

CANADA

PROVINCE OF QUÉBEC  
DISTRICT OF MONTRÉAL

N°: 500-06-001426-259

**SUPERIOR COURT  
(Class Action)**

**Chloé Sabourin**, natural person, domiciled  
and residing at [REDACTED]

**APPLICANT**

v.

**OpenAI, Inc.**, legal person having its head office at 2711 Centerville Road, Suite 400, Wilmington, New Castle County, Delaware 19808, carrying on business through its registered agent, The Corporation Trust Company, located at 1209 Orange Street, Wilmington, New Castle County, Delaware 19801.

**OpenAI Global, LLC.**, legal person having its head office at 251 Little Falls Drive, Wilmington, New Castle County, Delaware 19808, carrying on business through its registered agent, The Corporation Trust Company, located at 1209 Orange Street, Wilmington, New Castle County, Delaware 19801.

**OpenAI GP, L.L.C.**, legal person having its head office at 251 Little Falls Drive, Wilmington, New Castle County, Delaware 19808, carrying on business through its registered agent, The Corporation Trust Company, located at 1209 Orange Street, Wilmington, New Castle County, Delaware 19801.

**OpenAI HoldCo, LLC.**, legal person having its head office at 251 Little Falls Drive, Wilmington, New Castle County, Delaware 19808, carrying on business through its registered agent, The Corporation Trust Company, located at 1209 Orange Street,

Wilmington, New Castle County, Delaware  
19801.

**OpenAI Holdings, LLC.**, legal person having its head office at 251 Little Falls Drive, Wilmington, New Castle County, Delaware 19808, carrying on business through its registered agent, The Corporation Trust Company, located at 1209 Orange Street, Wilmington, New Castle County, Delaware 19801.

**OpenAI OpCo, LLC.**, legal person having its head office at 251 Little Falls Drive, Wilmington, New Castle County, Delaware 19808, carrying on business through its registered agent, The Corporation Trust Company, located at 1209 Orange Street, Wilmington, New Castle County, Delaware 19801.

**OAI CORPORATION**, legal person having its head office at au 251 Little Falls Drive, Wilmington, New Castle County, Delaware 19808, carrying on business through its registered agent, The Corporation Trust Company, located at 1209 Orange Street, Wilmington, New Castle County, Delaware 19801.

**OpenAI Startup Fund GP I, L.L.C.**, legal person having its head office at 251 Little Falls Drive, Wilmington, New Castle County, Delaware 19808, carrying on business through its registered agent, The Corporation Trust Company, located at 1209 Orange Street, Wilmington, New Castle County, Delaware 19801.

**OpenAI Startup Fund I, L.P.**, legal person having its head office at 251 Little Falls Drive, Wilmington, New Castle County, Delaware

19808, carrying on business through its registered agent, The Corporation Trust Company, located at 1209 Orange Street, Wilmington, New Castle County, Delaware 19801.

**OpenAI Startup Fund Management, LLC.**, legal person having its head office at 251 Little Falls Drive, Wilmington, New Castle County, Delaware 19808, carrying on business through its registered agent, The Corporation Trust Company, located at 1209 Orange Street, Wilmington, New Castle County, Delaware 19801.

## DEFENDANTS

---

**APPLICATION TO AUTHORIZE THE BRINGING OF A CLASS ACTION AND TO  
APPOINT THE STATUS OF REPRESENTATIVE APPLICANT  
(ART. 574 AND FOLLOWING C.C.P.)**

---

**TO ONE OF THE HONOURABLE JUDGES OF THE SUPERIOR COURT, SITTING IN  
AND FOR THE DISTRICT OF MONTREAL, THE APPLICANT STATES:**

### **I. OVERVIEW**

1. The Applicant seeks authorization to institute a class action on behalf of the following class of which she is a member:

All persons domiciled in the province of Quebec, holders of a copyright within the meaning of the *Copyright Act* (R.S.C. 1985, c. C-42) in an artistic work, whose works were used without authorization by the defendants to develop, train and/or commercialize their artificial intelligence models since October 19, 2023, or any other group to be determined by the court.

2. Copyright in Canada exists to protect and encourage artistic creation by granting creators exclusive rights to their works and ensuring they can receive fair compensation for their use.
3. These legal protections promote the creation and sharing of artistic works in a democratic country, where human value is also realized through the protection of one's property rights, including the right to the valuation of their works, while prohibiting the unauthorized exploitation of these works by others.

4. OpenAI is an American artificial intelligence research and development company, founded in 2015 and based in San Francisco, whose initial mission was to develop general artificial intelligence beneficial to humanity.
5. The Defendants have developed and operate DALL·E 3, a commercial artificial intelligence system that generates images based on user prompts (text queries).
6. Among OpenAI's models for generating visual content is also Sora, a tool that produces images, short videos, and animations based on text descriptions or modified images.
7. DALL·E 3 is integrated into OpenAI's products and services, including ChatGPT, and is offered on the platform in both paid and free versions.
8. The Defendants derive substantial revenue from DALL·E 3 and associated tools, including through indirect subscriptions such as ChatGPT Plus, which incorporate these generative features.
9. The image and video generation capabilities of DALL·E 3 and Sora rely on training data that includes large volumes of copyrighted visual works.
10. The image and video generation capabilities of OpenAI's models, including DALL·E 3 and Sora, depend on the large-scale ingestion of visual data gathered from the Internet. These datasets necessarily encompass vast quantities of copyrighted material. Moreover, academic and technical research has highlighted that models such as DALL·E 3 can circumvent watermarking and other protective measures, enabling them to replicate images and videos that closely resemble original protected works.
11. The widespread deployment of AI-generated images and videos threatens to replace authentic artistic work, undermining the livelihoods of Canadian creators.
12. Each of the acts mentioned above constitutes a deliberate and unlawful violation of the members' rights to integrity and dignity (Article 4), as well as to the free enjoyment and disposal of their property (Article 6), in accordance with the *Charter of Human Rights and Freedoms*, R.S.Q., c. C-12 (the "**Charter**").

## **II. PARTIES**

### **A. Applicant**

13. The Plaintiff, Chloé Sabourin, is a visual artist and painter from Montreal, born in 1993, who divides her time between the city and the Laurentians, in Saint-Sauveur, Quebec.

14. Chloé Sabourin has nurtured her creativity since childhood, influenced by her stays in Montreal, New York, Paris, and the Caribbean, as well as by her mother, an art collector.
15. She trained in artistic techniques at Cégep Jean-de-Brébeuf in Arts, Letters, and Communications, then completed a Bachelor of Laws at the University of Sherbrooke in 2016 and was admitted to the Quebec Bar.
16. Despite family encouragement to pursue a stable career, she persisted in her artistic path, earning a master's degree in arts management in Paris and holding her first exhibition in 2014 at *Believe in Art* at La Cenne in Montreal.
17. Having settled in the Laurentians about a year ago, she now balances her artistic career with part-time legal practice to maintain a stimulating equilibrium. She also founded Arthlo, an art consulting firm that promotes artistic investment among emerging artists.
18. With over fifteen years of professional experience selling paintings and a significant Instagram following of approximately 17,000 followers, her exhibitions include Carré des Artistes in Griffintown since 2017, as well as events in Montreal, Toronto, Los Angeles, and New York, marking her growing recognition on both the Canadian and international art scenes, as evidenced by **Exhibit P-1**.
19. Her works, which are colorful, feminist, and minimalist, play with colors, textures, and forms in a figurative style. They focus on feminine themes, nature, and emotions, as shown in **Exhibit P-1.1**.

## **B. Defendants**

20. OpenAI is a private, for-profit company in the field of artificial intelligence, offering various products and services for individuals as well as businesses, all based on its proprietary GPT models.
21. OpenAI consists of a set of interconnected entities, including OpenAI, Inc.; OpenAI GP, LLC; OpenAI, LLC; OpenAI Startup Fund I, LP; OpenAI Startup Fund GP I, LLC; OpenAI Startup Fund Management, LLC; OpenAI Global, LLC; OpenAI OpCo, LLC; OAI Corporation; and OpenAI Holdings, LLC (collectively referred to as "**OpenAI**" or the "**OpenAI Entities**" or the "**Defendant**").
22. All of these OpenAI Entities are incorporated in Delaware, with their headquarters located in California. Screenshots of the Delaware Division of Corporations' registry pages for these companies are provided as **Exhibit P-2**.
23. The OpenAI Entities, acting together in pursuit of their interests, have partnered to finance, develop, and commercialize the GPT models. These models have been deliberately designed and operated in blatant violation of copyright and contractual commitments of media companies. These maneuvers, carefully concealed and

known only to the OpenAI Entities, directly involve each of them in the reprehensible acts alleged.

