

The background features a dark blue gradient with abstract, wavy lines in shades of teal and purple. A network graph with small nodes and connecting lines is overlaid on these waves, extending across the width of the slide.

**Baker  
McKenzie.**

# Using Generative AI Tools – Legal And Ethical Risks

ACC SFBA | April 18, 2023



# Speakers



**Brad Newman**

Partner, Palo Alto  
Leader, AI Practice  
Chair, ABA AI Subcommittee  
bradford.newman  
@bakermckenzie.com



**Cortnie Abercrombie**  
**AI Truth**

CEO and Founder  
cortnie@aitruth.org



**Mackenzie Martin**

Partner, Dallas  
Member, AI Practice  
mackenzie.martin  
@bakermckenzie.com



**Justin Calderon**

Associate, Palo Alto  
Member, AI Practice  
justin.calderon  
@bakermckenzie.com



# WHAT YOU DON'T KNOW

AI'S UNSEEN INFLUENCE  
ON YOUR LIFE AND  
HOW TO TAKE BACK CONTROL

CORTNIE ABERCROMBIE



# Agenda

Note: This is not legal advice, nor should it be considered legal advice. This presentation is for educational and informational purposes only.

1 How Generative AI Tools Work

2 Legal And Ethical Questions to Consider

3 Use Cases

4 License Terms to Watch Out For

5 Risk of Violating Tool Use Restrictions

6 Overview of Current Lawsuits

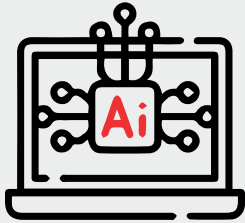
7 Best Practices and Ways to Mitigate Legal Risk



1

# How Generative AI Tools Work

# What is Generative AI?



Generative AI is a type of artificial intelligence that is capable of **creating new data** and content that did not previously exist.

It does this by learning the underlying patterns and structure of a particular dataset, and then **generating new content** that is similar to the original data.

# What is Generative AI?



Different from machine learning – **identifying a pattern** or correlation given a set of data and **making predictions** or **data specific insights**

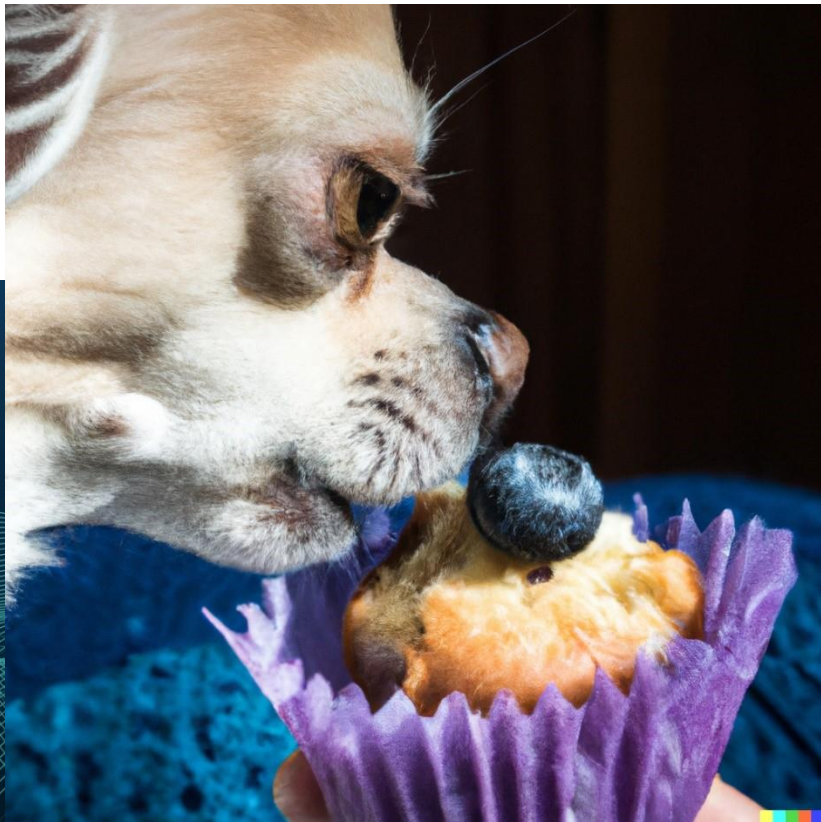


Generative AI **creates** new outputs based on the data they have been trained on.



