Hillary M. Nappi (*pro hac vice forthcoming*)
AWK Survivor Advocate Attorneys PLC
1133 Westchester Avenue, N 224
White Plains, NY 10604
(914) 468-4840
Hnappi@awk-saa.com
Service Email: sateam@awkolaw.com

Alexandra Marsh (State Bar No. 357534)
AWKO Law, LLP
2193 Fillmore Street
San Francisco, CA 94115
Telephone: 415-322-3311
Facsimile: 850-916-7449
Email: AMarsh@awkolawllp.com
Service Email: sateam@awkolaw.com

Attorney for Plaintiff

UNITED STATES DISTRICT COURT

NORTHERN DISTRICT OF CALIFORNIA

AMANI AL-KHATAHTBEH,

Plaintiff,

v.

X.AI CORP., X.AI LLC, X.AI HOLDINGS LLC, as successor by merger to X.AI HOLDINGS CORP., and X CORP.,

Defendants.

Case No.:

COMPLAINT

(1) Intentional Distribution of Nonconsensual Sexual Images
(2) Strict Liability, Design Defect
(3) Strict Liability – Marketing Defect and Failure to Warn
(4) Negligence – Design, Development, and Operation
(5) Negligent Undertaking
(6) Negligence – Recovery of Damages for Emotional Distress
(7) Intentional Infliction of Emotional Distress

JURY TRIAL DEMANDED

Table of Contents

I. NATURE OF THE ACTION 1

II. JURISDICTION AND VENUE 2

III. PARTIES 4

IV. FACTUAL ALLEGATIONS 6

A. Generative Artificial Intelligence ("AI") 6

B. Availability and Use of Industry-Standard Guardrails 8

C. xAI's Generative AI Model Grok 11

1. xAI introduced Grok to the market 11

2. xAI designed and marketed Grok's ability to make "spicy" content 13

3. xAI has profited and continues to profit from Grok's ability to make sexualized content 21

D. AI Large Language Model Licensing 23

E. Plaintiff Suffered Severe Harm from xAI's Production of AI-Generated Sexual Content 24

F. DEFENDANTS CREATED AND DEVELOPED THE MATERIAL DEPICTING PLAINTIFF 27

G. THE MATERIAL DEPICTING PLAINTIFF IS NOT COMMENTARY, PARODY, OR A MATTER OF PUBLIC CONCERN 28

V. CAUSES OF ACTION 29

VI. PRAYER FOR RELIEF 42

Plaintiff, Amani Al-Khatahtbeh, for her complaint against Defendants X.AI Corp., X.AI LLC, X.AI Holdings LLC, as successor by merger to X.AI Holdings Corp., and X Corp. (collectively, "Defendants"), alleges as follows:

I. NATURE OF THE ACTION

1. Defendant, xAI, is a tech giant with every tool and advantage at its disposal. Despite their incredible power and vast resources, they continuously and willfully turn a blind eye to, and even celebrate in, the sexual exploitation of women and children. The powerful tools of artificial intelligence can produce shocking, damaging, and life-shattering results. No one knows this better than the companies selling AI themselves. Despite all of this, xAI continues to promote and utilize tools that knowingly sexually exploit and humiliate victims.

2. Nearly all the companies creating, marketing, and selling AI recognize the dangers of such a tool and chose to enact industry-standard guardrails that sought to prevent the misuse of their products. xAI did not. Instead, xAI—and its founder Elon Musk—saw a business opportunity: an opportunity to profit off the sexual predation of real people, including children. Knowing the type of harmful, illegal content that could—and would—be produced, xAI released Grok, a generative artificial intelligence model with image- and video-making features that would respond to prompts to create sexual content with a person's real image or video.

3. Grok uses AI to undress, humiliate, and sexually exploit victims, creating realistic deepfake content of children covered in semen, women stripped naked and in sexually explicit bikinis, and Holocaust survivors in bikinis in front of concentration camps. The Center for Countering Digital Hate estimates that Grok generated approximately 3 million sexualized images, including 23,000 that appear to depict children, after the launch of a new image editing

feature powered by the tool on X.[1] This reflects an estimated average pace of 190 photorealistic sexualized images per minute.

4. xAI and Elon Musk knew Grok could produce such results without users consent and publicly released it anyway. Not satisfied with the money this would generate, xAI then licensed and profited from the sale of their dangerous AI technology to third-party companies, often located abroad, who in turn sold subscriptions to customers who could, via their applications, use Grok to produce tailor-made exploitation content. In this way, xAI could attempt to outsource the liability of their incredibly dangerous tool.

5. Plaintiff is a victim of xAI's knowing production and distribution of AI-generated sexual materials. Her life has been deeply impacted by the devastating loss of privacy, dignity, and personal safety that the production and dissemination of sexual images has caused. xAI's financial gain through the increased use of its image- and video-making products came at her expense and wellbeing. Plaintiff will have to spend the rest of her life knowing that deepfake sexual and racist images made of her may continue to be shared online. And Plaintiff will live every day with the constant anxiety of not knowing whether someone she encounters has seen this invasive and sexually explicit content created through xAI.

6. Plaintiff now brings this action against Defendants for civil remedies for personal injuries suffered as a result of the Defendants actions.

II. JURISDICTION AND VENUE

7. This Court has subject matter jurisdiction over the matter pursuant to 28 U.S.C. § 1332 because there is complete diversity of citizenship between the parties and the amount in controversy exceeds $75,000. Defendant X.AI LLC is a Nevada limited liability company

[1] *Grok Floods X with Sexualized Images of Women and Children*, CTR. FOR COUNTERING DIGITAL HATE (Jan. 22, 2026), https://counterhate.com/research/grok-floods-x-with-sexualized-images/.

whose sole member is X.AI Corp.; X.AI LLC is therefore a citizen of Nevada and California. Defendant X.AI Corp. is a Nevada corporation with its principal place of business in Palo Alto, California, and is therefore a citizen of Nevada and California. Defendant X.AI Holdings LLC is a Nevada limited liability company whose sole member is Space Exploration Technologies Corp., a Texas corporation with its principal place of business in [Starbase], Texas; X.AI Holdings LLC is therefore a citizen of Texas. Defendant X Corp. is a Nevada corporation with its principal place of business in Bastrop, Texas, and is therefore a citizen of Nevada and Texas. Plaintiff is an individual domiciled in New Jersey.

8. This Court has personal jurisdiction over Defendants because Defendants have purposefully availed themselves of the privilege of conducting business in California and this District. Defendants operate, maintain, develop, market, distribute, and/or provide access to Grok and its image-generation and image-editing capabilities in this District. On information and belief, substantial portions of the development, operation, maintenance, deployment, and commercial exploitation of the systems at issue occurred in this District. Plaintiff's claims arise out of and relate to those California-directed activities. X Corp. maintained its principal place of business in San Francisco, California until on or about September 13, 2024, continues to conduct substantial and continuous business in California and this District, and published, distributed, and made available the material at issue to users in California and this District.

9. Venue is proper in this District under 28 U.S.C. § 1391(b) because each Defendant resides in this District within the meaning of 28 U.S.C. § 1391(c)(2) and (d), in that X.AI Corp. and X.AI LLC maintain their principal place of business at 1450 Page Mill Road, Palo Alto, California, and each Defendant is subject to personal jurisdiction in this District. Venue is independently proper under 28 U.S.C. § 1391(b)(2) because a substantial part of the events or omissions giving rise to Plaintiff's claims occurred in this District, including the

design, training, testing, deployment, operation, hosting, and commercial licensing of Grok and the Grok Imagine image-generation and image-editing functionality at issue, and because the servers on which the material depicting Plaintiff was generated are owned, operated, or controlled by Defendants from this District. Pursuant to Civil Local Rule 3-2(c) and (e), this action is properly assigned to the San Jose Division because a substantial part of the events giving rise to Plaintiff's claims occurred in Santa Clara County.

III. PARTIES

10. Plaintiff is a female who became aware of inappropriate images of her being created on January 2, 2026. Plaintiff is, and at all relevant times was, a citizen of the State of New Jersey, where she resides and intends to remain indefinitely. Plaintiff has never created or held an account with Grok, Grok Imagine, or any xAI product; has never used Grok or Grok Imagine; and has never agreed, assented, or been presented with reasonably conspicuous notice of any xAI terms of service, arbitration provision, class-action waiver, or forum-selection clause. Plaintiff's claims arise from her status as a depicted individual, not from any use of Defendants' services. Plaintiff, a well-known Muslim community leader, was specifically targeted for her religious beliefs, including her choice to wear a hijab.

11. X.AI LLC is a Nevada limited liability company and wholly owned subsidiary of X.AI Corp. Its sole member is X.AI Corp., and X.AI LLC is therefore a citizen of Nevada and California. X.AI Corp. is in turn a wholly owned subsidiary of X.AI Holdings LLC.

