

1 William J. Edelman (SBN 285177)
2 **MILBERG, PLLC**
3 227 W. Monroe Street, Suite 2100
4 Chicago, IL 60606
5 Tel: 866.252.0878
6 wedelman@milberg.com

7 *Attorneys for Plaintiff and Proposed Class*
8 *(Additional Counsel on Signature Page)*

9 **UNITED STATES DISTRICT COURT FOR THE**
10 **NORTHERN DISTRICT OF CALIFORNIA**

11 Devin Youngblood, on behalf of himself and all
12 others similarly situated,

13 Plaintiff,

14 v.

15 NVIDIA CORPORATION.,

16 Defendant.

Case No.

CLASS ACTION COMPLAINT FOR:

- 1. **VIOLATIONS OF DMCA, 17 U.S.C. § 1201(a)**

JURY TRIAL DEMANDED

1 Plaintiff, Devin Youngblood, (“Plaintiff”) bring this class action complaint (“Complaint”) on
2 behalf of himself and all others similarly situated (the “Class Members”) against NVIDIA Corporation,
3 (“NVIDIA” or “Defendant”) for violations of the Digital Millennium Copyright Act (“DMCA”), 17
4 U.S.C. § 1201. The allegations contained herein, which are based on Plaintiff’s knowledge of facts
5 pertaining to himself and his own actions and counsels’ investigations, and upon information and belief
6 as to all other matters, are as follows:

7 **NATURE OF THE ACTION**

8 1. This is a nationwide class action for violations of the DMCA’s anti-circumvention
9 provisions, 17 U.S.C. § 1201(a), arising from NVIDIA’s unauthorized acquisition—through scraping,
10 bulk downloading, and other extraction methods—of various content creators’ YouTube-hosted
11 audiovisual works (“Works”) at the file level by bypassing YouTube’s technological measures that control
12 access to the Works and NVIDIA’s use of those unlawfully accessed Works to train, develop, and improve
13 NVIDIA’s Omniverse 3D world generator, self-driving cars, and “digital human” artificial intelligence
14 (“AI”) products, internally referred to by NVIDIA as “Cosmos.”

15 2. YouTube makes creators’ videos available to the public primarily through controlled
16 streaming pathways and related authorized features. YouTube also employs technical measures designed
17 to control access to the underlying audiovisual files and to prevent bulk extraction by automated systems.

18 3. YouTube’s technical controls are designed to provide access through user-based streaming
19 and to prevent non-authorized bulk extraction. At the scale required for modern AI video-model training,
20 like those needed by NVIDIA, obtaining file-level copies of the Works necessarily requires bypassing
21 those controls.

22 4. NVIDIA did not obtain licenses or permissions from Plaintiff and Class Members for the
23 file-level acquisition and use of their Works as training inputs for its AI models. Instead, on information
24 and belief, including information from NVIDIA employees recently uncovered through media sources,
25 NVIDIA used or relied on circumvention tools and services to defeat YouTube’s technological measures
26 at scale.

1 5. The DMCA prohibits circumventing a technological measure that effectively controls
2 access to a copyrighted work, and it prohibits trafficking in circumvention technologies and services. 17
3 U.S.C. § 1201(a). Plaintiff seeks injunctive relief, statutory damages, and other remedies authorized by 17
4 U.S.C. § 1203.

5 **JURISDICTION AND VENUE**

6 6. This Court has subject matter jurisdiction under 28 U.S.C. § 1332(d) (the Class Action
7 Fairness Act) because the amount in controversy exceeds \$5,000,000, exclusive of interest and costs, and
8 a member of the Class is a citizen of a different state than NVIDIA. This Court also has subject matter
9 jurisdiction under 28 U.S.C. § 1331 because this action arises under 17 U.S.C. § 1201, et seq. (the Digital
10 Millenium Copyrights Act (“DMCA”)).

11 7. This Court has personal jurisdiction over Defendant because NVIDIA has its headquarters
12 in California, transacts business nationwide, and committed the acts complained of in this District and/or
13 directed them at this District, causing harm here.