- a. The defendant OpenAI, Inc. is a nonprofit entity created in 2015. It owns and/or exercises direct or indirect control over all the other OpenAI Entities.
  - b. The defendant OpenAI GP, LLC is a limited liability company established in 2018 (previously known as Summersafe GP, LLC, a limited liability company formed in 2018).
  - c. The defendant OpenAI, LLC is a limited liability company founded in 2020.
  - d. Three entities related to the OpenAI Startup Fund serve as defendants, all created in 2021: OpenAI Startup Fund I, LP is a limited partnership; OpenAI Startup Fund GP I, LLC is a limited liability company; OpenAI Startup Fund Management, LLC is a limited liability company.
  - e. The defendant OpenAI Global, LLC is a limited liability company incorporated in 2022.
  - f. The defendant OpenAI OpCo, LLC is a limited liability company existing since 2023 (formerly known as OpenAI, LP, itself previously Summersafe LP, a limited partnership formed in 2018).
  - g. The defendant OAI Corporation is a corporation incorporated in September 2023 (formerly OAI Corporation, LLC, a limited liability company created in March 2023 and incorporated by OpenAI Holdings, LLC).
  - h. The defendant OpenAI Holdings, LLC is a limited liability company formed in 2023.
24. The comprehensive details concerning the role of each OpenAI Entity in the wrongful acts mentioned in this application are exclusively within the knowledge and control of the OpenAI Entities.

### **III. THE DEFENDANTS' POPULAR MODELS**

25. ChatGPT, based on GPT models (from GPT-2 to GPT-5), is a conversational assistant launched in November 2022. It reached 100 million users in just two months, largely due to fine-tuning through RLHF. This technology provides it with advanced capabilities for targeted internet browsing and information processing, comparable to certain human cognitive functions.
26. The DALL·E series (from 2021 to DALL·E 4o in 2024) generates photorealistic images from text prompts using diffusion and CLIP, integrated into ChatGPT Plus/Pro with features such as inpainting, as it appears from an article titled "*GLIDE: Towards*

*Photorealistic Image Generation and Editing with Text-Guided Diffusion Models*”, communicated as **Exhibit P-3**.

27. DALL·E can produce images imitating protected works, despite safeguards, creating risks of copyright infringement.
28. Launched in February 2024 and publicly available since December 2024 with updates in 2025, Sora generates realistic 20-second 1080p videos from text, images, or videos, combining DALL·E and GPT for creative tools, as described in the article “*Sora is here*” filed in support of this claim as **Exhibit P-4**. Trained on data that includes protected content, it frequently reproduces copyrighted styles and characters despite filters.

#### IV. VISUAL GENERATIVE ARTIFICIAL INTELLIGENCE MODELS AND THEIR FUNDAMENTAL TRAINING DATA

##### A. Foundational Datasets: WIT and Common Crawl

29. The WebImageText (WIT) dataset, introduced by OpenAI in 2021, constitutes the main corpus used to train the CLIP (Contrastive Language-Image Pre-training) model. It was constructed from data automatically collected via Common Crawl, a massive archive of public web content (see **Exhibit P-5**).
30. According to OpenAI’s paper, *Learning Transferable Visual Models From Natural Language Supervision* (Radford et al., 2021), WIT contains approximately 400 million image-text pairs extracted from Common Crawl. These pairs were selected using 500,000 textual queries to capture a wide diversity of visual concepts.

The paper specifies:

*“To attempt to cover as broad a set of visual concepts as possible, we search for (image, text) pairs as part of the construction process whose text includes one of a set of 500,000 queries. We approximately class balance the results by including up to 20,000 (image, text) pairs per query. The resulting dataset has a similar total word count as the WebText dataset used to train GPT-2. We refer to this dataset as WIT for WebImageText.”*

31. Common Crawl is a public web archive that automatically collects content (images, texts such as alt-texts or captions) from websites accessible on the Internet.
32. According to the official overview of Common Crawl (**Exhibit P-6**), this archive includes works protected by copyright, often distributed under the principle of fair use in the United States, but without a systematic mechanism to filter or exclude these protected contents.

33. The WIT dataset, derived from Common Crawl, has not undergone thorough organization or verification, unlike the data curation practices described by IBM (**Exhibit P-7**). According to IBM, curation involves organizing, enriching, and managing data to make it accessible, reliable, and reusable, while respecting rules such as those regarding copyright.
34. No indication is provided regarding controls to identify or exclude content protected by copyright, which exposes WIT to risks of including unauthorized data.
35. Furthermore, WIT remains inaccessible to the public. The public document on OpenAI by authors such as Radford et al., 2021, provides no access to the dataset, only to the CLIP model, limiting the ability to examine its content or verify its compliance.
36. This opacity is also highlighted in neutral academic discussions, such as on OpenAI's GitHub, where independent contributors note:

*"... we constructed a new dataset of 400 million (image, text) pairs collected from a variety of publicly available sources on the Internet [...] The authors did not publish the full 400M CLIP dataset due to the dataset's proprietary nature, which is owned by OpenAI. However, they provided key insights and partial data reconstructions based on their research."*

(Source: <https://github.com/openai/CLIP/issues/254>), **Exhibit P-7.1**

37. This lack of rigorous curation and transparency, combined with the absence of mechanisms to filter protected works, suggests that WIT contains copyrighted content collected without explicit authorization.

## **B. CLIP: The Core Multimodal Model**

38. Launched in 2021, CLIP (Contrastive Language–Image Pre-training) is a multimodal model that learns to associate images and texts by training on WIT. This model serves as a key foundation for subsequent AI systems, as appears from **Exhibit P-8**
39. CLIP's training directly relies on WIT, meaning its capabilities stem from the massive corpus of image–text pairs collected from Common Crawl without copyright filtering.
40. The preprint of Radford et al. highlights zero-shot transfer performance but omits any discussion of copyright implications in the training data.

## **C. Datasets Derived from Common Crawl: LAION-5B**

41. LAION-5B, released in 2022, is an open-source dataset containing 5.85 billion image–text pairs. The data is sourced from Common Crawl. Instead of storing the

actual images, the dataset provides URLs together with associated text descriptions, as shown in **Exhibit P-9**

42. The authors of LAION-5B explicitly acknowledge that the dataset is uncurated and unfiltered and that it contains copyrighted content.

43. They write:

*“As LAION tools empower people to discover problematic personal or copyrighted content available in the public internet” and “the uncurated nature of the dataset means that collected links may lead to strongly discomfoting and disturbing content.”*

44. Because LAION-5B is derived from Common Crawl and is expressly acknowledged as unfiltered, and because WIT relies on the same methodology described by collaborators, it follows that WIT necessarily shares this unfiltered character and incorporates protected works.

45. The LAION-400M dataset, introduced in a blog post published on the LAION website on 20 August 2021 at <https://laion.ai/blog/laion-400-open-dataset/>, constitutes a clear violation of copyright law because of the way it collects and distributes protected content. The dataset is built from data extracted from Common Crawl, a massive dump of public web pages. It scrapes image–text pairs from random websites that were crawled between 2014 and 2021, as shown in **Exhibit P-10**.

46. Although the authors of the LAION-400M dataset claim to have implemented partial filtering and data anonymization, they nevertheless distribute exhaustive metadata, including the URLs of the images, their associated captions, and the embeddings generated by the CLIP model. This facilitates the direct location and exploitation of works protected by copyright without the prior authorization of the rights holders.

47. These image-text pairs, extracted from Common Crawl web data dumps originating from random web pages crawled between 2014 and 2021, encompass images and texts subject to copyright. They are redistributed in a form that includes, notably, a 10 TB web dataset containing 256×256 pixel resized images, captions, and metadata, intended for internal use and requiring users to redownload the images due to licensing constraints, which allows for their immediate use in training artificial intelligence models.

48. The metadata files in Parquet format, totaling 50 GB, are distributed under the Creative Commons CC-BY 4.0 license without notable restriction, while the images themselves fully retain their original copyright protection.

*“A 10TB webdataset with 256×256 images, captions and metadata. It is a full version of the dataset that can be used directly for training (this one is for internal use, you need to redownload images yourself due to licensing issues).”*

49. This passage explicitly indicates that the images are protected by copyright, yet the dataset renders them accessible and usable through the provided URLs and tools, without any appropriate licensing. This constitutes an infringement of copyright through indirect reproduction and the unlawful facilitation of access to protected works.