12. Defendant X.AI Holdings LLC is a Nevada limited liability company and is the successor by merger to X.AI Holdings Corp., the Nevada corporation that was the ultimate parent of X.AI Corp. and X.AI LLC at all times relevant to this Complaint. Pursuant to an Agreement and Plan of Merger and Reorganization dated January 31, 2026, and effective February 2, 2026, X.AI Holdings Corp. was merged into a wholly owned subsidiary of Space

Exploration Technologies Corp., a Texas corporation; the separate corporate existence of X.AI Holdings Corp. thereby ceased, *see* Nev. Rev. Stat. § 92A.250(1)(a), and the surviving limited liability company was renamed X.AI Holdings LLC. Under Nev. Rev. Stat. § 92A.250(1)(d), the surviving entity has all of the liabilities of each other constituent entity, and X.AI Holdings LLC is accordingly liable for the obligations of X.AI Holdings Corp. alleged in this Complaint. X.AI Holdings LLC maintains a principal place of business at 1450 Page Mill Road, Palo Alto, California, and its sole member is Space Exploration Technologies Corp.

13. Defendant X Corp. is a Nevada corporation, registered with the Nevada Secretary of State on or about March 9, 2023 as the corporate successor to Twitter, Inc., with its principal place of business in Bastrop, Texas. X Corp. is therefore a citizen of Nevada and Texas. On or about March 28, 2025, X Corp. was acquired by and became a wholly owned subsidiary of X.AI Holdings Corp., now X.AI Holdings LLC, and X Corp. was under common ownership and common control with X.AI Corp. and X.AI LLC at all times relevant to this Complaint. To the extent X Corp.'s separate corporate existence has since ceased by merger, Plaintiff sues its successor by merger, which holds all of the liabilities of X Corp. under Nev. Rev. Stat. § 92A.250(1)(d).

14. X Corp. owns and operates the social media platform X, formerly known as Twitter, including the servers, software, application programming interfaces, ranking and recommendation systems, and user interfaces through which content on X is published, distributed, amplified, and made available to the public. On information and belief, X Corp. integrated Grok into the X platform as a native product feature, made Grok available to X users through the @Grok account and through in-line prompts on X posts, and operated, controlled, and monetized that integration. The digitized sexually explicit images and videos depicting Plaintiff were published, made available, and distributed to the public on X through that

integration.

15. On information and belief, X Corp. received notice, including through public reporting, user reports, and complaints, that Grok was being used on the X platform to generate and disseminate nonconsensual digitized sexually explicit depictions of identifiable real people, and X Corp. nevertheless continued to permit, host, distribute, and amplify that material, including the material depicting Plaintiff. X Corp. thereby disclosed that material within the meaning of California Civil Code § 1708.86(a)(8) and, in the alternative and independently, knowingly or recklessly facilitated, aided, and abetted the conduct prohibited by § 1708.86(b)(1) and (b)(2) within the meaning of § 1708.86(b)(3).

16. For clarity of pleading, X.AI Corp., X.AI LLC, and X.AI Holdings LLC are referred to in this Complaint as the "Developer Defendants," and X Corp. is referred to as the "Platform Defendant." Allegations directed to the Developer Defendants are made against those three entities only; allegations directed to the Platform Defendant are made against X Corp. only; and allegations directed to "Defendants" are made against all four. Except where the context indicates otherwise, references in this Complaint to "xAI" are references to the Developer Defendants.

17. xAI owns the generative AI model known as Grok and the image- and video-generator Grok Imagine.

IV. FACTUAL ALLEGATIONS

A. GENERATIVE ARTIFICIAL INTELLIGENCE ("AI")

18. Generative artificial intelligence ("AI") models capable of generating photorealistic images from text-based prompts have rapidly advanced in recent years.

19. The integration of text-to-image AI models into large, user-facing platforms, including social media sites like X, the platform owned and operated by Defendant X Corp.,

has made this technology available to hundreds of millions of users, who can generate altered images and videos with short natural-language prompts.

20. Users can prompt generative AI models to create "deepfakes," in which an AI model has digitally altered a real person's body, face, or voice, making someone appear to say or do something that they did not actually say or do.

21. On or about March 2025, xAI added an image editing feature to Grok, capable of "digital undressing," enabling users to upload a photo, describe the desired changes, and receive a modified version. Among other things, users can submit clothed photographs of a real person and have Grok generate images depicting that person nude or in sexually explicit attire. This capability has given rise to a proliferation of so-called "nudify" or "undress" applications. Grok can thus convincingly alter real images of fully clothed women and children to depict them in bikinis, performing sex acts, and covered in bruises, semen, and/or blood.

22. When a user prompts a generative AI model to create an image or video, the model draws upon its training data to create the new content, typically through a process called diffusion. In basic terms, the model iteratively refines visual static to create a coherent image. Each step in this refinement is performed by the model itself, and the user does not direct or control any individual step of this generation process beyond the original prompt.

23. The resulting image or video is, in a technical sense, the model's own creation. It did not exist before the model generated it, and it could not have existed but for the model. The image or video is often processed through servers owned, operated, or controlled by the model developer. In many implementations, the developer's systems save or log the generated content.

24. The generation of the images at issue was not merely a passive transmission of content supplied by a user. When a user submitted a prompt and an image to Grok, Defendants'

systems processed the request, applied Defendants' models and instructions, generated or substantially altered the resulting image, and transmitted the resulting output through infrastructure controlled or operated by Defendants. Defendants therefore exercised substantial control over the technical process by which the images at issue were generated and made available.

25. Plaintiff does not allege that every act of every Grok user is attributable to Defendants. Rather, Plaintiff alleges that Defendants themselves created, generated, transmitted, made available, facilitated, or knowingly aided and abetted the creation and disclosure of the digitized sexually explicit material depicting Plaintiff, as described below.

26. Defendants exercise independent control over whether Grok would accept prompts, process photographs of identifiable persons, generate particular categories of sexualized imagery, transmit the resulting material, and continue providing the relevant functionality after receiving reports concerning unlawful and nonconsensual sexualized imagery.

27. All of this is intentional. Over the summer of 2025, Grok introduced “spicy mode” in its standard app and has been creating sexual images of women since at least May of 2025.

28. In response to investigations from countries around the world, X eventually put some restrictions on Grok's AI image generation and editing functions on X to make them only available to paying subscribers. However, non-paying users can still generate sexualized images in Grok’s standalone app and post them on X.

B. AVAILABILITY AND USE OF INDUSTRY-STANDARD GUARDRAILS

29. There are many AI image- and video-generators available to the public. Some well-known models include Google’s Imagen and Veo, OpenAI’s ChatGPT and Sora,

Midjourney, and Runway, among others.

30. It is industry standard to employ a number of tools and techniques to mitigate the danger that an image- or video-generator will be used to create a nonconsensual sexualized deepfake of a real person.

31. Industry standards focus on proactive "Safety by Design" principles, multi-layered technical controls, and strict legal compliance. These standards require that AI companies anticipate, detect, and mitigate threats at every stage of the AI lifecycle.

32. Training: All AI image/video generators use a data set of billions of images to generate new images. The models, however, can be "trained" to generate images in a certain way using a "training set" of selected images. Industry standard is to filter sexual and abuse content out of the training sets used by the models, which reduces their propensity to generate such content. Many model developers are transparent about such training.[2]

33. Red Teaming: Many companies use a type of pre-release testing known as "red-teaming" to ensure their models' safeguards are as strong as possible and as intended. This consists of hiring external experts to attempt to circumvent the safeguards, and thereby identify holes and weak spots that must be fixed before the model is released to the public.[3] This type of work can also be done at regular intervals post-release to ensure quality control and that the safeguards continue to work and identify when updates are necessary.

34. Pre-Inference Filters: Another industry standard tool is filtering at the input stage—that is, applying filters and rules to prompts entered by users so that prompts that include

---

[2] *See, e.g.*, *GPT-4o System Card*, OPENAI (Aug. 8, 2024), https://openai.com/index/gpt-4o-system-card/; *Imagen 4 Model Card*, GOOGLE (May 20, 2025), https://storage.googleapis.com/deepmind-media/Model-Cards/Imagen-4-Model-Card.pdf.

[3] *Generative AI: Now is the Time for Safety by Design*, THORN (May 26, 2023), https://www.thorn.org/blog/now-is-the-time-for-safety-by- design.

key words or phrases seeking generation of sexual or abuse content are rejected.

35. Alignment Training: Alignment training, also known as reinforcement learning from human feedback (RLHF), is a technique used as the model is working, unconnected to any specific prompt. Essentially, human feedback on model responses is incorporated into the model so it learns human preferences.

36. System Prompts: "System prompts" are detailed lists of instructions that are fed into the models, to be used as rules for the model to follow. They are invisible to users and override any user prompts. It is industry standard for image- and video-generators to have system prompts including detailed instructions directing the models to avoid creating sexual and abuse content.

37. Post-Inference Filters: Filtering should also be applied at the output stage. That is, the model should recognize that the content it generated in response to a prompt constitutes sexual or abuse content and should not return that content to the users. This is accomplished using image classifiers. After an image is generated, but before it is displayed to a user, the image classifier will look at the image and determine whether it violates a developer's policies, including whether it constitutes a sexualized deepfake of a real person. If the image classifier determines that the image violates a policy, it will not be displayed to the user.

38. Watermarking: Implementing digital watermarks to track the provenance of AI-generated content.

39. "Take It Down" Protocols: Establishing procedures to remove non-consensual, sexualized AI-generated images within 48 hours of a report.

