to explore new ways to actively tackle child exploitation, including continuously evaluating its policies and investing in and soliciting feedback from relevant internal and external groups.

Meta has systems in place to remove policy-violating content when it finds it, regardless of the context or the person’s motivation for sharing it. Meta also has developed cutting-edge technology to prevent, find, remove, and report policy-violating content. In addition to this technology, Meta has invested in specially trained teams with backgrounds in law enforcement,

online safety, analytics, and forensic investigations to both find and review potentially violating content, accounts and adversarial networks.

A. Content Policies

Meta's policies define what content is and is not allowed on Meta's platforms.¹⁷ Meta has strict policies prohibiting child nudity, abuse, and exploitation, including Child Sexual Abuse Material ("CSAM"), inappropriate interactions with children, solicitation, and content that sexualizes children. When Meta finds content that violates its policies, the company removes it, regardless of the context or the person's motivation for sharing it. Under Meta's Community Standards, the following content, inter alia, is considered violating: content, activity, or interactions that threaten, depict, praise, support, provide instructions for, make statements of intent, admit participation in, or share links of the sexual exploitation of children; content that solicits sexual content or activity depicting or involving real or non-real children (CSAM); nude imagery of real or non-real children, and/or sexualized imagery of real or non-real children; content that solicits sexual encounters with children; content that constitutes or facilitates inappropriate interactions with children; content that attempts to exploit real children by coercing money, favors, or intimate imagery with threats to expose intimate imagery or information or by sharing, threatening, or stating an intent to share private sexual conversations or intimate imagery; content that sexualizes real or non-real children; content that depicts real or non-real child nudity; videos or photos that depict real or non-real non-sexual child abuse regardless of sharing intent; or content that praises, supports, promotes, advocates for, provides instructions for or encourages participation in non-sexual child abuse.¹⁸

¹⁷ Meta's Community Standards apply to content posted on Facebook and Instagram. See <https://transparency.meta.com/policies>; <https://www.facebook.com/help/instagram/477434105621119>.

¹⁸ <https://transparency.meta.com/policies/community-standards/child-sexual-exploitation-abuse-nudity/>.

B. Content Moderation

Meta uses a combination of machine learning technology, reports from external partners, reports from its users, and human review to help identify and remove content that violates Meta’s policies related to child exploitation and abuse. Content moderation—the processes by which Meta detects and enforces against content that violates its policies, including but not limited to its child safety policies—is a multi-layered process that may rely on a number of tools, including sophisticated machine learning and media matching tools and human review. A piece of potentially violating content can go through multiple steps of review, each of which can result in a range of possible enforcement actions depending on the content at issue. Content can be flagged for evaluation in two ways: by being automatically flagged for action or review by Meta’s machine learning and/or media matching technologies, or by being reported by a user for review. To illustrate the scale of Meta’s proactive review of the child sexual exploitation content Meta removed from its platforms in Q1 2024, the company detected 94% on Facebook and 93% on Instagram before it was reported by users. And in Q2 2024, the company detected more than 97% of such content on Facebook and 96% on Instagram before it was reported.¹⁹

C. Classifiers

Meta uses automated machine learning tools known as “classifiers” that can proactively identify, enqueue for review, and, in some instances, remove potentially violating content on Meta’s Services. In circumstances where these tools assess that a piece of content violates Meta’s policies, the content may be removed without further review before any user sees the content. Meta also proactively detects *known* child exploitation and sexualizing content, leveraging the type of technology available across industry for CSAM hash matching, including

¹⁹ <https://transparency.meta.com/reports/community-standards-enforcement/child-nudity-and-sexual-exploitation/facebook/>

methods that Meta developed and open-sourced. For example, Meta uses photo- and video-matching technologies that detect identical or near-identical photos and videos of known child exploitative content, and Meta uses Google’s Content Safety API to help better prioritize content that may contain child exploitation for Meta’s content reviewers to assess.

A “classifier” refers to a machine learning tool or algorithm that automatically reviews and categorizes content on Facebook and Instagram for further action, including but not limited to content removal, downranking, demoting, filtering or non-action, among others. Classifiers evaluate the content against Meta’s Policies and can make determinations about whether a particular piece of content is considered “policy-violating.” Classifiers run automatically and review content on Facebook and Instagram. Different classifiers review and detect different types of policy-violating content, also known as “harm areas.” Those harm areas include, but are not limited to, child safety.