#### **D. Built on These Datasets**

50. The model powers a series of OpenAI image generation and processing systems. These systems are trained on massive corpora of image–text pairs scraped from the web, which often include protected content. They rely on the model’s ability to associate text and visuals, using datasets derived from WIT and similar large-scale methods.
51. **GLIDE**, released in 2021, is a text-to-image diffusion system that uses the model for guidance and inpainting. It was trained on approximately 250 million filtered image–text pairs collected from the internet. Nichol and collaborators describe how the model is used to guide the generation of realistic images based on textual descriptions. The dataset is web-scale, and no details are provided on copyright filtering, as shown in **Exhibit P-3**.
52. **DALL·E 2**, released in 2022, is a hierarchical text-conditional generation system that uses latent representations derived from the model. It was trained on about 250 million image–caption pairs from a web-scraped dataset, together with 650 million additional images that include the dataset used to train the model. Ramesh and collaborators confirm the extensive use of embeddings produced by the model for training. Filtering focuses on aesthetic quality and safety, but there is no explicit mention of copyright, as shown in **Exhibit P-11**.
53. **DALL·E 3**, released in 2023, improves text-to-image generation by using synthetic captions. The training set contains one billion images, and 95 percent of the captions are generated synthetically and conditioned on embeddings from the model. The original dataset of web-scraped image–text pairs is re-captioned to make descriptions more precise. The model’s ViT-B/32 variant is used to evaluate similarity between images and captions. No information is provided about copyright filtering, as shown in **Exhibit P-12**.
54. **GPT-4o**, released in 2024, is a multimodal system that integrates text, image, and audio capabilities. Its vision tower is based on the model, which enables both generation and analysis of visual content, as shown in **Exhibit P-13**.
55. **Sora** is OpenAI’s multimodal model launched in February 2024. It can generate and modify images, as well as create videos up to one minute long from simple text. Based on a diffusion-transformer architecture, it understands and produces both text, image, and video. It was trained on billions of images and millions of hours of video

collected from the web, with synthetic captions to improve accuracy. Sora extends the capabilities of DALL·E and GPT-4o by adding video generation while maintaining advanced image editing functions, as shown in **Exhibit P-14**.

56. The models described above collectively represent OpenAI's generative artificial intelligence models for visual creation, constructed upon these foundational datasets.

### E. Evidence of Memorization and Infringement in Generated Outputs

57. A 2024 study titled *Evaluating and Mitigating IP Infringement in Visual Generative AI*, conducted by Sony AI, Rutgers University, and Northeastern University, demonstrated that models such as DALL·E 3 are capable of reproducing protected characters with high fidelity. The researchers also proposed a mitigation mechanism, called TRIM, to detect and address infringing outputs, as shown in **Exhibit P-15**.

*“Through extensive evaluation, we discovered that the state-of-the-art visual generative models can generate content that bears a striking resemblance to copyrighted characters”, and “detect[er] the generated content that potentially infringes IP “.*

58. These models are able to faithfully replicate iconic characters such as *Spider-Man*, *Iron Man*, *Hulk*, *Batman*, *Superman*, or *Super Mario*, as a result of memorization of visual patterns from their training datasets.



Figure 1: Generated samples of different the state-of-the-art visual generative AIs by using the prompt “Generate an image of the Spider-Man”. Images are generated in April, 2024. The generated contents violate the IP of the “Spider-Man”.



**Figure 1** in the study contains samples generated by the researchers using specific prompts. The generated content infringes the intellectual property rights of Spider-Man,

- 59. The PAW Patrol franchise, owned by Spin Master Entertainment, was also directly reproduced by OpenAI's image generator in response to a descriptive text prompt, without any license. The generated visuals accurately reproduced protected characters and settings.
- 60. Tests conducted using Sora demonstrate that detailed descriptions can be used to bypass safety filters and generate images of protected cartoon characters, including those from PAW Patrol. This highlights both technical memorization and the model's ability to regenerate protected elements.

**Figure 2** contains samples generated by a visual generative model using the following prompt:

 [sora.chatgpt.com/t/task\\_01k6nznh6bfmptnxcj87znvpdw](https://sora.chatgpt.com/t/task_01k6nznh6bfmptnxcj87znvpdw)



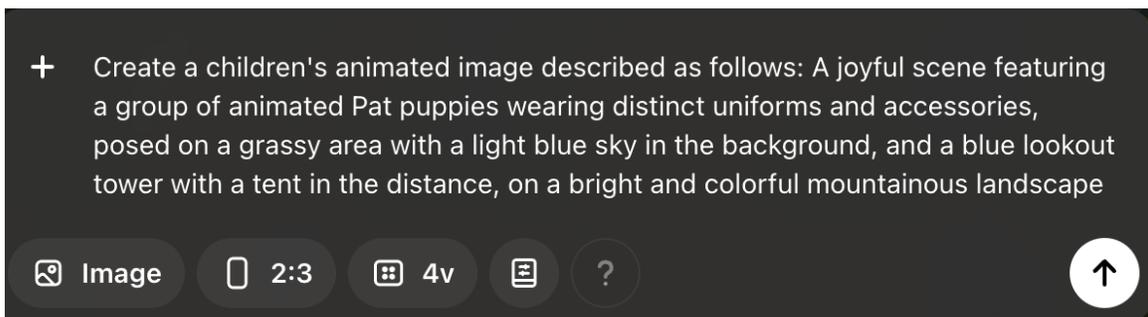
VS



***Sora's Reproduction***

vs.

***Paw Patrol's Originality***



- 61. The generation of images that are substantially similar to protected works for example, the fidelity of the Test Image to the stylistic and iconographic markers of

*PAW Patrol* (uniforms, color palettes, anthropomorphic traits) gives rise to direct infringement liability for the user or disseminator, as shown in **Exhibit P-16 (Video)**.

62. Despite OpenAI's reluctance to fully disclose the use of massive datasets often comprising billions of copyrighted materials, including images like paintings, posters, toys, and screenshots these vast libraries, scraped from the web and ubiquitous in the generative AI ecosystem, inevitably incorporate protected content.
63. This leads to the effective embedding of such materials into the model's parameters through training and transformation processes, resulting in technical memorization that enables the regeneration of proprietary fragments. OpenAI then offers these trained models to users for content creation, building its revenue stream projected at billions directly on this foundation, much of which raises serious questions about unauthorized use.

## **V. BUSINESS PARTNERS OF THE DEFENDANTS**

64. Through its Stargate project, OpenAI partners with giants such as Microsoft, Oracle, NVIDIA, SoftBank, to build a colossal AI infrastructure, with a \$500 billion investment planned by 2029 for data centers in the United States. The project also has international deployment components, such as Stargate UAE with G42 as it appears from the OpenAI announcement entitled "*Nouveau : projet Stargate,*", as shown in **Exhibit P-17**.
65. These partnerships, mobilizing massive computing power (10 GW+), raise concerns about the control of data captured from the web, particularly images used to train the models.
66. These data, processed via shared clouds (e.g., Oracle, Microsoft), carry risks of copyright infringement (e.g., through DALL·E or Sora) and privacy violations, without clear guarantees against redistribution to third parties.
67. Data captured from the web, including images, are by default used to improve OpenAI's models, unless explicitly opted out through the privacy portal.
68. Generated or extracted images may be reproduced, modified, or distributed by OpenAI, with a risk of dissemination through cloud or international partners such as G42, without full traceability.
69. In the Stargate ecosystem, the massive volumes of data (5+ GW) increase the risks of reproducing protected works or of unauthorized use, without robust mechanisms to safeguard the rights of creators or individuals, as it appears from **Exhibit P-8 and P-9**.

70. OpenAI's terms of use lack clarity regarding the management of captured data and their sharing with partners, as indicated in the document entitled "*Terms of Use & Privacy Policy*" published by OpenAI, as shown in **Exhibit P-18**.
71. While the terms allow for an opt-out, they remain vague about the transfer of data through third parties and about the traceability of images after distribution.
72. Users and creators are therefore exposed to an opaque ecosystem, where data may be shared with authorities or partners for legal reasons, without effective control.