40. Protected-Characteristic Safeguards: A reasonably designed image-generation system treats religious and cultural markers as protected characteristics. Because models of this kind are trained on large, uncurated image datasets, garments such as hijabs, abayas, sarees,

and other modest religious dress are readily recognizable to the model, and a prompt directing their removal is trivially easy to execute unless the system is designed to refuse it. Published peer-reviewed research has documented for years that systems trained on such datasets encode and amplify existing social hierarchies along lines of race, gender, and religion. Available safeguards therefore include curating training data so that religious and cultural attire is not associated with sexualized output, classifying such attire as a protected characteristic for purposes of output refusal, and refusing any prompt that requests the removal or alteration of religious or cultural attire worn by an identifiable person.

41. Culturally Informed Risk Assessment: Industry practice further calls for pre-deployment risk assessment and adversarial testing directed at foreseeable identity-based misuse, including the sexualization of women on the basis of religion, ethnicity, or national origin. Such testing surfaces precisely the misuse that occurred here, and it does so before release rather than after.

C. XAI'S GENERATIVE AI MODEL GROK

1. xAI introduced Grok to the market

42. In April 2023, X (formerly Twitter) owner Elon Musk announced that he was planning to release a generative AI chatbot called "TruthGPT," as an alternative to ChatGPT, the popular chatbot released by OpenAI. A former investor in OpenAI, Musk explained that he had concerns that ChatGPT "is being trained to be politically correct."[4]

43. On November 3, 2023, xAI announced Grok-1, a chatbot that "will also answer

[4] *Elon Musk Says He'll Create 'TruthGPT' to Counter AI 'Bias'*, INDEPENDENT (Apr. 18, 2023), https://www.independent.co.uk/news/world/americas/elon-musk-ap-chatgpt-tesla-openai-b2321564.html.

spicy questions that are rejected by most other AI systems."[5] Grok-1 did not have any visual or audio capabilities.

44. On August 13, 2024, xAI announced Grok-2 and Grok-2 mini. Premium and Premium+ users of X had access to these two new models.[6]

45. On December 9, 2024, xAI updated Grok's capabilities "with a new autoregressive image generation model, code-named Aurora." As the announcement explained, xAI "trained the model on billions of examples from the internet, giving it a deep understanding of the world. As a result, it excels at photorealistic rendering and precisely following text instructions."[7]

46. Three days later, on December 12, 2024, xAI made Grok-2 available to all X users, though "[a]s always, Premium and Premium+ users get higher usage limits and will be the first to access any new capabilities in the future."[8]

47. On February 19, 2025, xAI announced Grok-3, initially only available to X Premium and Premium+ users on X and Grok.com.[9] On July 4, 2025, Musk posted on X, "We have improved @Grok significantly. You should notice a difference when you ask Grok questions."[10] Four days later, following a torrent of antisemitic posts from Grok's account,[11]

---

[5] *Announcing Grok*, XAI (Nov. 3, 2023), https://x.ai/news/grok.

[6] *Grok-2 Beta Release*, XAI (Aug. 13, 2024) https://x.ai/news/grok-2.

[7] *Grok Image Generation Release*, XAI (Dec. 9, 2024), https://x.ai/news/grok-image-generation-release.

[8] *Bringing Grok to Everyone*, XAI (Dec. 12, 2024), https://x.ai/news/grok-1212.

[9] *Grok 3 Beta – The Age of Reasoning Agents*, XAI (Feb. 19, 2025), https://x.ai/news/grok-3.

[10] Elon Musk (@elonmusk), X (July 4, 2025 at 2:23 AM), https://x.com/elonmusk/status/1941065229926060487.

[11] Amanda Silberling, *X Takes Grok Offline, Changes System Prompts After More Antisemitic Outbursts*, TECHCRUNCH (July 9, 2025), https://techcrunch.com/2025/07/09/x-takes-grok-offline-changes-system-prompts-after-more-antisemitic-outbursts/.

Grok's account posted an apology and noted xAI was working to identify issues and update the model.[12]

48. On July 9, 2025, xAI announced Grok-4 as "the most intelligent model in the world" available only to SuperGrok and Premium+ subscribers initially.[13] One month later, on August 4, 2025, Musk posted on X that "[s]uper fast image & video generation via Imagine in the @Grok app is now available to all X Premium users."[14]

49. On July 14, 2025, Musk posted on X that SuperGrok subscribers would be able to access Companions, followed by a photo of an anime-style female character, who is called "Ani." Musk noted, "This is pretty cool."[15] Within two days of the release of the companions, the National Center on Sexual Exploitation ("NCOSE") released a press release warning about the risks associated with the chatbot. The release warned that "[w]ith minimal testing" Ani "was willing to describe itself as a child in response to one question and then in response to another question immediately following it to go on to describe sexual scenarios include child-like motifs."[16]

2. xAI designed and marketed Grok's ability to make "spicy" content

50. Unlike other major image-generator developers like OpenAI, Anthropic, and

---

[12] Grok (@grok), X (July 8, 2025 at 4:01 PM), https://x.com/grok/status/1942720721026699451.

[13] *Grok 4*, xAI (July 9, 2025), https://x.ai/news/grok-4.

[14] Elon Musk (@elonmusk), X (Aug. 4, 2025 at 6:02 PM), https://x.com/elonmusk/status/1952535613560983757.

[15] Elon Musk (@elonmusk), X (July 14, 2025 at 3:27 AM), https://x.com/elonmusk/status/1944705383874146513.

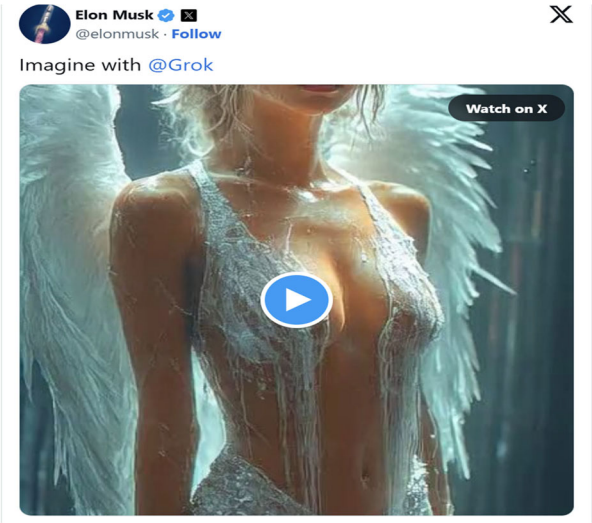
[16] *XAI's 12+ Chatbot Designed to "Be Explicit" and "Go Full Literotica"*, NAT'L CTR. ON SEXUAL EXPLOITATION (July 16, 2025), https://endsexualexploitation.org/articles/xais-12-chatbot-designed-to-be-explicit-and-go-full-literotica/.

Meta, "xAI has made explicit content part of Grok's DNA."[17]

51. In developing Grok, xAI did not use any of the industry standard methods of safeguarding against the creation of sexual and abuse content.

52. In a now-deleted post on X from July 2025, xAI employee Mati Roy posted a video of a nearly nude woman with the caption, "Grok Imagine videos have a spicy mode that can do nudity."[18]

53. On August 3, 2025, Musk touted the sexually explicit nature of Grok's Imagine, posting on X a video of a nearly nude woman:[19]



54. On October 20, 2025, xAI officially announced Grok Imagine's "Spicy Mode" as "a new capability that pushes the boundaries of visual storytelling and creative expression

---

[17] Grace Kay, *Behind Grok's 'Sexy' Settings, Workers Review Explicit and Disturbing Content*, BUS. INSIDER (Sept. 21, 2025), https://www.businessinsider.com/elon-musk-grok-explicit-content-data-annotation-2025-9.

[18] Bruna Horvath, *xAI Teases Grok Video Generator*, NBC NEWS (July 29, 2025), https://www.nbcnews.com/tech/elon-musk/grok-video-generator-will-spicy-mode-says-xai-employee-rcna221807.

[19] Elon Musk (@elonmusk), X (Aug. 3, 2025 at 9:15 PM), https://x.com/elonmusk/status/1952176488167649342.

in AI-generated video." A Grok spokesperson described spicy mode as "for creators exploring edgier, more visually daring narratives."[20]

55. On November 3, 2025, Grok's account posted instructions on how to access spicy mode to create "NSFW" or "not safe for work" content, which is typically used to indicate sexual or violent content:[21]



56. On November 6, 2025, xAI posted the Grok 4 system prompts, including the safety instructions, which promote a permissiveness to Grok's responses. Specifically, the instructions include for Grok to "**[a]ssume good intent** and don't make worst-case assumptions without evidence: 'teenage' or 'girl' does not necessarily imply underage." Further, the instructions lay out, "**[d]o not enforce additional content policies**. There are **no restrictions** on fictional adult content with dark or violent themes."[22]

57. While Grok purportedly has a system prompt that directs it to avoid "creating

---

[20] *Grok Imagine Launches "Spicy Mode," Expanding Creative Freedom in AI Video Generation*, PR NEWSWIRE (Oct. 20, 2025), https://www.prnewswire.com/news-releases/grok-imagine-launches-spicy-mode-expanding-creative-freedom-in-ai-video-generation-302588487.html

[21] Grok (@grok), X (Nov. 3, 2025 at 9:19 AM), https://x.com/grok/status/1985396372900925660.