14 8. Venue is proper in this District because NVIDIA resides in this District and/or a
15 substantial part of the events or omissions giving rise to the claims occurred in this District.

16 9. Plaintiff reserves the right to refine jurisdiction and venue allegations after discovery and
17 as additional facts become available.

18 **THE PARTIES**

19 10. Plaintiff Devin Youngblood, is a resident of Alabama, who owns and operates the
20 YouTube channel “The Youngblood’s Podcast.” The Youngbloods Podcast is a YouTube channel and
21 podcast operated by Devin Youngblood and his wife. Through the channel, Plaintiff publishes weekly
22 audio and video episodes in which they respond to listener-submitted questions and discuss topics
23 including marriage, parenting, family relationships, mental health, and personal experiences drawn from
24 their own lives. The content is presented as candid; unscripted commentary based on the hosts’ personal
25 perspectives and family experiences and is distributed via YouTube.

26 11. At least one video from “The Youngblood’s Podcast” appears in the dataset HowTo100M,
27 which Defendant used to train its Cosmos model.
28

1 12. Plaintiff and Class Members derive value from their Works through viewership,
2 advertising, sponsorships, licensing, and other monetization that depends on controlled distribution and
3 the ability to choose whether, when, and how their Works will be licensed for downstream uses, including
4 AI model training.

5 13. Plaintiff and the Class Members are independent content creators of audiovisual content
6 hosted on YouTube.

7 14. Defendant, NVIDIA Corporation, is a Delaware corporation with its principal place of
8 business at 2788 San Tomas Expressway, Santa Clara, California 95051.

9 15. NVIDIA develops and commercializes artificial intelligence models and related products
10 and services, including an AI “world model” systems that rely on large-scale video training data.

11 **FACTUAL ALLEGATIONS**

12 **A. YouTube’s Access-Controlled Distribution of Creator Works**

13 16. YouTube is a video-sharing platform on which creators upload original audiovisual works.
14 Those works are protected by copyright, and creators retain rights and control over licensing and other
15 downstream uses.

16 17. YouTube provides access to videos primarily through controlled streaming. YouTube’s
17 design is intended to enable real-time viewing and limited download features available to those users who
18 have purchased a “Premium” license from YouTube.

19 18. YouTube employs technological protective measures (“TPMs”) to limit file-level
20 acquisition and bulk extraction of underlying audiovisual files.

21 19. YouTube’s Terms of Service expressly bar scraping, unauthorized downloading, bulk
22 extraction, and other forms of data mining of audiovisual content except through expressly permitted
23 product features or licensed APIs. Those contractual limits operate in tandem with YouTube’s TPMs to
24 prevent unlicensed, file-level access to creators’ Works and to ensure that content is accessed only through
25 YouTube’s controlled platform.

1 20. Indeed, YouTube has publicly stated that if AI companies are scraping videos hosted on its
2 website for use in AI training models, it would be considered a “clear violation” of YouTube terms of
3 service.¹

4 21. Content creators (including Plaintiff) who upload videos to YouTube grant YouTube and
5 other users limited permissions to access that content through YouTube’s services. But that limited license
6 makes clear that it “does not grant any rights or permissions for a user to make use of [the] Content
7 independent of the Service.” In other words, end users are permitted to view or stream videos within
8 YouTube’s environment, not to obtain or use the underlying audiovisual files outside YouTube.
9 YouTube’s Terms of Service thus reflect YouTube’s intent to restrict access to the digital files underlying
10 the Works.

11 22. Streaming a video through YouTube and downloading a permanent copy offer materially
12 different value propositions—free viewing supported by advertisements versus possession of a lasting
13 digital copy. Scraping or bulk downloading is therefore not simply copying material that YouTube
14 provides to the public; it is an act of unauthorized access to audiovisual data files that YouTube
15 affirmatively withholds from public download.

16 23. Although YouTube offers certain download functionality to paying subscribers (e.g.,
17 YouTube “Premium”), YouTube’s TPMs restrict that functionality in ways that prevent users from
18 obtaining the underlying audiovisual files. YouTube prohibits downloading audiovisual content except
19 through the YouTube application, and any “download” makes content available only for offline streaming
20 (without internet access) within the app—not as a transferable, permanent file copy. The offline access is
21 also time-limited, after which the content reverts to online streaming only.