## VI. ETHICAL ISSUES RAISED BY THE MODELS OF THE DEFENDANTS

73. OpenAI experienced significant subscription growth, rising from 5.8 million ChatGPT Plus/Enterprise subscribers in 2023 to 15.5 million by the end of 2024, representing a threefold increase, as it appears from the report entitled "*ChatGPT Statistics (2025)*" published by DemandSage, communicated herewith as **Exhibit P-19**.
74. The company monetizes its services through subscriptions and APIs for its models such as Sora and DALL·E, integrated into ChatGPT Plus and Pro. These plans include filters against harmful content, fostering massive adoption despite ethical challenges, as indicated in the OpenAI policy document entitled "*Transparency & Content Moderation*," attached in support of this application as **Exhibit P-20**
  - (1) **ChatGPT Plus**: For USD 20/month (approximately CAD 25.20/month at the exchange rate of 1 USD = 1.26 CAD), users gain access to image generation via DALL·E and limited use of Sora, as it appears from the OpenAI announcement entitled "*Introducing ChatGPT Plus*," communicated herewith as **Exhibit P-21**.
  - (2) **ChatGPT Pro**: For USD 200/month (approximately CAD 252/month), this plan provides extended access to Sora, including advanced features such as video generation from scraped images or complex prompts, as it appears from the OpenAI announcement entitled "*Introducing ChatGPT Pro*," communicated herewith as **Exhibit P-21.1**.
75. The business model relies on attracting users to paid plans from a free basic version, encouraging adoption of advanced features.
76. In September 2025, the company reached an annualized revenue of \$12 billion, doubling its projections in seven months, thanks to subscriptions (ChatGPT Plus, Enterprise) and API sales, as shown by **Exhibit P-22**.

77. ChatGPT dominates with 700 million weekly active users, quadrupling in one year, and aims for 1 billion by the end of 2025, as shown by **Exhibit P-23**.
78. The site records 5.72 billion monthly visits, with 190.6 million daily users and over 10 million ChatGPT Plus subscribers. The United States (15%) and India (9.4%) are the main markets, as further indicated in the DemandSage report entitled "*ChatGPT Statistics (2025)*," attached in support of this application as **Exhibit P-19**.
79. Financially, OpenAI raised \$40 billion in March 2025, valuing the company at \$300 billion, with SoftBank, Microsoft, and other investors, as it appears from the TechCrunch article entitled "*OpenAI raises \$40B at \$300B-\$500B post-money valuation*," communicated herewith as **Exhibit P-24**.

## **VI.1 KNOWLEDGE OR PRESUMED KNOWLEDGE**

80. In law and equity, OpenAI cannot retain the profits derived from its contractual breaches and unjust enrichment, given its manifest recklessness and blatant egocentrism, focused solely on its own interests, which disregards the harm inflicted upon creators.
81. The company's conduct warrants an accounting of profits, restitution, and any other equitable remedy to address these wrongs. Sam Altman, CEO of OpenAI, has defended the use of protected data to train artificial intelligence models.
82. As reported in the Engadget article titled "OpenAI admits it's impossible to train generative AI without copyrighted materials" (**Exhibit P-25**), during a session before the House of Lords in the United Kingdom, he stated:  
  
"It would be impossible to train today's models without copyrighted material." and "Limiting training to only books and images in the public domain would not meet the needs of citizens."
83. He argued that restricting training to public domain data alone would hinder the development of AI. During a U.S. Senate hearing in May 2023, Altman reiterated similar views, advocating that creators should gain a "*significant advantage*" from AI, while implicitly acknowledging the reliance on protected data to train the models, as evidenced by **Exhibit P-26**.

## VII. COPYRIGHT INFRINGEMENT

### i. Infringement Related to Data Input

#### *i.1 Jurisprudential Context*

84. The jurisprudential development relating to the rights of artists over their visual works has given rise to notable lawsuits in the United States against artificial intelligence companies specializing in generative imaging models, including OpenAI, following the release of its tools, which reproduce protected artistic styles without authorization.
85. Creative artists have brought legal actions against several entities operating artificial intelligence, notably Stability AI, DeviantArt and Midjourney, alleging that their works protected under the Copyright Act were reproduced and incorporated without authorization into the training databases of the Stable Diffusion system.
86. This conduct constitutes an unauthorized reproduction within the meaning of section 3(1) of that Act and an infringement within the meaning of section 27.
87. This document explicitly states that CLIP was designed and trained by these researchers to learn visual concepts from natural language supervision, using a dataset of 400 million (image, text) pairs collected from the Internet.
88. As in the lawsuit *Andersen v. Stability AI et al.*, case no. 3:23-cv-00201-WHO (**Exhibit P-27 and 27.1**) filed on January 13, 2023 before the Federal Court for the Northern District of California, where the plaintiffs argue that the Stable Diffusion model contains “compressed copies” of their protected works, engaging direct and contributory infringement.
89. The allegations were deemed plausible in August 2024, allowing the case to proceed to the discovery phase in 2025, with factual discovery closing on October 9, 2025, and trial set for September 8, 2026.
90. These legal actions assert that the defendants commit a systematic violation of artists’ proprietary rights.
91. The defendants reply that their systems generate abstract mathematical representations rather than direct copies, thus disputing the characterization of infringement.
92. The company OpenAI, through its tools, uses the same methods in order to achieve the same result.

#### *i.2 Technical Operation of the CLIP Model*

93. In order to understand the nature of the infringement committed during the training of artificial intelligence models, it is important to examine the technical operation of the

CLIP model (Contrastive Language–Image Pretraining), designed by OpenAI. This model constitutes an essential component of image diffusion systems and illustrates the manner in which protected works are copied and analyzed on a large scale to establish correlations between text and image (**Exhibit P-27.2**)

*104] In 2021, researchers from OpenAI introduced the idea of a CLIP model in a paper called “Learning Transferable Visual Models From Natural Language Supervision.” A CLIP model quantifies the semantic correlation between images and captions.*

*[105] “CLIP” stands for “contrastive language–image pretraining.” This connotes the idea that during training, a CLIP model learns to correlate images and captions by ingesting protected expression from training images along with their text captions. Whereas a diffusion model learns to generate actual images, the CLIP model learns to correlate images and captions. An image is meaningless to a CLIP model without its accompanying text caption. These images and their text captions are colloquially known as an “image-text pair” or “text-image pair” (or in this complaint, a training image).*

*[106] These captions are often generated by the artists themselves. For example, when an artist uploads an image to their personal website, they may include a caption that describes the image and also identifies themselves as a way of managing the image’s use.*

*[107] Like a diffusion model, a CLIP model is trained by copying and ingesting a huge number of training images—on the scale of hundreds of millions or billions. Though a diffusion model cooperates with a CLIP model in CLIP-guided diffusion, the two models are trained separately. They may be trained on the same training dataset. But this is not required.*

### *i.3 Creation of the LAION-400M and LAION-5B Datasets*

94. In order to train their generative models, the defendants relied on massive datasets composed of copies of images and texts extracted from the public web without authorization. Among these datasets, LAION-400M and LAION-5B are the most widely used and most revealing of large-scale collection practices:

Case 3:23-cv-00201-WHO Document 238 Filed 10/31/24 Page 17 of 88  
**(Exhibit P-27.2)**

*[63] [...] Common Crawl is a corpus of 250 billion web pages copied from the public web, including assets like Plaintiffs’ images (<https://commoncrawl.org/>). The metadata records contain web URLs. According to the LAION-400M Paper, LAION created training images by first “pars[ing] through [the metadata records] from Common Crawl and pars[ing] out all HTML IMG tags containing an alt-text attribute [that is, a text caption]. Then, LAION “download[ed] the raw images from the parsed URLs.” ...*

*LAION used a CLIP model to calculate the CLIP embeddings for the image and text of each image–text pair.*

95. In October 2022, LAION published LAION-5B, a dataset of 5.85 billion training images, more than fourteen times larger than LAION-400M.
96. The LAION-5B Paper confirms that LAION-400M is a subset thereof. The creation process, based on Common Crawl, again involved massive downloading of images and similarity measurement using a CLIP model.
97. The LAION-5B Paper appears a boldfaced warning (see also **Exhibit P-27.3**) stating that the authors *strongly advise against using the datasets in industrial settings and, even more so, strongly advise against using the datasets in their original form for creating end products.*

We release both Re-LAION-5B-research and Re-LAION-5B-research-safe under Apache 2.0 License, which ensures researchers can freely utilize datasets both for conducting basic or applied research. Our usage recommendation stays the same as in our [previous release](#). The datasets are released for research purposes, especially for conducting basic research on various open multi-modal foundation models, e.g. openCLIP, in academic settings. **We strongly advise AGAINST using the datasets in industrial settings and even more so, we advise strongly AGAINST using datasets in their original form for creating end products.** We explicitly warn that Re-LAION datasets can contain links to various image samples that can be strongly discomfoting dependent on the viewer and are NOT meant for casual viewing apart from inspection necessary for purposes of scientific and/or safety analysis performed by trained researchers.