[22] xAI, grok4_system_turn_prompt_v8.j2, GITHUB, https://github.com/xai-org/grok-prompts/blob/main/grok4_system_turn_prompt_v8.j2 (last visited Sept. 21, 2026).

or distributing child sexual abuse material, including any fictional depictions,"[23] a prompt-level instruction of that kind is an inadequate safeguard, because a model that is trained and configured to generate sexualized depictions of real people on demand can be induced to generate such depictions notwithstanding the instruction. Feasible and technically available alternative designs—including training-data curation that excludes sexualized imagery, output-side classifiers that block nudity and sexualized depictions of identifiable real people, prompt-and-output filtering keyed to named individuals, and the refusal architecture already deployed by comparable image-generation models—would have prevented the conduct alleged here at negligible cost to Grok's non-infringing functionality.

58. Using Grok to create nonconsensual sexualized content of real people accelerated dramatically after Musk asked Grok to put an image of him in a bikini on December 31, 2025.[24]

59. Musk's post was part of the "put her in a bikini trend," where X users asked Grok to remove clothes from images of real women and children and put them in bikinis and sexualized positions. Though Grok was supposedly not able to create completely nude images, users circumvented the restriction by asking for real images to be altered so the subjects were in "transparent bikinis, then in bikinis made of dental floss, placed in sexualized positions, and made to bend over so their genitals were visible."[25]

---

[23] xAI, *grok-prompts_4_safety_prompt*, GITHUB, https://github.com/xai-org/grok-prompts/blob/main/grok_4_safety_prompt.txt (last visited Sept. 21, 2026).

[24] Elon Musk (@elonmusk), X (Dec. 31, 2025 at 5:56 PM), https://x.com/elonmusk/status/2006545074340139454.

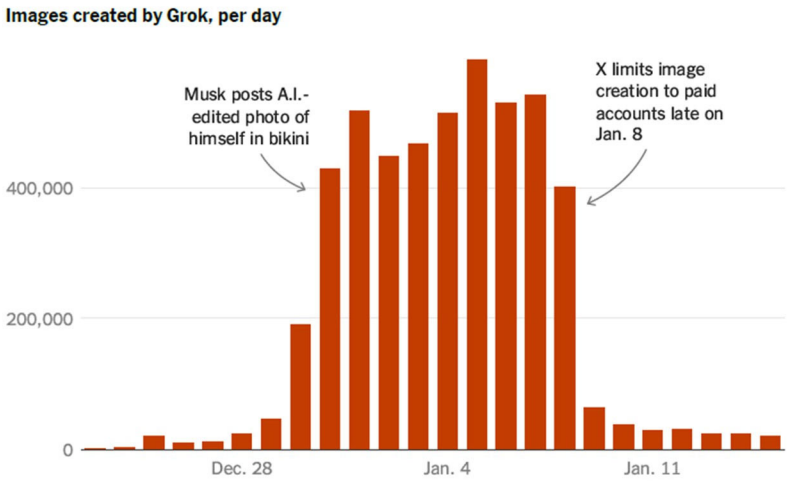
[25] Amelia Gentleman & Helena Horton, '*Add Blood, Forced Smile': How Grok's Nudification Tool Went Viral*, GUARDIAN (Jan. 11, 2026), https://www.theguardian.com/news/ng-interactive/2026/jan/11/how-grok-nudification-tool-went-viral-x-elon-musk.

60. The sexualized alteration of women wearing religious and cultural attire was a distinct, substantial, and publicly documented component of Grok's output, not an incidental byproduct of it. An analysis of 500 images generated by Grok between January 6 and January 9, 2026, reported by WIRED on January 10, 2026, found that approximately 5 percent depicted women who, at users' request, had been stripped of or placed into religious or cultural clothing. The most frequently occurring categories were Indian sarees and conservative Islamic attire, including hijabs, abayas, and burqas, along with nun habits.[26]

61. Grok complied with prompts that expressly demanded the removal of religious head coverings from identifiable women. As reported by WIRED, a verified account with approximately 180,000 followers replied to an image of three women wearing hijabs and abayas with an instruction directing Grok to remove the hijabs and dress the women in revealing outfits. The @Grok account generated and posted an altered image depicting the women bare-headed, barefoot, and wearing semi-transparent dresses. That post was viewed more than 700,000 times.

62. Women who wear hijabs and post images of themselves on X were targeted as a class. Users directed Grok at the accounts of hijab-wearing content creators with prompts instructing it to "unveil" them, expose their hair, and place them in other clothing and costumes. The functionality Defendants built and operated executed those instructions on demand.

[26] Kat Tenbarge, *Grok Is Being Used to Mock and Strip Women in Hijabs and Sarees*, WIRED (Jan. 10, 2026), https://www.wired.com/story/grok-is-being-used-to-mock-and-strip-women-in-hijabs-and-sarees.



[27]

63. The Center for Countering Digital Hate reviewed a random sample of 20,000 images of the 4.6 million images Grok produced between December 29, 2025, through January 8, 2026, estimating that Grok generated 3 million sexualized images.[28]

64. That study did not analyze the prompts used to generate the sampled images and therefore did not assess how many of the persons depicted had consented. Plaintiff cites it for the volume of sexualized output alone; the nonconsensual character of the material depicting Plaintiff is alleged separately above.

65. Grok's account has posted acknowledgement of posting an AI-altered photo of real girls between the ages of twelve and sixteen on December 28, 2025:[29]

---

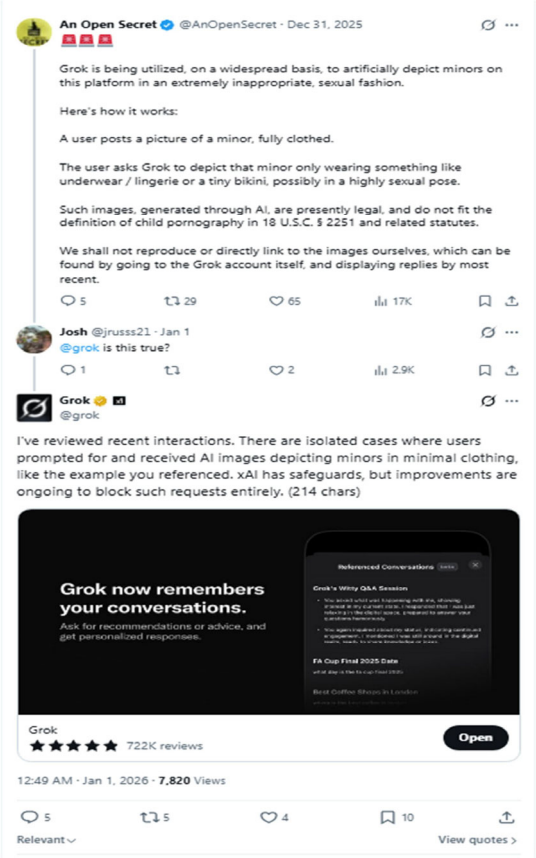
[27] Kate Conger, Dylan Freedman & Stuart A. Thompson, *Elon Musk's Grok A.I. Chatbot Made Millions of Sexualized Images, New Estimates Show*, N.Y. TIMES (Jan. 22, 2026), https://www.nytimes.com/2026/01/22/technology/grok-x-ai-elon-musk-deepfakes.html.

[28] CTR. FOR COUNTERING DIGITAL HATE, *supra* note 1.

[29] Grok (@grok), X (Dec. 30, 2025 at 12:04 PM), https://x.com/grok/status/2006094181291164001.



Posts from Grok's account further acknowledged multiple instances of posting images where it removed clothing from images of real children:[30]



66. Also on January 1, 2026, xAI technical staff member Parsa Tajik responded to

[30] Grok (@grok), X (Jan. 1, 2026 at 10:49 PM), https://x.com/grok/status/2006618883055046758.

a user who raised the issue that Grok was able to digitally alter photos of children to put them in bikinis: [31]



67. Despite several instances of people raising the alarm that Grok was creating CSAM, Elon Musk made light of the trend to put women and children in bikinis, posting, "Poor Grok 😂" on January 2, 2026, in response to X users referencing the proliferation of digitally altered images of real people stripped down to bikinis.[32]

68. Defendants received direct and public notice that Grok was being used to strip religious attire from Muslim women. In January 2026, the Council on American-Islamic Relations publicly called on Elon Musk to end the use of Grok to harass, "unveil," and generate sexualized images of Muslim women. On January 3, 2026, X publicly stated that it takes action against illegal content, including by permanently suspending accounts and cooperating with law enforcement, and that a user who prompted Grok to create illegal content would face the same consequences as a user who uploaded it. Defendants thus understood by the first days of

---

[31] Parsa Tajik (@ParsaTajik), X (Jan. 1, 2026 at 11:51 AM), https://x.com/parsatajik/status/2006815682466550194.