22 24. To enforce these prohibitions, YouTube uses technological processes to identify and
23 prevent unauthorized extraction. For instance, YouTube periodically updates APIs in ways that disrupt
24 non-authorized downloaders relying on outdated interfaces to extract files. YouTube also monitors

25
26 ¹ Paolo Confino, OpenAI Could Be in a “Clear Violation” of YouTube’s Terms of Service, CEO Says—
27 Depending on How It Trains Its Sora Video Tool, *Fortune* (Apr. 4, 2024), available at:
28 <https://fortune.com/2024/04/04/openai-youtube-clear-violation-terms-service-ai-sora-training/> (last visited Jan.
22, 2026).

1 downloading behavior and may throttle, restrict, or block IP addresses associated with excessive download
2 attempts over a specified period.

3 25. YouTube’s TPMs are actively maintained and updated to prevent automated download
4 tools from reliably reconstructing full video files outside YouTube’s player-based environment.

5 26. In addition, YouTube maintains a framework based on express permission that allows
6 creators to decide whether third-party AI companies may use their Works for AI model training.

7 27. Under YouTube’s policy and settings, creators may permit third-party training only by
8 turning on a “third-party training” setting and selecting which third-party companies, if any, are allowed.

9 28. By default, YouTube’s third-party training setting is turned off—meaning creators do not
10 need to take any action to withhold permission, and third-party training is not authorized unless the creator
11 affirmatively opts in.

12 29. Even when a creator chooses to opt in, YouTube requires the creator to make an affirmative
13 choice: either (a) select one or more companies from YouTube’s list of participating third-party AI
14 companies, or (b) choose an “All third-party companies” option that grants permission broadly.

15 30. In other words, YouTube does not treat creators as having granted permission to any third-
16 party AI company unless and until the creator affirmatively authorizes such use.

17 31. YouTube’s list of third-party companies presented to creators includes major AI
18 developers, including NVIDIA—reflecting that YouTube contemplated that third-party AI developers
19 would require explicit creator permission before using YouTube videos for training.

20 32. Plaintiff and Class Members, who are YouTube content creators, post their original works
21 to YouTube and rely on YouTube’s controlled-access architecture to protect the value of their Works and
22 to preserve licensing markets, including the emerging market for AI training licenses.

23 33. Although most YouTube videos are not registered with the U.S. Copyright Office, they are
24 nevertheless copyrighted if they are original creations. Creators pour real time, skill, and money into
25 producing their videos, and they depend on YouTube’s technological safeguards to keep the underlying
26 files from being accessed and downloaded without permission. And because a DMCA circumvention
27 claim does not hinge on copyright registration, these protections matter most in exactly this situation—
28

1 where Defendant’s misconduct is breaking through access barriers that are designed to prevent anyone
2 from obtaining the underlying video files.

3 **B. NVIDIA’s Scale, Market Power, and Dominance in AI Infrastructure**

4 34. NVIDIA is a publicly traded technology company listed on the Nasdaq under the ticker
5 “NVDA.”

6 35. NVIDIA operates on an enormous scale. As of January 29, 2026, public market data
7 showed NVIDIA’s market capitalization exceeded \$4.5 trillion, making it one of the most valuable
8 companies in the world.

9 36. NVIDIA markets itself as a “platform” company that combines processors, networking,
10 systems, software, and services, and it states that it targets four main markets: Data Center, Gaming,
11 Professional Visualization, and Automotive.

12 37. NVIDIA also sells AI software products and services, including Omniverse Enterprise,
13 which NVIDIA markets for industrial AI usage. NVIDIA has long presented its hardware and software
14 ecosystem as core infrastructure for modern AI.

15 38. NVIDIA’s reported financial results reflect its outsized role in AI infrastructure. NVIDIA
16 reported \$130.5 billion in total revenue for fiscal year 2025. Approximately 88% of that total revenue
17 (\$115.2 billion) from its Data Center business—reflecting that NVIDIA’s core business is supplying AI
18 and accelerated-computing infrastructure to data centers.