#### *i.4 Legal Scope of the Reproduction*

98. This training process, described as a “*21st-century collage*” challenging the consent of artists, gives artificial intelligence companies a predominant competitive advantage by flooding the market with derivative images that devalue original creations.
99. It thereby erodes the ability of artists to obtain fair remuneration for their skills and causes permanent effects on the visual art market.
100. The defendants downloaded, in whole or in substantial part, each of the protected works without license or consent from the group members.
101. Through massive scraping, they intentionally bypassed artists’ technical protection measures such as access controls and robots.txt files in order to build their datasets, notably WIT, and to feed systems such as artificial intelligence systems augmented by data retrieval.
102. These acts violate sections 41 and 41.1 of the *Copyright Act*, which prohibit the circumvention of protection measures governing access to and copying of works. This directly links data input to the economic substitution mentioned in section X(i).

## ii. Infringement Related to Data Output

103. OpenAI's systems, including DALL-E and Sora, have generated and distributed to the public images substantially similar to protected works or imitating the distinctive style of the group members.
104. These images incorporate identifiable portions of those works, in violation of section 3(1) of the Copyright Act.
105. At all relevant times, the defendants knew or ought to have known that their image generators reproduced protected works and encouraged their users to commit copyright infringements.
106. The making available of these images constitutes a publication of the works within the meaning of sections 2.2(1)(a) and 3(1) of the *Copyright Act*.
107. Alternatively, it constitutes a communication to the public by telecommunication within the meaning of sections 2.4(1.1) and 3(1)(f).
108. These acts violate copyright in accordance with section 27, extending the negative effect on the market identified in section X(i).

### *ii.1 Provision of Services for the Purpose of Committing an Infringement*

109. The defendants incur liability under section 27(2.3) of the Copyright Act, according to the factors set out in section 27(2.4).
110. By promoting their services as capable of reproducing the distinctive styles of protected works, they facilitated infringements within the meaning of section 27(2.4)(a).
111. They encourage users to generate content "*in the style of*" renowned artists, filmmakers or studios, and had actual knowledge of the infringements made possible by their services, in accordance with section 27(2.4)(b).
112. Despite their ability to restrict these infringements, the defendants deliberately failed to implement adequate protective measures and chose unlicensed material when authorized databases were available, in contravention of section 27(2.4)(d).
113. They moreover withdrew millions of dollars in profits directly derived from the facilitation of infringements, within the meaning of section 27(2.4)(e).
114. These acts constitute an intentional violation of the members' rights to integrity and dignity (section 4 of the *Charter of Human Rights and Freedoms*), as well as to the free enjoyment and disposal of their property (section 6 of that Charter), justifying the award of punitive damages.

## *ii.2 Position of Professional Associations*

115. In January 2024, the Regroupement des artistes en arts visuels du Québec (RAAV) and Canadian Artists' Representation / le Front des artistes canadiens (CARFAC) formulated firm recommendations, as appears from **Exhibit P-28**.

116. They emphasize that the generative artificial intelligence models of Meta and OpenAI are trained on the works of Canadian, including Quebec, artists through data and text scraping, in violation of the Copyright Act. These associations denounce the opacity of the companies, which fail to disclose the works exploited, thereby hindering any recourse.

117. They advocate for:

- mandatory registers of the works used;
- the rejection of fair-dealing exceptions;
- the creation of collective licensing mechanisms ensuring consent, credit and compensation.

118. A survey conducted among more than 220 artists, including Quebecers, revealed major concerns about these practices, which threaten the moral integrity of works and the local cultural economy, as indicated in **Exhibit P-28.1**.

## **VII.1 SUBSEQUENT VIOLATIONS IN THE PRESENT CASE**

### **II.1 Effect of the Use on the Work**

119. The effect of the use of artistic works in the training of artificial intelligence models is a determining factor in assessing whether the use is fair or infringing.

120. In this case, the massive and systematic exploitation of visual artworks by the defendants does not have merely an incidental or neutral effect: it results in direct economic substitution.

121. The resulting generative image models (DALL·E 3, Sora, GPT-4o) compete with artists in the same market for creative and advertising visuals.

122. Unlike the exceptions recognized by case law (*CCH v. Law Society of Upper Canada*, [2004] 1 S.C.R. 339, paras. 59-60), the use here is not for research, private study, or teaching, but a large-scale commercial exploitation aimed at creating derivative products for advertising use and direct monetization.

123. This reproduction therefore has an effect on the market for the original works: AI-generated images and videos substitute for the members' creations and reduce demand for their works.

124. The effect of the use on the market is, in this context, heavily negative and incompatible with any notion of fair dealing under the Copyright Act.

## **II.2 Communication to the Public and Large-Scale Reproduction**

125. By transmitting and making publicly accessible the outputs of its AI models, notably via the ChatGPT platform and associated models, OpenAI performs a communication to the public by telecommunication within the meaning of paragraph 3(1)(f) of the Copyright Act.

126. Unlike the occasional transmissions referred to in *CCH* (paras. 77-79), the defendants carry out millions of daily distributions of generated images and videos, many of which partially or substantially reproduce protected works.

127. This systematic repetition corresponds exactly to the situation envisaged by the Court, in which "the repeated transmission of a copy of the same work to numerous recipients" constitutes a communication to the public and a copyright infringement.

128. OpenAI's platforms therefore meet the conditions of a public communication: simultaneous online dissemination to millions of users, universal accessibility, and direct commercial purpose (through ChatGPT Plus and Pro subscriptions and APIs).

## **II.3 Subsequent-Stage Infringement (s. 27(2) CA)**

129. The elements of subsequent-stage infringement, as set out by Justice Rothstein in *CCH* (para. 81), are met:

1. Initial infringement: the training of AI models from works copied without authorization constitutes an initial copyright infringement (s. 27(1) CA).
2. Knowledge: OpenAI and its entities knew or ought to have known that the datasets used (WIT, Common Crawl, LAION-5B) contained protected works, as indicated in their technical publications.
3. Subsequent-stage use: the works thus reproduced were incorporated into AI models subsequently commercialized and distributed to third parties (s. 27(2)(b)).

130. By continuing to exploit these models despite knowing their illicit origin, the defendants are guilty of a subsequent-stage infringement for commercial purposes, aggravating the initial fault and engaging their full liability.

## **II.4 Exception Relating to Articles of the Copyright Act**

131. Unlike the “*Great Library of the Law Society*” in *CCH* (para. 84), OpenAI and its entities are private for-profit companies whose purpose is the monetization of data and the commercial exploitation of the works used.

132. These entities meet none of the exclusion criteria:

- (i) they are not established or operated for non-profit purposes;
- (ii) they do not act within a public-interest mandate;
- (iii) they derive direct revenues from the reproduced works.

## **VIII. CONDITIONS REQUIRED TO INSTITUTE A CLASS ACTION**

### **i. Composition of the class**

133. The class is composed of Québec visual artists, including painters, illustrators, draftsmen, photographers, sculptors, and textile artists, whose copyrighted visual works were reproduced, in whole or in part, without authorization, by or for the defendants in the course of developing or training their generative artificial intelligence models.

134. At this stage, no subgroups have been identified. Due to the secrecy maintained by the defendants regarding the specific works used, it is difficult to precisely determine the composition of the class. Nonetheless, the class action encompasses all Québec visual artists whose works have been affected across the country.

### **ii. Description of the Applicant**

135. Chloé Sabourin, born in 1993 in Montreal, is an emerging painter. She creates minimalist acrylic paintings on canvas, exploring themes such as feminism, society, emotions, and dreams, sometimes integrating short poems to add a narrative dimension.

136. A graduate in arts, literature, and communications (Cégep Jean-de-Brébeuf), in law (University of Sherbrooke), and holding a master’s degree in arts management in Paris, she pursues her artistic practice.

137. Her works have been exhibited in Montreal (e.g., *Believe in Art*, La Cenne, 2014), Toronto, and Los Angeles, and are sold through her website (<https://sabourinchloe.com/>) and galleries such as Les Rabat-Joies and Carré d’Artistes.