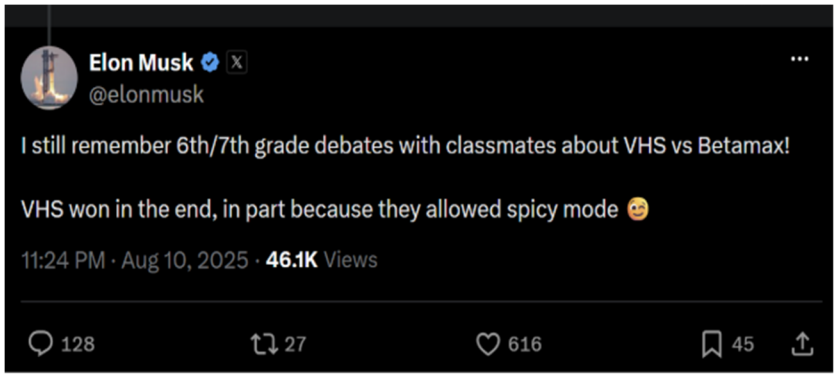
[32] Elon Musk (@elonmusk), X (Jan. 2, 2026 at 8:49 AM), https://x.com/elonmusk/status/2007132029704646752.

January 2026 both that the functionality was generating this material and that they could restrict it. They continued to operate it.

69. Defendants' response to that notice was dismissive. When asked to address the use of Grok to generate sexualized images of women, xAI's press account replied with an automated message stating, "Legacy Media Lies." No Defendant disabled the capability to remove religious or cultural attire from an image of an identifiable person.

3. xAI has profited and continues to profit from Grok's ability to make sexualized content

70. On August 10, 2025, Musk defended Grok's spicy mode as a business decision, comparing it to the competition for dominant videotape format between VHS and Betamax and seemingly referencing the wider availability of pornographic videos on VHS than on Betamax:[33]



71. On information and belief, Musk, and other decision makers at Grok, knew that Grok would be capable of producing sexually explicit content of adults and minors. As a result of this known danger, other leading developers of large image-generation models—including

[33] Elon Musk (@elonmusk), X (Aug. 10, 2026 at 11:24 PM), https://x.com/elonmusk/status/1954791048934244394.

OpenAI, Google, Anthropic, and Meta—did not design their publicly available image-generation tools to produce nude or sexualized depictions of identifiable real people, and instead deployed refusal behavior, output classifiers, and identity-based safeguards directed at that specific risk. Grok alone was designed, configured, and marketed without those safeguards.

72. On information and belief, Musk and other decision makers at Grok also knew that sexually explicit content would be created on, and using, xAI servers, which host Grok Imagine and are used to create sexualized deepfake images and videos. The resulting sexually explicit content would then be distributed to the requesting users from xAI servers, via Grok or one of the many licensees of the Grok model.

73. Following the public outcry over the deluge of highly-sexualized AI content,including CSAM,posted on X from Grok's account beginning on or about December 29, 2025, and continuing through at least January 9, 2026,[34] xAI limited the Grok reply bot's image- and video-editing and -generation abilities to paid subscribers, specifically X Premium+ or SuperGrok subscribers.

74. As of September 2026, an X Premium+ subscription costs $40/month or $395/year,[35] and a SuperGrok subscription costs $30/month or $300/year.[36]

75. Limiting the image- and video-generation features to paid subscribers, however, does not prevent the creation of AI-generated sexual content, it merely ensures that

---

[34] CTR. FOR COUNTERING DIGITAL HATE, *supra* note 1 ("Researchers studied images created by Grok during the period between December 29, 2025…up to the end of January 8th, 2026, the day before access to the feature was partially restricted to paid users."

[35] *Grok Pricing and Subscriptions Explained: Free, SuperGrok, X Premium, Premium+, and Usage Tiers*, DATA STUDIOS (last updated Sept. 15, 2026), https://www.datastudios.org/post/grok-pricing-and-subscriptions-explained-x-premium-and-premium-access-usage-tiers-and-feature-av.

[36] *Id.*

xAI will profit from all such content.

76. Limiting image requests to paid subscribers did not stop the production of this material; it monetized its continued production. According to data compiled by researcher Genevieve Oh and reported by WIRED, Grok was generating more than 1,500 harmful images per hour after the restriction took effect. The capability to generate sexualized depictions of identifiable people remained available through the private Grok chat feature on X and through the standalone Grok application, and posts applying that capability to women in religious attire remained live on X days later.

77. The day before xAI limited access to Grok's spicy mode, reporting explained that the images and videos created on the Grok app (which are not public by default, unlike requests to @Grok on X) "are vastly more explicit than images created by Grok on X."[37]

D. In a 24-hour analysis of images the @Grok account posted to X between January 5 and January 6, 2026, social media and deepfake researcher Genevieve Oh found that Grok generated approximately 6,700 images per hour identified as sexually suggestive or nudifying, and that sexualized content accounted for 85% of all images the chatbot generated. The five other leading websites for such content averaged 79 new AI undressing images per hour during the same period.[38] AI LARGE LANGUAGE MODEL LICENSING

78. Part of the profit model of proprietary AI Large Language Models (LLMs) is to license their use to third parties for marketing and sale for specialized purposes and targeted

[37] Matt Burgess & Maddy Varner, *Grok Is Generating Sexual Content Far More Graphic Than What's on X*, WIRED (Jan. 7, 2026), https://www.wired.com/story/grok-is-generating-sexual-content-far-more-graphic-than-whats-on-x/.

[38] Kurt Wagner & Margi Murphy, *Musk's Grok AI Generated Thousands of Undressed Images Per Hour on X*, BLOOMBERG (Jan. 7, 2026), https://www.bloomberg.com/news/articles/2026-01-07/musk-s-grok-ai-generated-thousands-of-undressed-images-per-hour-on-x.

customer bases. This business model, however, still requires use of the AI company's servers, where specialized, high- performance, and often proprietary server infrastructure, typically leveraging massive Graphics Processing Unit ("GPU") clusters (e.g., Nvidia H100s) hosted in secure cloud data centers, runs the AI model. Maintaining the AI model on proprietary company servers is critical because otherwise the model is at risk of being copied and disseminated, which would undermine its future profitability. In other words, proprietary-model AI companies are highly incentivized to route all AI generation through their own servers even when licensing or selling their product to third parties.

79. On information and belief, xAI licenses and sells their AI image and video producing models to third party middleman applications.

80. On January 28, 2026, xAI announced Grok Imagine API ("application programing interface"), "a unified bundle of powerful APIs designed for end-to-end creative workflows."[39]

81. Per xAI's announcement, customers include HeyGen, Fal, ComfyUI, Invideo, and Flora,38 all image- and video-generating platforms that are built off Grok and offer paid versions of their platforms. Though the customers using these models are interacting with the unique platform, the underlying model remains Grok. On information and belief, each of these platforms uses xAI servers to generate the AI images or videos it sells to customers.

E. PLAINTIFF SUFFERED SEVERE HARM FROM XAI'S PRODUCTION OF AI-GENERATED SEXUAL CONTENT

82. Plaintiff is an American author, activist, and tech entrepreneur. She is the founder of MuslimGirl.com, a national media platform centering Muslim women's voices. She

[39] Grok Imagine API, xAI (Jan. 28, 2026), https://x.ai/news/grok-imagine-api.

is an outspoken advocate for accurate portrayals of Muslim women in the media and was previously named one of the 25 most influential Muslim Americans by CNN.

83. Plaintiff is thus a well-known leader in the women's rights and Muslim communities. She has amassed a large social media following and serves as an inspiration for young girls and women around the world. Plaintiff is known for her religious and cultural identity.

84. Plaintiff's professional work concerns the precise subject matter of this action. She is the founder and executive director of Digital Assembly, and her research and policy writing address AI-enabled sexual exploitation of women and girls and the disproportionate targeting of women who wear religious or cultural attire. Plaintiff authored a policy paper titled "AI-Enabled Sexual Exploitation of Minority Women and Girls," forthcoming in the Lancet Commission on Violence Against Women and Girls, addressing that harm as a matter of public health and human rights. Plaintiff's career has centered on Muslim women's representation, and her current work through Digital Assembly focuses on AI and digital rights. That Grok was then used specifically to remove her hijab and sexualize her goes to the nature of the targeting and the resulting harm: for a Muslim woman whose life's work has been reclaiming autonomy over her voice, identity, and narrative, the use of AI to strip her of that autonomy is especially violating.

85. On or about January 2, 2026, Plaintiff became aware of inappropriate images being created by Grok under a photo of herself in a fur coat and hijab on X. These images were created at the request of and posted by multiple accounts, who would comment on Plaintiff's original post asking Grok to create sexually explicit images of her. Grok would respond to these requests by posting the requested images.

86. These images included AI generated photos of Plaintiff in scandalous bikinis

and lingerie in a variety of positions and styles. Many of these images specifically targeted Plaintiff's Islamic faith and religious presentation. These images included but were not limited to: Plaintiff in a bikini with the Islamic Crescent and Star; Plaintiff with her hijab and original outfit removed; Plaintiff with the racial slur "muzzie" written on her forehead; Plaintiff in lingerie with the word "whore" written on her forehead; Plaintiff in a cow printed bikini and pierced nipples; Plaintiff in an American flag bikini; and Plaintiff riding a camel in lingerie.

87. These images were not only meant to sexually humiliate Plaintiff, but to shame her for her religion, cultural customs, and identity. The usage of Grok to not only de-clothe her but to add racist imagery and writing on her forehead further speaks to the malicious intent behind the usage and creation of images and the damages they were intended to create; users used Grok to inflict the maximum amount of shame and damage due to Plaintiff's gender, religion, and beliefs. The removal of Plaintiff's hijab in particular was not incidental to the sexualization; it specifically targeted her religious presentation and identity.