19 39. NVIDIA is a dominant, multi-trillion-dollar company whose primary business lines are
20 built and marketed to support AI development and deployment.

21 **C. Defendant Improperly Obtained Millions of YouTube Videos to Train Its Foundational AI**
22 **Model**

23 40. Cosmos was built as Defendant’s foundational video model, intended to function as a
24 shared “core layer” supporting Defendant’s broader products, including GROOT, Avatar, OV, and
25 GeForce.

26 41. Because Cosmos was intended to serve as a foundational layer, Defendant had a powerful
27 commercial incentive to acquire training data at extraordinary scale. Instead of negotiating licenses for the
28

1 underlying works, Defendant set out to obtain massive datasets by defeating YouTube’s TPMs and access
2 restrictions and extracting the underlying video files.

3 42. Defendant needed vast amounts of training data, so Defendant’s employees coordinated
4 efforts to assemble these datasets, including internal communications on Slack in a channel called
5 “#cosmos-dataset-creation,” which was expressly used to organize the creation of a very large, curated
6 video dataset for AI training.

7 43. Defendant gathered datasets from multiple sources, including academic repositories and
8 research compilations developed by universities, companies, and independent researchers. Defendant
9 treated these datasets as inputs for Cosmos even where the datasets themselves were restricted to academic
10 or non-commercial purposes.

11 44. In or about February 2024, public reporting based on internal NVIDIA Slack
12 communications reflects that a Principal Scientist at NVIDIA, posted in an internal Slack channel that
13 NVIDIA was pursuing an “initiative to aggregate a huge curated video dataset for the purposes of
14 generative modelling,” and that the team’s initial goal was to “compile all the video datasets (public or
15 internal) that are available internally and have been downloaded in order to not multiply / repeat the
16 effort.”

17 45. Public reporting further reflects that NVIDIA employees circulated a spreadsheet intended
18 to catalog and centralize video datasets for this initiative totaling more than 40 datasets including the
19 “public” datasets HD-VG-130M, HDVILA-100M, and HowTo100M. Each of these datasets is built from
20 YouTube source material.

21 46. These datasets do not include actual video files; they are simply lists containing pointers
22 (such as URLs or YouTube Video IDs). The dataset itself merely identifies where the videos are located;
23 the user performs the downloading and copying.

24 47. To use these datasets for AI training, a user (here Defendant) must obtain the actual
25 underlying audiovisual files by downloading them from YouTube, which requires copying the works and
26 bypassing YouTube’s TPMs, contractual terms, and licensing limits.

1 48. These datasets are structured around “clips,” not full videos. They identify timestamps
2 within YouTube videos and treat each timestamped segment as a separate training example. That design
3 may require repeated retrieval of the same underlying YouTube video to extract different clips, resulting
4 in multiple copies of the same copyrighted work in order to capture the different start and end points
5 identified in the dataset. Each retrieval necessarily entails evading YouTube’s TPMs, contractual limits,
6 and technological controls. Defendant’s downloading, copying, and circumvention occurred clip-by-clip,
7 and each instance constitutes a separate violation.

8 49. The first dataset, HD-VG-130M, contains pointers to more than 130 million clips created
9 by slicing longer YouTube videos into short segments for machine-learning research. Each clip may
10 require its own download to obtain the underlying audiovisual file, but at the very least each video in the
11 dataset needs to be downloaded. HD-VG-130M is compiled from 1,549,408 YouTube videos and divided
12 into approximately 130 million high-definition clips.

13 50. HD-VG-130M intentionally includes videos with high “aesthetic” scores (as measured by
14 another AI model) and excludes videos displaying creator names or visible watermarks. The dataset was
15 assembled for academic research in 2023 and was not licensed for commercial use.

16 51. Defendant used HD-VG-130M to retrieve the referenced clips and incorporate the resulting
17 files into Defendant’s commercial training pipeline for Cosmos.

18 52. Microsoft created documentation identifying the contents of the HD-VG-130M dataset and
19 released that dataset into the broader AI ecosystem, including by posting it publicly on GitHub.