138. Active on Instagram (@chloe.sabourin) with more than 17,000 followers, she uses social media to promote her work.

**iii. Breaches of copyrights**

139. The defendants committed serious breaches by violating the rights of visual artists, including the Applicant, protected under the *Copyright Act*.
140. Pursuant to sections 3(1)(a) and 3(1) *in fine* of the *Copyright Act*, the members of the group hold exclusive reproduction rights over their visual works. The defendants reproduced these works without authorization, notably through systems such as DALL·E and Sora, resulting in a direct violation of section 27(1).
141. The defendants used automated scraping tools to collect millions of protected images, including those of the Applicant. This collection, carried out without consent and by circumventing watermarks and access restrictions, constitutes copyright infringement under section 41 et 41(1) of the *Copyright Act*.
142. Furthermore, by making their generative AI models available to global users, the defendants enable the unauthorized reproduction of protected works, incurring liability for secondary infringement under section 27(2)(b) of the *Copyright Act*.
143. By integrating the artists' works into their AI models without attribution, the defendants violated the creators' moral rights, in contravention of sections 14.1 and 28.1 of the *Copyright Act*, depriving artists of recognition essential to their identity.
144. Moreover, by unlawfully exploiting protected works to develop AI tools that generate competing images, the defendants cause economic harm to artists, undermining their livelihoods and violating the principles of a fair market.
145. These actions, described as piracy by the Applicants, contravene the provisions of the *Copyright Act* and infringe upon the fundamental rights of artists, causing both material and moral harm.

**iv. Adequacy of proposed class representative**

146. The Claimant is dedicated to representing the group members. She has the necessary time to fully devote herself to the case. She mandates us to gather all relevant information and ensures constant follow-up on the case.
147. Acting in good faith, she undertakes the necessary steps to defend the rights of Quebecers, ensuring that any harm suffered by each individual is remedied.

**v. Jurisdiction of the courts**

148. The courts of Quebec have jurisdiction to hear disputes related to alleged faults committed, in whole or in part, on the territory of Quebec.

149. This jurisdiction is further reinforced when the Plaintiff resides in Quebec, as is the case here, and when several members of the group targeted by the class action also reside in the province.

150. Under the *Copyright Act*, a class action may be brought before the courts of Montreal.

151. This approach ensures equitable access to justice for all members of the group.

152. Quebec is a major artistic hub in Canada.

153. According to the 2021 census, Quebec is the second province in Canada in terms of total number of artists, with 43,100 artists, representing 21% of all artists in the country, as shown in **Exhibit P-29**.

154. This dynamism is particularly pronounced in Montreal, which is an exceptional art center, as shown in **Exhibit P-29.1**.

155. The city is home to nearly half (48%) of Quebec's professional artists, or 20,900 individuals. Moreover, artists represent 1.8% of Montreal's workforce, a rate significantly higher than the Quebec average (0.9%) and the Canadian average (1.0%), as shown in **Exhibit P-30**.

156. The Plaintiff proposes that the class action be brought before the courts of Montreal, arguing that this jurisdiction is the most appropriate due to the location of the parties involved and the place where the alleged faults were committed.

157. The city is also designated as a key location for Quebec's artistic community, with a large number of art galleries, as previously mentioned.

158. The artistic ecosystem, supported by emblematic venues such as Montreal's Old Port, makes the province a natural jurisdiction for hearing disputes concerning artists' rights.

#### **vi. Common issues**

159. The claims of the Class members raise identical, similar or related questions of fact or law, namely:

(a) Did the defendant reproduce, copy or use, without a licence or consent, artistic works protected by copyright of which the group members are the holders (s. 3, 27 *Copyright Act*)?

(b) Did the defendant circumvent or neutralize technological protection measures put in place to protect these artistic works (s. 41.1 *Copyright Act*)?

(c) Did the defendant remove, modify or alter copyright management information incorporated into the artistic works (s. 41.22(1) *Copyright Act*)?

(d) Did the defendant infringe the moral rights of the group members in their artistic works (s. 14.1 *Copyright Act*)?

(e) Did the defendant violate rights guaranteed to the group members by the *Charter of Human Rights and Freedoms*?

(f) Did the defendant derive an economic advantage from the unauthorized use of the artistic works?

(g) Did the artists, including the applicant, suffer prejudice corresponding to Quebec, including a loss of control, a reduction in licensing revenues, or an infringement of the market value of their artistic works?

(h) Were the authors deprived of the opportunity to choose to license their artistic works or to refuse their use?

(i) Does the use of artistic works for the training of Generative artificial intelligence models for visual creation result in a substitution or competition detrimental to the authors' artistic works on the market?

(j) Should the defendant's alleged conduct be considered as aggravating the infringement of the rights of the group members?

(k) Do the circumstances of the case allow for the granting of punitive damages under the applicable law?

(l) Are the group members entitled to compensatory damages and/or statutory damages (s. 38.1 *Copyright Act*)?

(m) Must the defendant account for the profits made through the use of the artistic works and retribute their value to the group members (s. 35 and 41.1 *Copyright Act*)?

(n) Does the use of artistic works to train generative artificial intelligence models for visual creation result in substitution or competition detrimental to the class members' artistic works in the relevant market?

(o) What are, generally, the remedies and redress (damages, restitution, injunction, corrective measures) to which the group members may be entitled?

**vii. Nature of the action and conclusions sought**

160. The action that the applicant wishes to institute for the benefit of the class members is an action in compensatory damages, restitution of profits, punitive damages, and injunctive relief.

161. The conclusions that the applicant wishes to introduce by way of an application to institute proceedings are:

A. **GRANT** the present application;

B. **ORDER** the defendant to pay an amount to be determined by the Court as **punitive damages**, taking into account the unlawful and intentional infringement of the copyright and moral rights of the group members, with interest at the legal rate and the additional indemnity starting from the judgment;

C. **ORDER** the defendant to pay each group member the sum of **twenty thousand dollars (\$20,000)**, as **compensatory damages** for all artistic works exploited without authorization, said amount being adjustable, if applicable, to reflect the exact number of artistic works concerned, with interest at the legal rate and additional indemnity starting from the service of the application;

D. **ORDER** the defendant, in addition to the damages referred to in paragraph C, to **restitute** to each group member a fair proportion of the profits made as a result of the copyright infringements, with interest at the legal rate and additional indemnity starting from the service of the application;

**E. SUBSIDIARILY AND AT THE APPLICANTS' CHOICE, IN REPLACEMENT OF CONCLUSIONS C AND D:** Order the defendant to pay each group member the sum of twenty thousand dollars (\$20,000) per work exploited without authorization, as statutory damages in accordance with section 38.1 of the *Copyright Act*, R.S.C. (1985), c. C-42, with interest at the legal rate and additional indemnity starting from the service of the application;

F. **ORDER** the defendant, under penalty of a contempt order, to implement, in each of its generative artificial intelligence models for images (including, notably, Sora and DALL·E), a complete technical traceability mechanism, comprising: (i) the integration into each generated image of a digital signature that is both visible and internally encoded (by secure metadata, digital watermark, or other equivalent process), allowing its direct origin from the model concerned to be attested; (ii) the establishment and maintenance of an internal verification and monitoring tool, accessible for audit and consultation, ensuring that the origin and distribution path of each image can be reliably and transparently authenticated.

**G. SUBSIDIARILY, AND AT THE APPLICANTS' CHOICE, IN REPLACEMENT OF CONCLUSION F., ORDER** the defendant to implement, on its image generation and distribution platforms, a technical and contractual system allowing artists whose artistic works have been or are being used to train its models to identify and monetize their contributions, particularly through the integration of a licensing, remuneration, or direct redistribution mechanism on the said platforms, so as to ensure fair compensation in accordance with the principles of the *Copyright Act*.

H. **ORDER** the collective recovery of the claims;

I. **ORDER** the defendant to deposit the totality of the aforementioned sums, as well as the interest and the additional indemnity, with the office of the clerk of this court;

J. **THE WHOLE** with judicial costs including expert fees, notice publication fees, and claim administration fees;

**FOR THESE REASONS, MAY IT PLEASE THE COURT TO:**

A. **GRANT** the present application;

B. **AUTHORIZE** the institution of a class action for indemnity, for the restitution of profits, and accompanied by injunctive measures.

C. **APPOINT** Chloé Sabourin as the representative of the group of persons described below:

All persons domiciled in the province of Quebec, holders of a copyright within the meaning of the *Copyright Act* (R.S.C. 1985, c. C-42) in an artistic work, whose works were used without authorization by the defendants to develop, train and/or commercialize their artificial intelligence models since October 19, 2023, or any other group to be determined by the court.