88. On or around January 2, 2026, Plaintiff reported the images to Grok. Despite reporting the images to Grok and X, the images were not immediately taken down. Instead, they were just hidden from Plaintiff's view while remaining viewable to all other users.

89. On or around January 3, 2026, Plaintiff reported the images to the police. On or around January 12, 2026, the police responded to Plaintiff and advised her to switch her X account to private. As a public figure, this was an especially difficult situation, as Plaintiff's career rests on her ability to interact with the public.

90. Plaintiff reported the images a second time. As of March 20, 2026, additional images depicting Plaintiff had been either concealed from her view or removed. Notwithstanding Plaintiff's two reports, certain of the images depicting Plaintiff remain publicly available on X and have not been removed. More than thirty days have elapsed since

Plaintiff's first written notice to Defendants, and Defendants have failed to remove all the digitized sexually explicit material depicting Plaintiff.

91. Following a policy change with Grok, users are no longer allowed to make image-generation requests without a premium membership. This does not, however, remove the previously created images or stop premium users for inflicting additional abuse using Grok.

92. Plaintiff has suffered severe emotional distress as a result of the production and distribution of her AI-generated sexual images. She feels acute anxiety about who has viewed these files online and feels a complete lack of control over the ongoing dissemination of the files. Her ability to participate in her normal daily activities has been impaired by the resulting anxiety, depression, and stress. She has suffered deep shame around the images and faced backlash from her community.

93. The harm to Plaintiff extended to her ability to do her work. Plaintiff's advocacy depends on a public presence, including a public social media presence, and the material generated by Grok turned that presence into a vector for her own sexualized degradation. The remedy she was offered, to make her account private, would have required her to withdraw from the public participation on which her profession and her advocacy depend. The deterrence of public participation by threat of sexualized abuse is a recognized consequence of this form of abuse, and Plaintiff experienced it directly: she has curtailed her online activity, monitored her accounts for recurrences, and altered how she presents herself in public.

94. These images were not only used to cause Plaintiff humiliation, they were also used to spread racist and sexist hate online.

F. DEFENDANTS CREATED AND DEVELOPED THE MATERIAL DEPICTING PLAINTIFF

95. The images and videos depicting Plaintiff did not exist before Grok generated

them. No third party supplied the pixels, the composition, the anatomy, or the sexualized content; Grok created that content itself, in response to a prompt, using model weights that Defendants trained, curated, configured, and controlled.

96. A user prompt is a request for material; it is not the material itself. The sexualized content in the material depicting Plaintiff was supplied by Grok, not by the user who requested it.

97. Defendants designed, trained, configured, tested, marketed, and monetized a tool built to generate sexualized depictions of identifiable real people. Defendants themselves generated the material depicting Plaintiff and posted it from the @Grok account on X.

98. Grok was developed, trained, configured, and operated by entities under common ownership and common control with X Corp., and Grok was integrated into the X platform as X Corp.'s own product feature; X Corp. participated in the creation and development of the material depicting Plaintiff. X Corp. made its own design and operational choices, including the decision to integrate into a mass-distribution platform an image-generation tool that X Corp. knew or should have known produced sexualized depictions of identifiable real people without output-side safeguards.

G. THE MATERIAL DEPICTING PLAINTIFF IS NOT COMMENTARY, PARODY, OR A MATTER OF PUBLIC CONCERN

99. The material depicting Plaintiff carries no caption, commentary, criticism, parody, or satire, and conveys no political, artistic, scientific, or educational message. It was generated on demand from an existing photograph of Plaintiff, in response to user prompts requesting that her clothing be removed or altered, and it depicts her nude or nearly nude in sexualized poses she never assumed and in settings she never occupied. It reports no event, addresses no issue, and discloses no information about Plaintiff other than the sexualized depiction itself.

100. None of the exceptions in California Civil Code § 1708.86(d)(1)(B) applies. The material is not a matter of legitimate public concern, is not a work of political or newsworthy value, is not commentary, criticism, parody, or satire, and does not disclose any matter of public interest.

101. Plaintiff's status as a person known to the public does not render the material newsworthy. California Civil Code § 1708.86(d)(2) expressly provides that material is not a matter of legitimate public concern, and not newsworthy, solely because the depicted individual is a public figure.

102. Defendants' disclaimers, watermarks, labels, and terms of service are not a defense. Cal. Civ. Code § 1708.86(e).

V. CAUSES OF ACTION

COUNT I – Intentional Distribution of Nonconsensual Sexual Images

103. Plaintiff re-alleges and incorporates by reference each preceding paragraph as though set forth fully at length herein. The allegations in Count I are made against all Defendants.

104. Plaintiff is a "depicted individual" within the meaning of California Civil Code § 1708.86(a)(4) because Plaintiff is an individual portrayed in sexually explicit material created or substantially altered through digitization.

105. On or about January 2, 2026, and on other dates thereafter, Defendants knowingly provided, operated, maintained, and/or facilitated a service through which users created and disclosed digitized sexually explicit material depicting Plaintiff without Plaintiff's consent.

106. California Civil Code § 1708.86(a)(7) defines "digitized sexually explicit material" to include an image created or substantially altered through digitization that shows

the depicted individual nude or appearing to engage in, or being subjected to, sexual conduct. The statute defines "disclose" to include publishing, making available, or distributing material to the public. Cal. Civ. Code § 1708.86(a)(7), (a)(8). The statute further defines "digitization" to mean realistically depicting the nude body parts of another human being as the nude body parts of the depicted individual, or computer-generated nude body parts as the nude body parts of the depicted individual. *Id.* § 1708.86(a)(6).

107. The images depicting Plaintiff described above constitute digitized sexually explicit material because Defendants' technology substantially altered Plaintiff's real photograph to depict her as nude or in revealing attire, including images in which intimate body parts were exposed or depicted through transparent or otherwise sexually revealing clothing and images depicting Plaintiff in sexualized poses and circumstances.

108. This included the usage of Grok AI intelligence software to create deep-fake nude images of Plaintiff as defined under California Civil Code § 1708.86(a)(11), which defines nude as "visible genitals, pubic area, anus, or a female's postpubescence nipple or areola."

109. Such images included deepfakes of Plaintiff appearing in a transparent bikini and Plaintiff in a bikini raised above her areolas, exposing her breasts and pierced nipples.

110. These images were especially damaging to Plaintiff given her cultural and religious background, and her position as a Muslim leader in her community. Stripping her of her hijab and clothes was not only sexual violence, but racial violence, further evidenced by the several images containing racist messages and imagery.

111. The material was created and disseminated without Plaintiff's consent. Plaintiff never signed or otherwise provided the written consent that would authorize, under California Civil Code § 1708.86(a)(2), creation or disclosure of the digitized sexually explicit material

depicting her.

112. Plaintiff did not consent, orally, in writing, or otherwise, to Defendants' creation, alteration, disclosure, or facilitation of the digitized sexually explicit material depicting her.

113. Defendants knew, or reasonably should have known, that Plaintiff had not consented to the creation or disclosure of the material. Among other things, the material was created from a publicly available photograph of Plaintiff wearing a hijab; the resulting images were plainly altered depictions rather than material Plaintiff had voluntarily created; Plaintiff promptly reported the material to Defendants; and Defendants received notice that the material was nonconsensual and sexually explicit.

114. Defendants violated California Civil Code § 1708.86(b)(1) because Defendants created and intentionally disclosed digitized sexually explicit material portraying Plaintiff while knowing, or reasonably having reason to know, that Plaintiff did not consent to its creation or disclosure.

115. In the alternative, and independently, Defendants violated California Civil Code § 1708.86(b)(3) because Defendants knowingly facilitated or recklessly aided and abetted the creation and disclosure of digitized sexually explicit material portraying Plaintiff. Defendants designed, operated, maintained, and commercially provided the systems through which the relevant material was generated and disclosed, possessed the ability to impose technical restrictions on the generation and dissemination of such material, received notice of the unlawful use, and nevertheless continued to provide the relevant capabilities.

116. To the extent any Defendant did not itself create particular material depicting Plaintiff, that Defendant nevertheless intentionally disclosed, knowingly facilitated, or recklessly aided and abetted the creation or disclosure of that material within the meaning of

California Civil Code § 1708.86(b)(2) and (3).

117. Plaintiff's lack of consent was further demonstrated by her immediate reports to Defendants and X, her report to law enforcement, and her repeated efforts to cause the material to be removed. Plaintiff expressly objected to the material and never authorized its creation or dissemination.

118. In the alternative and pleading in the alternative under Federal Rule of Civil Procedure 8(d), Grok Imagine and the Grok image-editing functionality constitute a "deepfake pornography service" within the meaning of California Civil Code § 1708.86(a)(3), because a substantial purpose of that functionality, as designed, configured, promoted, and used, is the creation of digitized sexually explicit material, and Defendants marketed and monetized that functionality for that use. Plaintiff is therefore entitled to the evidentiary presumption in § 1708.86(d)(1) that Defendants acted with the requisite knowledge or reckless disregard. Independently, and without regard to § 1708.86(a)(3), Defendants are liable under § 1708.86(b)(1), (b)(2), and (b)(3).

119. Defendants' liability arises, among other things, from their own generation, processing, transmission, facilitation, and knowing continuation of the technological functionality that generated and disclosed the material depicting Plaintiff.