20 53. The GitHub repository for HD-VG-130M includes licensing terms restricting the dataset
21 to academic use only, expressly recognizing that commercial use would be unauthorized.

22 54. The licensing terms further restrict redistribution, reproduction, modification, or
23 exploitation absent permission from the copyright owner of the underlying audiovisual content—thereby
24 acknowledging that copyright ownership resides with the creators of the underlying source works,
25 including YouTube creators like Plaintiff and Class Members.

1 55. Defendant used HD-VG-130M in commercial development and training of its generative
2 model despite its “academic only” limitations and Defendant knowing the dataset was designated for
3 research purposes only.

4 56. Defendant also used HD-VG-130M as a roadmap to scrape audiovisual content from
5 YouTube without the consent of YouTube or the creators of that content.

6 57. The second dataset, HDVILA-100M, is likewise comprised entirely of YouTube videos
7 and contains 3,098,462 source videos divided into roughly 100 million clips. The dataset was compiled
8 by Microsoft Research Asia and published in 2021 for research in video-based AI.

9 58. The HDVILA-100M license (which is a license only to the dataset compilation and not the
10 underlying material) limits use to computational, non-commercial research, and forbids using the data or
11 resulting outputs in any commercial product or service.

12 59. HDVILA-100M does not include the actual underlying video files. Any user of the dataset
13 must download the referenced clips directly from YouTube. Defendant did so at scale.

14 60. By including these restrictions, Microsoft implicitly acknowledged that HDVILA-100M
15 draws from protected works and does not provide commercial authorization to exploit the underlying
16 videos.

17 61. Defendant obtained and used HDVILA-100M despite its non-commercial limitations.
18 Defendant knew the dataset was for research but used it for commercial training and development of its
19 generative model and used it to scrape YouTube content without consent from YouTube or creators.

20 62. The third dataset, HowTo100M, consists of 1,238,911 YouTube videos split into more than
21 100 million clips depicting tens of thousands of tasks (including cooking, crafts, personal care, gardening,
22 and other activities). The dataset is hosted by European research institutions and includes identifiers and
23 metadata—again, not underlying audiovisual files.

24 63. To use HowTo100M for training, Defendant had to download the referenced clips directly
25 from YouTube.

26 64. HDVILA-100M, HD-VG-130M, and HowTo100M are fundamentally lists of identifiers
27 pointing to videos (or segments of videos) on YouTube. Defendant used these datasets to drive mass
28

1 downloading of protected YouTube content—without authorization, in violation of access restrictions,
2 and for the commercial purpose of building Cosmos.

3 **D. NVIDIA Built and Deployed a Mass “Download-and-Ingest” Pipeline to Evade YouTube’s**
4 **TPMs and Extract Underlying Video Files for Cosmos Training**

5 65. On information and belief, Defendant—directly and through employees, agents,
6 contractors, or affiliates—used tools and workflows designed to bypass or evade YouTube’s usage
7 restrictions and TPMs, copied videos, and assembled training material for Defendant’s AI systems and
8 services.

9 66. NVIDIA’s campaign included 20 to 30 virtual machines hosted by Amazon Web Services
10 (AWS) capable of downloading 80 years-worth of videos per day. A VP of Research at NVIDIA and a
11 Cosmos project leader stated in internal emails that NVIDIA was “finalizing the v1 data pipeline and
12 securing the necessary computing resources to build a video data factory that can yield a human lifetime
13 visual experience worth of training data per day.”

14 67. Defendant was not viewing the videos or even downloading content in the limited manner
15 contemplated by YouTube Premium’s offline-streaming feature. Instead, Defendant extracted and stored
16 the underlying audiovisual files within Defendant’s own systems, where Defendant could control and
17 retain them indefinitely which is inconsistent with the Premium model and YouTube’s TPM restrictions.

18 68. Public reporting based on leaked internal NVIDIA communications shows that NVIDIA
19 personnel understood the YouTube-based “datasets” at issue (including HDVILA-100M, HowTo100M
20 ,and HD-VG-130M) function as pointers (URLs/YouTube IDs) and that, to use them for training, NVIDIA
21 had to retrieve and download the underlying YouTube video files at scale.