D. Human Review

Meta trains and relies on thousands of human reviewers around the world to enforce its Community Standards and review content that may potentially violate Meta’s policies. For all policies and protocols, Meta provides comprehensive and rigorous training for content reviewers.

Meta also leverages automated detection to optimize the efficiency of human review by automatically scanning and prioritizing for review content that is more likely to be potentially harmful and/or more likely to spread across its platforms. Meta’s review systems use technology to prioritize high-severity content with the potential for offline harm and viral content which is spreading quickly.

E. Text-Based Interventions

Meta employs keyword or text-based interventions to identify and prevent users from searching for problematic content. For example, when a user's search matches one of these keywords or phrases, they encounter an Empty Search Engine Results Page (SERP) and receive a warning message. For searches that match one of these keywords or phrases, typeahead recommendations or related searches are disabled. Meta maintains a unified keyword list (currently tens of thousands of terms) across Services and interventions to aid in detecting potentially violating content. This unified list is used to proactively detect violating words and phrases across its Services.

Meta has processes in place to regularly update its unified keyword list to reflect and incorporate new terminology identified through detection and enforcement mechanisms as well as external input. Meta works regularly with professionals to aid in its understanding of evolutions in coded language and to identify new and evolving terms, phrases, slang, and emojis that could be used to evade its detection systems and bypass its policies. Meta uses these signals to train its technology in an effort to proactively identify new trends in behavior and language and take action. Meta's teams also use technology to identify misspellings and spelling variations of inherently violating terms and coded language, which allows Meta to proactively identify and disrupt efforts to evade the protections in place.

F. Media/Image-Based Interventions

Since 2011, Meta has used Microsoft's PhotoDNA, a technology that creates a unique digital signature (known as a "hash") of an image which is then compared against hashes of other photos to find copies of the same or nearly duplicative image. In addition to using photo and video matching technologies to identify known CSAM, Meta uses, among other things, machine learning

tools to proactively detect child nudity and previously unknown child exploitative content when it is uploaded.

Meta employs its Media Match Service (“MMS”) to scan for, proactively detect, and enforce against copies and near-copies of known, policy-violating content, including potential CSAM and other violating images across media uploads on its Services. Every piece of content uploaded to Facebook and Instagram is checked against MMS banks.

G. User Reports

Reporting is an essential tool for users to stay safe and help Meta respond to abuse and violating content effectively. Meta encourages its users to report content that they believe violates Meta’s policies using the dedicated tools Meta has designed for its services. This includes Pages, Groups, profiles, individual posts, and comments on Facebook and Instagram and accounts and chats on Messenger and Instagram DMs.

Facebook and Instagram users can report Community Standards and Community Guidelines violations in a variety of ways, including through an option available on individual pieces of content. If a user sees and reports a piece of content, or if the content is flagged by classifiers (*see* Part C), it may be removed or marked for non-distribution by Meta’s automated systems. If the reviewer confirms that the reported content violates Meta’s policies, it is removed.

Meta works constantly to find, remove, and report policy-violating content, including comments that praise or support the sexual exploitation of children; deactivate accounts of predators; and dismantle abusive networks. The company goes beyond legal requirements and uses technology to proactively seek out abusive material, and as a result, finds and reports more inappropriate content than anyone else in the industry. As one example, in Q1 2024, Meta took action against 14.4 million pieces of child sexual exploitation content on Facebook and 2.7 million

on Instagram—over 93% before it was reported to Meta. From August 1, 2023 to December 31, 2023, Meta disabled more than 2.6 million accounts for violating its child safety policies. Meta’s teams disrupted 37 abusive networks between 2020 and 2023 and removed nearly 200,000 accounts associated with those networks.

H. External Reporting

Since 2009, as required by federal law, Meta has reported apparent instances of child exploitation identified on Meta’s platforms anywhere in the world to the National Center for Missing and Exploited Children (“NCMEC”), which coordinates with law enforcement authorities across the globe. Meta responds to valid law enforcement requests for information with data, including email addresses and phone numbers, and traffic data, such as IP addresses, that can be used in criminal investigations. Meta provides operational guidelines to law enforcement who seek records from Facebook or Instagram.