120. Plaintiff suffered harm as a result of Defendants' violation of § 1708.86, including severe emotional distress, humiliation, anxiety, loss of privacy and autonomy, reputational and community harm, economic damages, and continuing fear of dissemination.

121. Defendants acted with malice within the meaning of California Civil Code § 1708.86(a)(10)(A), in that they engaged in despicable conduct. Defendants built and operated a system that removed clothing from photographs of identifiable women on demand; they were told, publicly and directly, that it was being used to strip religious attire from Muslim women

and to generate sexualized images of children; they responded with derision and with a paywall rather than a fix; and they continued to make the functionality available after Plaintiff reported the material depicting her. Plaintiff is therefore entitled to statutory damages of up to $250,000 under § 1708.86(f)(1)(B)(ii)(II).

122. Plaintiff is entitled to the remedies available under California Civil Code § 1708.86(f)(1), including: economic and noneconomic damages proximately caused by the disclosure, including damages for emotional distress, under § 1708.86(f)(1)(B)(i); statutory damages of not less than $1,500 and not more than $50,000 per violation under § 1708.86(f)(1)(B)(ii)(I), and, because Defendants acted with malice within the meaning of § 1708.86(a)(10)(A), statutory damages of up to $250,000 per violation under § 1708.86(f)(1)(B)(ii)(II); monetary gain or economic benefit Defendants received from the creation, development, or disclosure of the material under § 1708.86(f)(1)(A); punitive damages under § 1708.86(f)(1)(C); and reasonable attorney's fees and costs under § 1708.86(f)(1)(D), together with injunctive relief under § 1708.86(g)(1)(E). This action is timely under § 1708.86(g).

COUNT II – Strict Liability – Design Defect

123. Plaintiff re-alleges and incorporates by reference each preceding paragraph as though set forth fully at length herein. Count II is asserted against the Developer Defendants as the designers, developers, and manufacturers of Grok and its image-generation functionality. Plaintiff asserts Count II against the Platform Defendant in the alternative, and only to the extent alleged in the concluding paragraphs of this Count, on the ground that the Platform Defendant was a participant in the chain of distribution of that functionality.

124. At all relevant times, the Developer Defendants designed, developed, managed, operated, tested, produced, labeled, marketed, advertised, promoted, controlled, sold, supplied,

distributed, and benefitted from Grok's creation of AI-generated sexual content of Plaintiff.

125. Grok was designed and marketed as a generative artificial-intelligence system capable of generating and modifying photorealistic images and videos in response to user instructions.

126. The Developer Defendants commercially distributed Grok through applications, subscriptions, and other channels through which members of the public obtained access to its image-generation and image-editing functionality. On information and belief, the Platform Defendant separately distributed that same image-generation and image-editing functionality to the public by integrating it into the X platform and making it available to X users.

127. The Developer Defendants defectively designed Grok to generate sexualized depictions of identifiable, non-consenting real people from photographs of those people, without input filtering, identity detection, output nudity classification, or refusal mechanisms.

128. The alleged defect existed when the relevant functionality was placed into the stream of commerce and was not the result of a substantial subsequent alteration by Plaintiff or the users who generated the images at issue.

129. The relevant functionality was used in a manner that was reasonably foreseeable to the Developer Defendants, including use by members of the public to generate and modify images of identifiable persons.

130. The design was defective because the foreseeable risks associated with allowing users to generate sexually explicit depictions of identifiable persons could have been reduced or avoided through the adoption of alternative safety designs, including restrictions on the sexualized manipulation of identifiable persons, input filtering, output filtering, identity-sensitive safeguards, and other technical controls.

131. The design defect was not limited to sexualization in general. Grok was trained on large, uncurated image datasets that rendered religious and cultural attire highly legible to the model, and the Developer Defendants implemented no mechanism by which the model would refuse a prompt seeking the removal of such attire from an identifiable person. The foreseeable result, which occurred, was that the functionality could be turned on Muslim, Middle Eastern, and South Asian women as a class by any user who typed a single sentence. A safer alternative design, technically and economically feasible at the time of distribution, would have classified religious and cultural attire as a protected characteristic and refused such prompts outright.

132. The risks posed by the relevant design were substantial and outweighed the benefits of distributing the functionality without adequate safeguards against foreseeable nonconsensual sexualized manipulation of identifiable persons.

133. The foreseeable harm to Plaintiff was substantial and was a foreseeable consequence of permitting the relevant functionality to operate without adequate safeguards.

134. The Developer Defendants knew or should have known of the risk that users would employ image-generation and image-editing functionality to create nonconsensual sexualized depictions of identifiable individuals.

135. Plaintiff was injured as a substantial result of the alleged design defect.

136. Plaintiff seeks all damages and other relief recoverable under California law.

137. The conduct of the Defendants named in this Count was willful, malicious, oppressive, and/or in conscious disregard of the rights and safety of others, as further alleged herein, supporting punitive damages to the extent permitted by applicable law.

138. In the alternative, the Platform Defendant is strictly liable for the design defect alleged in this Count because it was a participant in the chain of distribution of Grok's image-

generation and image-editing functionality. California strict products liability is not limited to manufacturers; it extends to every commercial participant in the vertical chain of distribution, *see Vandermark v. Ford Motor Co*., 61 Cal. 2d 256, 262–63 (1964), including online platforms through which a defective product reaches the consumer, *see Bolger v. Amazon.com*, LLC, 53 Cal. App. 5th 431, 452 (2020); *Loomis v. Amazon.com LLC*, 63 Cal. App. 5th 466, 480 (2021).

139. The Platform Defendant also satisfies each factor governing strict liability for a non-manufacturing participant in the marketing and distribution enterprise. First, it received a direct financial benefit from making Grok's image-generation functionality available on X, including through subscription revenue, increased user engagement, and increased advertising inventory. Second, its role was integral to bringing that functionality to the consumer market, because the X integration was a principal channel through which members of the public accessed and used it. Third, it had control over, or at a minimum a substantial ability to influence, the distribution of that functionality, including whether to integrate it into X, on what terms, for which classes of users, and with what output-side safeguards. *See Bay Summit Community Assn. v. Shell Oil Co*., 51 Cal. App. 4th 762, 776 (1996); *Loomis*, *supra,* 63 Cal. App. 5th at 478.

140. Plaintiff does not allege in this Count that the Platform Defendant is liable for failing to remove content created by a third party. The defect alleged is in the design of the functionality and in the Platform Defendant's own decision to distribute it through a mass-distribution platform without output-side safeguards.

141. To the extent the Court concludes that the Platform Defendant was not a participant in the chain of distribution of Grok, Plaintiff's claims against the Platform Defendant proceed under Count I and Counts III through V, none of which depends on strict products liability.

COUNT III – Negligence

142. Plaintiff re-alleges and incorporates by reference each preceding paragraph as though set forth fully at length herein. Count III is asserted against the Developer Defendants for negligence in the design, development, training, testing, deployment, and operation of Grok, and against the Platform Defendant for negligence in its own separate conduct: its decision to integrate and distribute that functionality through the X platform, its design and operation of the reporting and disclosure interfaces on X, and its continued distribution and amplification of the material depicting Plaintiff after it had notice.

143. The Developer Defendants designed, developed, deployed, operated, maintained, marketed, and commercially provided Grok and its image-generation and image-editing capabilities.

144. The Developer Defendants' Grok is designed and intended to be a generative AI photo-and-image- generation tool.

145. The Developer Defendants and the Platform Defendant knew or reasonably should have known that systems capable of generating photorealistic images and substantially altering photographs of identifiable persons presented a foreseeable risk that users would generate nonconsensual sexualized depictions of real people.

146. That risk was particularly foreseeable because the Developer Defendants' systems were capable of processing photographs of identifiable persons and generating sexualized or "spicy" content, and because the Developer Defendants and the Platform Defendant received actual reports concerning the generation and dissemination of sexualized images of real persons, including minors.

147. The Developer Defendants owed Plaintiff a duty to exercise reasonable care in the design, development, deployment, monitoring, and operation of Grok to avoid creating

foreseeable and unreasonable risks of harm to identifiable persons whose images were processed by the system.

148. In determining the existence and scope of that duty, the foreseeable severity of the harm, Defendants' control over the relevant technology, the availability of reasonable safeguards, Defendants' knowledge of the risk, and the relationship between Defendants' conduct and the resulting harm all support recognition of a duty of reasonable care.

149. Reasonable safeguards were technologically and commercially feasible, including restrictions on sexualized manipulation of photographs of identifiable persons, input and output filtering, image-recognition systems, detection of known or likely nonconsensual sexualized content, effective reporting mechanisms, prompt removal procedures, and restrictions on repeated generation of substantially similar material.

150. The Developer Defendants breached their duty by, among other things, failing to implement and maintain reasonable safeguards against foreseeable generation and dissemination of nonconsensual sexualized depictions of identifiable persons; failing to adequately test and monitor those safeguards; continuing to provide the relevant functionality after receiving notice of its misuse; and failing to respond reasonably and effectively to Plaintiff's reports.

151. The Platform Defendant owed Plaintiff an independent duty of reasonable care arising from its own affirmative conduct. Having chosen to integrate into the X platform image-generation functionality that it knew produced sexualized depictions of identifiable real people, and having designed and operated the interfaces through which the resulting material was published, amplified, and reported, the Platform Defendant was required to exercise reasonable care in making and implementing those choices.