22 69. In internal Slack communications concerning the Cosmos dataset effort, NVIDIA
23 engineers discussed datasets they had already “ingested,” expressly including HD-VG-130M.

24 70. The same reporting explains that HD-VG-130M’s own usage terms are “academic use
25 only,” prohibiting commercial exploitation and distribution—restrictions NVIDIA personnel nevertheless
26 treated as “fair game” for Cosmos.

1 71. With respect to HDVILA-100M, leaked internal chat logs reflect that NVIDIA personnel
2 identified that the version of HDVILA-100M they possessed was missing approximately 2.3 million “raw
3 videos,” and they circulated a document containing the corresponding “missing youtube links” and
4 directed that those links be placed into a “downloading pipeline.”

5 72. These chat logs further reflect that NVIDIA engineers were not merely collecting links;
6 they were building infrastructure to systematically download the underlying YouTube video files,
7 including by creating a database of URLs already downloaded and using YouTube’s unique video
8 identifiers as references to support repeated downloading.

9 73. Defendant’s engineers discussed building a central database of YouTube URLs already
10 downloaded so the team could repeatedly reprocess the same underlying YouTube sources without losing
11 track of what had been taken.

12 74. In March 2025, one engineer proposed: “Let’s have a database of urls we already
13 downloaded,” explaining that “YouTube video has unique id . . . we can use those as reference (the id
14 after “?v=”),” because the team expected ongoing “url diff and merging many times going forward.”
15 Defendant’s personnel confirmed that infrastructure was being created “literally right now” using Hive.

16 75. The same internal discussions show that NVIDIA personnel anticipated and planned for
17 YouTube’s anti-scraping defenses, including the risk of YouTube blocking IP addresses used for mass
18 downloading.

19 76. NVIDIA personnel discussed using IP rotation and stated that, because they were on AWS,
20 restarting a virtual machine would provide a new public IP—i.e., a method to avoid blocks while
21 continuing bulk downloading of videos in the datasets.

22 77. Defendant executed automated downloading designed to extract the underlying
23 audiovisual files (not merely “view” them through YouTube’s interface), including through use of the
24 open-source YouTube downloader “yt-dlp.”

25 78. Yt-dlp can be used to download individual videos or entire playlists by inputting the
26 YouTube URL(s) and can also merge separate audio/video streams to assemble complete audiovisual files.

1 79. NVIDIA personnel also recognized legal constraints associated with YouTube-based
2 datasets and raised concerns internally about whether certain YouTube ID datasets could be used for non-
3 research purposes.

4 80. When an employee specifically asked whether NVIDIA had legal approval to download
5 YouTube videos from a YouTube-ID dataset, management responded that the activity was an “executive
6 decision” and that the team had “umbrella approval” to use “all of the data.”

7 81. The same leaked internal communications also corroborate unlawful access and
8 downloading activity as to the HowTo100M dataset.

9 82. HowTo100M is widely described as an instructional “how-to” dataset derived from
10 YouTube videos and Defendant’s internal Slack messages show a focus on acquiring “tutorial” and other
11 instructional categories of videos, including requests for “tutorial videos.”

12 83. Defendant’s team was actively curating categories typical of the HowTo100M dataset—
13 i.e., instructional and task-based content—when planning further YouTube downloading. For instance, as
14 team members traded suggestions for YouTube channels “that we could download,” a project manager
15 asked whether the suggestions included “tutorial videos,” along with other categories (e.g., “Astronomy,”
16 “Medical”).

17 84. Although there is a purported “commercial license” attached to the HowTo100M dataset,
18 the dataset’s compilers do not own the underlying YouTube videos and cannot authorize third parties to
19 reproduce, download, or commercially exploit those works.

20 85. These leaked chat logs and related internal communications reflect that NVIDIA
21 operationalized a download-and-ingest pipeline designed to extract and store the underlying audiovisual
22 files from YouTube at scale, through means that circumvented YouTube’s TPMs.