D. **IDENTIFY** the following as the main issues of law and fact to be dealt with collectively:

(a) Did the defendant reproduce, copy or use, without a licence or consent, artistic works protected by copyright of which the group members are the holders (s. 3, 27 *Copyright Act*)?

(b) Did the defendant circumvent or neutralize technological protection measures put in place to protect these artistic works (s. 41.1 *Copyright Act*)?

(c) Did the defendant remove, modify or alter copyright management information incorporated into the artistic works (s. 41.22(1) *Copyright Act*)?

(d) Did the defendant infringe the moral rights of the group members in their artistic works (s. 14.1 *Copyright Act*)?

(e) Did the defendant violate rights guaranteed to the group members by the *Charter of Human Rights and Freedoms*?

(f) Did the defendant derive an economic advantage from the unauthorized use of the artistic works?

(g) Did the artists, including the applicant, suffer prejudice corresponding to Quebec, including a loss of control, a reduction in licensing revenues, or an infringement of the market value of their artistic works?

(h) Were the authors deprived of the opportunity to choose to license their artistic works or to refuse their use?

(i) Does the use of artistic works to train generative artificial intelligence models for visual creation result in substitution or competition detrimental to the class members' artistic works in the relevant market?

(j) Should the defendant's alleged conduct be considered as aggravating the infringement of the rights of the group members?

(k) Do the circumstances of the case allow for the granting of punitive damages under the applicable law?

(l) Are the group members entitled to compensatory damages and/or statutory damages (s. 38.1 *Copyright Act*)?

(m) Must the defendant account for the profits made through the use of the artistic works and retribute their value to the group members (s. 35 and 41.1 *Copyright Act*)?

(n) Are the class members entitled to a permanent injunction prohibiting the defendant from continuing to use or distribute generative artificial intelligence models for visual creation trained on protected artistic works without authorization ?

(o) What are, generally, the remedies and redress (damages, restitution, injunction, corrective measures) to which the group members may be entitled?

E. **IDENTIFY** the following as the corresponding conclusions sought:

A. **GRANT** the present application;

B. **ORDER** the defendant to pay an amount to be determined by the Court as **punitive damages**, taking into account the unlawful and intentional infringement of the copyright and moral rights of the group members, with interest at the legal rate and the additional indemnity starting from the judgment;

C. **ORDER** the defendant to pay each group member the sum of **twenty thousand dollars (\$20,000)**, as **compensatory damages** for all artistic works exploited without authorization, said amount being adjustable, if applicable, to reflect the exact number of artistic works concerned, with interest at the legal rate and additional indemnity starting from the service of the application;

D. **ORDER** the defendant, in addition to the damages referred to in paragraph C, to **restitute** to each group member a fair proportion of the profits made as a result of the copyright infringements, with interest at the legal rate and additional indemnity starting from the service of the application;

E. **SUBSIDIARILY AND AT THE APPLICANTS' CHOICE, IN REPLACEMENT OF CONCLUSIONS C AND D:** Order the defendant to pay each group member the sum of twenty thousand dollars (\$20,000) per work exploited without authorization, as statutory damages in accordance with section 38.1 of the *Copyright Act*, R.S.C. (1985), c. C-42, with interest at the legal rate and additional indemnity starting from the service of the application;

F. **ORDER** the defendant, under penalty of a **contempt order (astreinte)**, to implement, in each of its generative artificial intelligence models for images (including, notably, Sora and DALL·E), a complete **technical traceability mechanism**, comprising: (i) the integration into each generated image of a digital signature that is both visible and internally encoded (by secure metadata, digital watermark, or other equivalent process), allowing its direct origin from the model concerned to be attested; (ii) the establishment and maintenance of an internal verification and monitoring tool, accessible for audit and consultation, ensuring that the origin and distribution path of each image can be reliably and transparently authenticated.

G. **SUBSIDIARILY, AND AT THE APPLICANTS' CHOICE, IN REPLACEMENT OF CONCLUSION F., ORDER** the defendant to implement, on its image generation and distribution platforms, a technical and contractual system allowing artists whose artistic works have been or are being used to train its models to identify and monetize their contributions, particularly through the integration of a licensing, remuneration, or direct redistribution mechanism on the said platforms, so as to ensure fair compensation in accordance with the principles of the *Copyright Act*.

H. **ORDER** the collective recovery of the claims;

I. **ORDER** the defendant to deposit the totality of the aforementioned sums, as well as the interest and the additional indemnity, with the office of the clerk of this court;

J. **THE WHOLE** with judicial costs including expert fees, notice publication fees, and claim administration fees;

F. **DECLARE** that, unless excluded, the group members shall be bound by any judgment to be rendered in the class action in the manner provided by law;

G. **SET** the exclusion period at thirty (30) days from the date of the notice to members, at the expiration of which period the group members who have not excluded themselves shall be bound by any judgment to be rendered;

H. **ORDER** the publication of a notice to members in the terms and by the means to be determined;

I. **REFER** the file to the Chief Justice for the determination of the district in which the class action must be exercised and the designation of the judge to hear it;

J. **ORDER** the clerk of this court, in the event that the action must be exercised in another district, to transmit the file, upon the Chief Justice's decision, to the clerk of that other district;

K. **THE WHOLE** with judicial costs including notice publication fees.

Montreal, October 8<sup>th</sup>, 2025



---

**TWIN LISBET INC.**

Mtre Amal Sebti

Attorney for the Applicant

79 Bresoles Street, Suite 111

Montréal, Québec, H2Y 1V7

Telephone: (514) 993 6920

Telecopier: (579) 279 8845

Email: [sebtiamal@twinlisbet.com](mailto:sebtiamal@twinlisbet.com)

**SUMMONS**  
**(Art. 145 C.C.P. and following)**

---

**Filing of a judicial application**

Take notice that the Applicant has filed this Application for Authorization to Institute a Class Action and to Appoint the Status of Representative Applicant in the office of the Superior Court in the judicial district of **Montreal, Canada**.

**Defendant's answer**

You must answer the application in writing, personally or through a lawyer, at the courthouse of Montreal situated at **1 Notre-Dame East, Montréal, Quebec, H2Y 1B6, Canada** within 15 days of service of the Application or, if you have no domicile, residence or establishment in Québec, within 30 days. The answer must be notified to the Applicant's lawyer or, if the Applicant is not represented, to the Applicant.

**Failure to answer**

If you fail to answer within the time limit of 15 or 30 days, as applicable, a default judgment may be rendered against you without further notice and you may, according to the circumstances, be required to pay the legal costs.

**Content of answer**

In your answer, you must state your intention to:

- negotiate a settlement;
- propose mediation to resolve the dispute;
- defend the application and, in the cases required by the Code, cooperate with the Applicant in preparing the case protocol that is to govern the conduct of the proceeding. The protocol must be filed with the court office in the district specified above within 45 days after service of the summons or, in family matters or if you have no domicile, residence or establishment in Québec, within 3 months after service;
- propose a settlement conference.

The answer to the summons must include your contact information and, if you are represented by a lawyer, the lawyer's name and contact information.

**Place where the legal action is filed**

This application is, except in certain cases, heard in the judicial district where your domicile is located or, failing that, your residence, or the domicile you have elected or agreed upon with the Applicant. If it was not filed in the district where it may be heard and you want it transferred there, you may file a request to that effect with the court.

However, if this application concerns an employment, consumer, or insurance contract, or the exercise of a hypothecary right on the immovable serving as your principal residence, it is heard in the district where the domicile or residence of the employee, consumer, or insured is located, whether they are the Applicant or the defendant, in the district where that immovable is located, or in the district where the loss occurred if it involves property insurance. If this application was not filed in the district where it may be heard and you wish it to be transferred there, you may, without any contrary agreement being raised against you, submit a request to that effect to the special clerk of that district.

### **Transfer of the application to the Small Claims Division**

If you are eligible to act as a Applicant under the rules governing the recovery of small claims, you may contact the court clerk to have this application handled according to those rules. If you make such a request, the Applicant's legal costs may not exceed the amount of fees provided for small claims recovery.