Every year, Meta submits tens of millions of reports to the NCMEC CyberTipline. Meta makes submissions to the CyberTipline by populating fields in an application programming interface (“API”)—a software intermediary that allows Meta and NCMEC’s applications to communicate. The vast majority of CSAM on Meta’s Services reported to NCMEC—in excess of 90%—is detected and reported automatically, using sophisticated technology and tools, including the media-matching technologies described above. By this association of CSAM to hash values, Meta can detect certain content that was previously determined to be apparent CSAM and report it to NCMEC automatically, without human involvement at the time of the reporting.

Because this is an industry-wide concern, Meta also educates people about various tools that help prevent nude or sexual photos from being shared or reshared online. For example, in February 2023, Meta announced that it would be a founding member of “Take It Down,” a free

HIGHLY CONFIDENTIAL

tool operated and run by NCMEC that can help individuals stop the online sharing of nude, partially nude, or sexually explicit images or videos taken when the individual was under 18 years old. Take It Down works by assigning a hash value (as described above), to nude, partially nude, or sexually explicit images or videos of people under the age of 18. Meta and other participating online platforms can use hash values to detect these images or videos on their services, including Instagram and Facebook, and remove this content. This all happens without the image or video ever leaving a user's device.

SUPPLEMENTAL RESPONSE: Subject to and without waiving the foregoing objections, Meta provides the following additional information: Meta has produced information under N.M. R. Civ. P. Dist. Ct. 1-033(E) from which answers broadly responsive to this interrogatory can be derived. Meta incorporates and refers to the documents produced at the Bates ranges listed below. Specifically, at these Bates ranges, Meta has produced documents reflecting iterations of past Facebook Community Standards, Instagram Community Guidelines, and Community Standards Enforcement Reports.

METANMAG-01 Bates Beg	METANMAG-02 Bates End
METANMAG-007-00379176	METANMAG-007-00379177
METANMAG-007-00461847	METANMAG-007-00461861
METANMAG-003-00277147	METANMAG-003-00277149

METANMAG-003-00277150	METANMAG-003-00277152
METANMAG-003-00277153	METANMAG-003-00277155
METANMAG-003-00277156	METANMAG-003-00277158
METANMAG-003-00277159	METANMAG-003-00277161
METANMAG-003-00277162	METANMAG-003-00277163
METANMAG-003-00277164	METANMAG-003-00277165
METANMAG-003-00277166	METANMAG-003-00277168
METANMAG-003-00277169	METANMAG-003-00277170
METANMAG-003-00277171	METANMAG-003-00277172
METANMAG-003-00277173	METANMAG-003-00277174
METANMAG-003-00277175	METANMAG-003-00277178
METANMAG-003-00277179	METANMAG-003-00277181

METANMAG-003-00129043	METANMAG-003-00129049
METANMAG-003-00131047	METANMAG-003-00131047
METANMAG-003-00131052	METANMAG-003-00131055
METANMAG-003-00131056	METANMAG-003-00131056
METANMAG-003-00131057	METANMAG-003-00131060
METANMAG-003-00131061	METANMAG-003-00131062

Pursuant to Rule 1-033(E), Meta will continue to produce additional information from which answers broadly responsive to this Interrogatory may be derived.

Interrogatory 16. Identify all steps, procedures, and technology You employ to prevent a previously banned User from creating a new account on Meta’s Platforms or to determine whether a previously banned User created a new account on Meta’s Platforms, and, for each step, procedure, or technology, identify the time frame(s) it was in use.

RESPONSE: Meta incorporates its Responses and Objections Applicable to All Requests by reference as though fully set forth herein. Meta further objects to this Interrogatory on the grounds that it is overly broad, seeks to impose an undue burden on Meta that is not proportional to the needs of the case, seeks discovery that is not relevant to the claims and defenses in this action, and is not limited to New Mexico. *See* Rule 1-026(B)(1), (2); *see also* *Lara*, 1999-NMCA-