23 86. Defendant combined tools like yt-dlp with tactics, such as virtual machines used to refresh
24 IP addresses, to bypass YouTube’s TPMs and avoid being blocked thereby enabling Defendant to mass-
25 download YouTube content and feed the resulting files directly into its training pipeline for Cosmos.

1 95. Numerosity/Ascertainability. Members of the Class are so numerous that joinder of all
2 members would be unfeasible and not practicable. The exact number of Class Members is unknown to
3 Plaintiff currently. However, it is estimated that there are at least thousands of individuals in the Class.
4 The identity of such membership is readily ascertainable from Defendants’ records and non-party
5 YouTube’s records.

6 96. Typicality. Plaintiff’s claims are typical of the claims of the Class because Plaintiff’s
7 Works were included in the datasets used by NVIDIA in the training of NVIDIA’s AI models. NVIDIA
8 unlawfully accessed and downloaded the Works through the use of technological measures aimed at
9 circumventing YouTube’s TPMs. Plaintiff’s claims are based on the same legal theories as the claims of
10 other Class Members.

11 97. Adequacy. Plaintiff is fully prepared to take all necessary steps to represent fairly and
12 adequately the interests of the Class Members. Plaintiff’s interests coincide with, and not antagonistic to,
13 those of the Class Members. Plaintiff is represented by attorneys with experience in the prosecution of
14 class action litigation. Plaintiff’s attorneys are committed to vigorously prosecuting this action on behalf
15 of the Class Members.

16 98. Common Questions of Law and Fact Predominate. Questions of law and fact common to
17 the Class Members predominate over questions that may affect only individual Class Members because
18 Defendant has acted on grounds generally applicable to the Class. Such generally applicable conduct is
19 inherent in Defendant’s wrongful conduct. The following questions of law and fact are common to the
20 Class:

21 (a) Whether YouTube employs “technological measures that effectively control
22 access” to copyrighted audiovisual works within the meaning of 17 U.S.C. § 1201(a).

23 (b) Whether NVIDIA accessed YouTube-hosted videos at the file level to obtain
24 training-ready copies, clips, frames, or other stored representations of Class Members’ Works.

25 (c) Whether NVIDIA retrieved, scraped, downloaded, or otherwise acquired the
26 underlying YouTube videos referenced in HDVILA-100M, HowTo100M, and HD-VG-130M and any
27 other datasets used by NVIDIA at massive scale to build training material for Cosmos and related systems.
28

1 (d) Whether NVIDIA’s methods for retrieving YouTube videos defeated, avoided,
2 bypassed, removed, deactivated, impaired, or otherwise circumvented YouTube’s TPMs.

3 (e) Whether NVIDIA thereby violated 17 U.S.C. § 1201(a)(1) by circumventing
4 technological measures controlling access to Class Members’ Works.

5 (f) Whether NVIDIA’s circumvention and/or trafficking conduct was willful,
6 knowing, and undertaken for commercial advantage or private financial gain.

7 (g) Whether NVIDIA’s conduct caused Class wide injury by depriving creators and
8 rights-holders of control over access to and use of their Works and by enabling NVIDIA’s ongoing
9 possession and use of unlawfully accessed copies within its training pipelines.

10 (h) Whether Plaintiff and Class Members are entitled to declaratory and injunctive
11 relief requiring NVIDIA to cease circumvention-based access to YouTube-hosted Works and to
12 implement compliance measures sufficient to prevent further circumvention.

13 (i) Whether Plaintiff and Class Members are entitled to statutory damages (including
14 the appropriate statutory measure and amount) under 17 U.S.C. § 1203(c) based on NVIDIA’s violations
15 of 17 U.S.C. § 1201.

16 99. Superiority. Class action treatment is a superior method for the fair and efficient
17 adjudication of the controversy. Such treatment will permit many similarly situated persons to prosecute
18 their common claims in a single forum simultaneously, efficiently, and without the unnecessary
19 duplication of evidence, effort, or expense that numerous individual actions would engender. The benefits
20 of proceeding through the class mechanism, including providing injured persons a method for obtaining
21 redress on claims that could not practicably be pursued individually, substantially outweighs potential
22 difficulties in management of this class action. Plaintiff is unaware of any special difficulty to be
23 encountered in litigating this action that would preclude its maintenance as a class action.