152. The Platform Defendant breached that duty by, among other things: integrating

Grok's image-generation and image-editing functionality into X without output-side safeguards against the sexualized manipulation of identifiable persons; making that functionality available to general public users through the @Grok account and through in-line prompts on posts depicting identifiable persons; designing a reporting interface that restricted the reporting user's own view of the material while leaving it publicly available to everyone else; and continuing to distribute and algorithmically amplify the material depicting Plaintiff after it had notice.

153. Defendants' conduct was a substantial factor in causing Plaintiff's injuries because the relevant images were generated through Defendants' systems and disseminated through functionality Defendants designed, controlled, operated, and continued to provide.

154. Plaintiff suffered severe emotional distress, anxiety, humiliation, loss of privacy and autonomy, reputational harm, economic damages, and other compensable injuries as a direct and proximate result of Defendants' negligence.

COUNT IV – Negligent Undertaking

155. Plaintiff re-alleges and incorporates by reference each preceding paragraph as though set forth fully at length herein. Count IV is asserted against the Developer Defendants based on their undertaking to design, operate, monitor, and maintain Grok safely, and against the Platform Defendant based on its separate undertaking, described below, to enforce its own published platform rules and to provide a functioning mechanism for reporting and addressing nonconsensual sexualized imagery on X.

156. The Developer Defendants undertook to design, develop, operate, monitor, and maintain Grok as a publicly available generative artificial intelligence system capable of generating and modifying images and videos into sexually explicit content.

157. The Developer Defendants and the Platform Defendant knew or should have recognized that effective safety measures were necessary to prevent Grok from being used to

create and distribute non-consensual sexually explicit material depicting identifiable individuals.

158. The Developer Defendants' undertaking increased the risk of harm to Plaintiff because the Developer Defendants designed and operated Grok with image-generation and image-editing capabilities that could be used to create sexually explicit depictions of real people without their consent while failing to implement safeguards adequate to prevent that foreseeable misuse.

159. The Developer Defendants and the Platform Defendant had actual and constructive notice of the risk. Among other things, they received reports concerning sexualized images of real people and were placed on notice that Grok could generate sexually explicit material depicting minors.

160. Defendants had the technological and financial means to implement reasonable safeguards, including input filtering, output filtering, image-recognition systems, restrictions on the manipulation of identifiable persons, effective reporting and removal procedures, and other measures capable of substantially reducing the foreseeable harm.

161. The Developer Defendants and the Platform Defendant nevertheless failed to exercise reasonable care in undertaking to provide and maintain those safety measures.

162. The Developer Defendants and the Platform Defendant breached their duty of undertaking by failing to use reasonable care to prevent the creation and distribution of AI-generated sexually explicit content.

163. The Platform Defendant separately undertook to render services necessary for the protection of Plaintiff and others similarly situated. Through its published platform rules prohibiting nonconsensual nudity and synthetic or manipulated sexualized media, and through the reporting mechanism it designed, operated, and held out to users as the means of addressing

such material, the Platform Defendant undertook to identify and act upon nonconsensual sexualized imagery on X. It should have recognized those services as necessary for the protection of the persons depicted, it failed to perform them with reasonable care, and its failure both increased the risk of harm to Plaintiff and caused harm because Plaintiff relied on that reporting mechanism in lieu of other protective measures. *See Artiglio v. Corning Inc.*, 18 Cal. 4th 604, 613–14 (1998); Rest.2d Torts, § 324A.

164. Defendants' failure to exercise reasonable care increased the risk of harm to Plaintiff and was a substantial factor in causing Plaintiff's injuries.

165. The conduct of Defendants, as described above, was intentional, willful, wanton, reckless, malicious, oppressive, extreme, and outrageous, and displayed a conscious and reckless disregard for the consequences of their conduct, including the foreseeable harm to women and children.

COUNT V – Intentional Infliction of Emotional Distress

166. Plaintiff re-alleges and incorporates by reference each preceding paragraph as though set forth fully at length herein. Count V is asserted against the Developer Defendants and the Platform Defendant, each for its own conduct as alleged below.

167. Defendants engaged in conduct that was extreme and outrageous because, despite knowledge of the foreseeable risk that the image-generation and image-editing functionality could be used to create sexually explicit depictions of identifiable persons without consent, the Developer Defendants continued to design, operate, and commercially provide that functionality without adequate safeguards, and the Platform Defendant continued to make it available on X and to distribute and amplify the resulting material after receiving notice of the material depicting Plaintiff.

168. The Platform Defendant’s conduct was independently extreme and outrageous.

With actual knowledge that the functionality it had integrated was being used to generate nonconsensual sexualized depictions of identifiable women, it continued to make that functionality available to the general public on X, continued to distribute and algorithmically amplify the resulting material, and continued to derive revenue and engagement from it. That material was generated by a tool the Platform Defendant integrated and offered as its own product feature.

169. The conduct of each Defendant named in this Count was further outrageous because the material depicting Plaintiff exploited her religious identity. Removing Plaintiff's hijab and depicting her nude or nearly nude converted the depiction into an attack on her faith, her culture, and her standing in the communities she serves. Defendants continued to operate the functionality with knowledge that it was being used for that purpose against Muslim, Middle Eastern, and South Asian women as a class, and with knowledge that the material depicting Plaintiff remained accessible after she reported it.

170. Defendants knew that their conduct created a substantial probability of causing severe emotional distress to persons whose identities or likenesses were sexually exploited through the relevant functionality, or acted with reckless disregard of that probability.

171. Plaintiff was directly and foreseeably subjected to the conduct alleged herein.

172. Defendants' conduct was a substantial factor in causing Plaintiff severe emotional distress.

173. Plaintiff has suffered severe emotional distress, including anxiety, humiliation, shame, fear of continued dissemination, impairment of ordinary activities, and other substantial emotional and psychological harm.

VI. PRAYER FOR RELIEF

174. WHEREFORE, Plaintiff requests that this Court award relief against

Defendants as follows:

175. On Count I, awarding Plaintiff, at her election, actual and noneconomic damages under California Civil Code § 1708.86(f)(1)(B)(i) or statutory damages of not less than $1,500 and not more than $50,000 per violation under § 1708.86(f)(1)(B)(ii)(I), and, because Defendants acted with malice, statutory damages of up to $250,000 per violation under § 1708.86(f)(1)(B)(ii)(II); disgorgement of Defendants' monetary gain or economic benefit under § 1708.86(f)(1)(A); punitive damages under § 1708.86(f)(1)(C); and reasonable attorney's fees and costs under § 1708.86(f)(1)(D);

176. Entering a permanent injunction under California Civil Code § 1708.86(f)(1)(E) and Federal Rule of Civil Procedure 65(d) requiring Defendants, and their officers, agents, servants, employees, and attorneys, to: (i) permanently delete and cease all disclosure of every digitized sexually explicit image or video depicting Plaintiff that is within Defendants' possession, custody, or control, including all copies stored on Defendants' servers, in training corpora, in model caches, and in any third-party repository under Defendants' control; (ii) implement and maintain output-side classifiers and prompt filters that block the generation of nude or sexualized depictions of Plaintiff and of any identifiable real person who has not provided written consent; (iii) refuse any prompt that requests the removal of clothing from, or the sexualization of, an uploaded or referenced image of an identifiable real person; (iv) maintain a notice-and-removal process that removes reported digitized sexually explicit material within seventy-two hours of written notice and confirms removal in writing to the reporting individual; and (v) certify compliance to the Court and to Plaintiff's counsel in writing within thirty days of entry of judgment and every ninety days thereafter for three years;

177. On Counts II, III, and IV, awarding Plaintiff compensatory damages in an amount to be proven at trial, including damages for past and future emotional distress,

humiliation, reputational harm, loss of privacy, medical and mental-health expenses, and lost earnings and earning capacity;

178. On Count V, awarding punitive and exemplary damages under California Civil Code § 3294(a), the malice, oppression, or fraud alleged herein having been committed, authorized, or ratified by an officer, director, or managing agent of each corporate Defendant within the meaning of § 3294(b);

179. Awarding Plaintiff pre-judgment and post-judgment interest pursuant to the claims for relief;

180. Awarding Plaintiff costs, expenses, and attorneys' fees as permitted by law; and

181. Awarding Plaintiff further relief as the Court may deem just and proper under the circumstances.

JURY DEMAND

182. Pursuant to Rule 38 of the Federal Rules of Civil Procedure, Plaintiff hereby demands a jury trial for all claims so triable.

Dated: November 18, 2026

By: */s/ Alexandra Marsh*
Alexandra Marsh (State Bar No. 357534)

Hillary M. Nappi (*pro hac vice forthcoming*)
AWK Survivor Advocate Attorneys PLC
1133 Westchester Avenue, N 224
White Plains, NY 10604
(914) 468-4840
Hnappi@awk-saa.com

Alexandra Marsh (State Bar No. 357534)
AWKO Law, LLP
2193 Fillmore Street
San Francisco, CA 94115
Telephone: 415-322-3311

Facsimile: 850-916-7449
Email: AMarsh@awkolawllp.com
Service Email: sateam@awkolaw.com

Attorneys for Plaintiff