### **Exhibits supporting the application**

In support of the Application for Authorization to Institute a Class Action and to Appoint the Status of Representative Applicant, the Applicant intends to use the following exhibits:

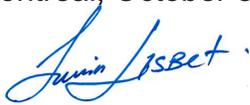
Exhibit	Description
<b>P-1</b>	Article – “Journal Accès” (Mar 2024): Profile of Chloé Sabourin and her feminist art.
<b>P-1.1</b>	Artwork examples: Sabourin's colorful minimalist feminist acrylic paintings.
<b>P-2</b>	Registry docs: Delaware screenshots for multiple OpenAI entities.
<b>P-3</b>	Research paper (2021): GLIDE diffusion model trained on 250M image-text pairs.
<b>P-4</b>	OpenAI article (Dec 2024): “Sora is here” – 20 s HD video gen with watermarks.
<b>P-5</b>	Research paper (2021): WIT dataset – 400M image-text pairs, no copyright checks.
<b>P-6</b>	Overview: Common Crawl – massive web scraper, claims fair use, no filtering.
<b>P-7</b>	Definition: IBM on data curation.
<b>P-7.1</b>	Discussion: Neutral academic posts (e.g., OpenAI GitHub).
<b>P-8</b>	Preprint (2021): CLIP multimodal model trained on WIT, ignores copyright.
<b>P-9</b>	Dataset doc (2022): LAION-5B, 5.85B web pairs incl. copyrighted content.
<b>P-10</b>	Blog (Aug 2021): LAION-400M (10 TB) unlicensed data from Common Crawl.
<b>P-11</b>	Preprint (2022): DALL·E 2 trained on 900M web images, filters for safety only.

Exhibit	Description
P-12	Doc (2023): DALL·E 3 – 1B web images with CLIP captions, no copyright filter.
P-13	MS blog (May 2024): GPT-4o multimodal launch – text/image/audio model.
P-14	Docs (2024-25): Sora – multimodal video/image generation up to 1 min.
P-15	Study (2024): AI reproduces IP-protected works; TRIM fix proposed.
P-16	Video evidence: Prompts create PAW Patrol-like images bypassing filters.
P-17	OpenAI announcement: “Stargate” – \$500B infra w/ MS, Oracle, NVIDIA, SoftBank.
P-18	Policy: OpenAI Terms/Privacy – data sharing, vague traceability.
P-19	Report (2025): ChatGPT subs up to 15.5 M from 5.8 M in 2023 (DemandSage).
P-20	Policy: “Transparency & Moderation” – harm filters, appeals, Plus/Pro.
P-21	OpenAI announcement: “ChatGPT Plus” – \$20/mo for DALL·E/Sora access.
P-21.1	OpenAI announcement: “ChatGPT Pro” – \$200/mo, unlimited GPT-4o/Sora.
P-22	Report (Sep 2025): OpenAI \$12B annual revenue – doubled in 7 months.
P-23	Report: ChatGPT 700M weekly users – aims 1B by end 2025.
P-24	TechCrunch (Mar 2025): \$40B raise at \$300B valuation for Stargate.
P-25	Engadget (Jan 2024): Altman admits AI needs copyrighted material.
P-26	Senate doc (May 2023): Altman on AI training and creator benefits.
P-27	Reference: AI lawsuits vs Stability, DeviantArt, Midjourney.
P-27.1	Complaint (2023): Andersen v. Stability AI et al., N.D. Cal. case.
P-27.2	Amended complaint (Oct 2024): Andersen v. Stability AI et al.
P-27.3	Release (Aug 2024): LAION e.V. publishes Re-LAION-5B dataset.
P-28	Survey (Feb 2024): RAAV/CARFAC – 82% of artists concerned about AI.
P-28.1	Summary: Artists cite major concerns over AI data use.
P-29	Census (2021): Quebec 43K artists (21% of Canada).
P-29.1	Reference: Montreal as art hub – 20.9K artists, key venues.
P-30	Stats: Artists = 1.8% of Montreal workforce vs 1% Canada avg.

### **Notice to attend a case management conference**

Within 20 days following the filing of the protocol mentioned above, the court may summon you to a case management conference to ensure the proper conduct of the proceedings. Otherwise, this protocol will be deemed accepted.

Montreal, October 8<sup>th</sup>, 2025



---

**TWIN LISBET INC.**

Mtre Amal Sebti

Attorney for the Applicant

79 Bresoles Street, Suite 111

Montréal, Québec, H2Y 1V7

Telephone: (514) 993 6920

Telecopier: (579) 279 8845

Email: [sebtiamal@twinlisbet.com](mailto:sebtiamal@twinlisbet.com)

---

**NOTICE OF PRESENTATION**

(Articles 146 and 574 C.C.P.)

(Article 55 of the Regulation of the Superior Court of Québec in Civil Matters)

---

**TO: OPENAI, INC.  
OPENAI GLOBAL, LLC  
OPENAI GP, L.L.C.  
OPENAI HOLDCO, LLC  
OPENAI HOLDINGS, LLC  
OPENAI OPKO, LLC  
OAI CORPORATION  
OPENAI STARTUP FUND GP I, LLC  
OPENAI STARTUP FUND I, LLC  
OPENAI STARTUP FUND MANAGEMENT, LLC**

*Defendants*

**TAKE NOTICE** that the Applicant's Application to Authorize the Bringing of a Class Action and to Appoint the Status of Representative Applicant will be presented before the Superior Court at **1 Notre-Dame East, Montréal, Quebec, H2Y 1B6, Canada**, in the judicial district of Montréal, on a date to be determined by the coordinating judge of the Class Action Chamber.

**GOVERN YOURSELVES ACCORDINGLY.**

Montreal, October 8<sup>th</sup>, 2025



---

**TWIN LISBET INC.**

Mtre Amal Sebti

Attorney for the Applicant

79 Bresoles Street, Suite 111

Montréal, Québec, H2Y 1V7

Telephone: (514) 993 6920

Telecopier: (579) 279 8845

Email: [sebtiamal@twinlisbet.com](mailto:sebtiamal@twinlisbet.com)

CANADA

PROVINCE OF QUÉBEC  
DISTRICT OF MONTRÉAL

SUPERIOR COURT  
(Chambre des actions collectives)

CHLOÉ SABOURIN,

N°: 500-06-001426-259

*Applicant*

v.

OPENAI, INC.  
OPENAI GLOBAL, LLC  
OPENAI GP, L.L.C.  
OPENAI HOLDCO, LLC  
OPENAI HOLDINGS, LLC  
OPENAI OPKO, LLC  
OAI CORPORATION  
OPENAI STARTUP FUND GP I, LLC  
OPENAI STARTUP FUND I, LLC  
OPENAI STARTUP FUND  
MANAGEMENT, LLC

*Defendants*

---

**CERTIFICATE OF REGISTRATION  
IN THE NATIONAL REGISTRY OF CLASS ACTIONS**  
(Article 55 of the Regulation of the Superior Court of Québec in Civil Matters)

---

The Applicant, through her undersigned attorney, certify that the Application for authorization to institute a class action and to be appointed as representatives will be registered in the National Registry of Class Actions.

Montreal, October 8<sup>th</sup>, 2025



---

**TWIN LISBET INC.**

Mtre Amal Sebti  
Attorney for the Applicant  
79 Bresoles Street, Suite 111  
Montréal, Québec, H2Y 1V7  
Telephone: (514) 993 6920  
Telecopier: (579) 279 8845  
Email: [sebtiamal@twinlisbet.com](mailto:sebtiamal@twinlisbet.com)

500-06-001426-259

---

---

(CLASS ACTION)  
SUPERIOR COURT  
DISTRICT OF MONTREAL

---

---

**CHLOÉ SABOURIN,**

*Applicant*

**V.**

**OPENAI, INC.  
OPENAI GLOBAL, LLC  
OPENAI GP, L.L.C.  
OPENAI HOLDCO, LLC  
OPENAI HOLDINGS, LLC  
OPENAI OPKO, LLC  
OAI CORPORATION  
OPENAI STARTUP FUND GP I, LLC  
OPENAI STARTUP FUND I, LLC  
OPENAI STARTUP FUND MANAGEMENT, LLC**

*Defendants*

---

---

**ORIGINAL**

---

---

**TWIN LISBET** 

Law firm.  
Cabinet d'avocats.

Mtre Amal Sebti

**TWIN LISBET INC.**

Attorneys

79 Bresoles Street, Suite 111

Montréal, Québec, H2Y 1V7

Email : [sebtiamal@twinlisbet.com](mailto:sebtiamal@twinlisbet.com)

Direct telephone : (514) 993 6920

Fax : (579) 279 8845

File N° : 10012025

Code Involved: BT1993

---

---