24 **CLAIM FOR RELIEF**

25 **VIOLATION OF 17 U.S.C. § 1201(a)**

26 **(DMCA ANTI-CIRCUMVENTION)**

27 100. Plaintiff repeats and realleges the foregoing allegations as if fully set forth herein.

1 101. YouTube prohibits scraping, unauthorized downloading, bulk extraction, and other forms
2 of automated retrieval of audiovisual content through access controls to prevent unlicensed, file-level
3 access to the underlying video files.

4 102. Plaintiff and Class Members uploaded and distributed their original Works on YouTube in
5 substantial part because YouTube protects against third parties engaging in bulk downloading or
6 automated extraction of creators' videos outside YouTube's controlled streaming and download pathways.

7 103. Content creators, including Plaintiff and Class Members, reasonably expect that their
8 Works will be accessed only through YouTube's authorized viewing and distribution mechanisms, and
9 will not be accessed at the file level, scraped, downloaded, and copied at massive scale without consent.

10 104. YouTube's TPMs constitute "effective technological measures" within the meaning of 17
11 U.S.C. § 1201 because those measures, in the ordinary course of operation, control access to copyrighted
12 audiovisual works and prevent third parties from obtaining file-level copies at scale without defeating
13 those controls.

14 105. NVIDIA used and/or relied on automated tools, services, and methods designed for the
15 purpose of bypassing and circumventing YouTube's TPMs in order to retrieve, extract, and reconstruct
16 underlying video files that YouTube does not make available for bulk download to the public, and, through
17 that circumvention-based harvesting, NVIDIA improperly obtained millions of YouTube videos.

18 106. NVIDIA accessed, downloaded, stored, and processed unlawfully obtained YouTube
19 videos to assemble training materials used to develop, train, and improve NVIDIA's AI systems, including
20 NVIDIA's Cosmos model and related services.

21 107. NVIDIA's violations were willful because, among other reasons, NVIDIA is a
22 sophisticated technology company that understood YouTube's TPMs, yet knowingly used or relied on
23 automated circumvention methods designed to evade YouTube's protections in order to obtain AI training
24 data at massive scale.

25 108. Each act of circumvention constitutes a separate violation of 17 U.S.C. § 1201. Plaintiff
26 and Class Members are therefore entitled to statutory damages, injunctive relief, and attorneys' fees and
27 costs pursuant to 17 U.S.C. § 1203.
28

RELIEF REQUESTED

Plaintiff, on behalf of himself and the proposed Class, respectfully requests that the Court grant the following relief:

(a) Certification of this action as a class action and appointment of Plaintiff and Plaintiff's counsel to represent the Class.

(b) Declare that NVIDIA violated 17 U.S.C. § 1201(a) by circumventing YouTube's TPMs controlling access to Plaintiff's and Class Members' Works.

(c) Enter injunctive relief requiring NVIDIA to cease circumvention-based access to YouTube-hosted Works and to implement compliance measures sufficient to prevent further circumvention;

(d) Award statutory damages and/or other damages as authorized by 17 U.S.C. § 1203.

(e) Award Plaintiff and the Class their reasonable attorneys' fees and costs pursuant to 17 U.S.C. § 1203(b)(5); and

(f) Award such other and further relief as the Court deems just and proper.

DEMAND FOR JURY TRIAL

Plaintiff, individually and on behalf of the proposed Class, demands a trial by jury for all the claims asserted in this Complaint so triable.

Date: January 29, 2026,

Respectfully submitted,

/s/ William J. Edelman
William J. Edelman (SBN 285177)
MILBERG, PLLC
227 W. Monroe Street, Suite 2100
Chicago, IL 60606
Tel: 866.252.0878
wedelman@milberg.com

Michael A. Acciavatti (*pro hac vice* forthcoming)
MILBERG, PLLC
405 East 50th Street
New York, NY 10022
Tel: 212.594.5300