"Blueberry muffin or  
Chihuahua?"

# What is Generative AI?



DALL-E prompt: "Chihuahua eating a blueberry muffin"





# Sources of AI Training Data



**Public datasets:** essentially anything available online (copyrighted or not).



**Private datasets:** collected by individuals or organizations for proprietary use, such as customer data.



**User-generated content:** social media posts, reviews.



**Custom data:** specific to intended application, such as medical AI trained with a dataset of medical images.



**Synthetic data:** generated using computer graphics or simulation techniques.



# Sources of AI Training Data



**ChatGPT:** around 45TB of text from the internet (web crawlers, Wikipedia, Reddit, etc.) and books



**DALL-E:** hundreds of millions of captioned images (training set not released).



**Stable Diffusion:** internet-acquired data (from websites like Pinterest, WordPress, Blogspot, Flickr, DeviantArt, etc.).



**Midjourney:** internet-acquired data, whether copyrighted or not.





2

# Legal And Ethical Questions to Consider

# Legal + Ethical Questions to Consider



Does the use of Generative AI tools infringe third-party IP, privacy, or other rights?

- Are you infringing anyone's rights?
- Who owns the output?
- Is the output copyrightable or patentable?
- Are there privacy concerns?
- What is the potential liability for the user?
- What are the implications for your company in using Generative AI output?



# Legal + Ethical Questions to Consider



Are you infringing anyone's rights?

- Generative AI is trained by data all across the internet.
- Should AI developers be made to pay every owner of every IP-protected material they scrape from the internet – like Napster in the early days of digital music?
- Are you, the game developer, affected by the copyright/trademark/patent of the material used to train the AI tool?
- Copyright lawsuit filed by Getty Images against Stability AI, creators of open source text-to-image generator Stable Diffusion.



AI generated image with the iconic Getty Images watermark

# Legal + Ethical Questions to Consider



## VentureBeat, Why are Getty and Shutterstock on opposite sides of the AI legal debate?, Jan. 17, 2023.

“...Bradford Newman, who leads the machine learning and AI practice of global law firm Baker McKenzie, maintains that while Getty has compared the current legal landscape of generative AI to the early days of digital music and Napster, the analogy does not hold.

‘Unlike in the music business, there are not large publishing houses and similar existing organizations with whom these AI developers can enter into licensing or royalty agreements,’ he said. ‘It is impractical for the developers to have to enter into tens of millions of individual agreements with the originator of each image scraped.’”

IN THE UNITED STATES DISTRICT COURT  
FOR THE DISTRICT OF DELAWARE

GETTY IMAGES (US), INC. )  
Plaintiff, )  
v. ) C.A. No.:  
STABILITY AI, INC. ) **DEMAND FOR JURY TRIAL**  
Defendant. )

COMPLAINT

Plaintiff Getty Images (US), Inc. (“Getty Images” or “Plaintiff”), by and through its undersigned attorneys, for its Complaint against Defendant Stability AI, Inc. (“Stability AI” or “Defendant”), hereby alleges as follows:

NATURE OF ACTION

1. This case arises from Stability AI’s brazen infringement of Getty Images’ intellectual property on a staggering scale. Upon information and belief, Stability AI has copied more than 12 million photographs from Getty Images’ collection, along with the associated captions and metadata, without permission from or compensation to Getty Images, as part of its efforts to build a competing business. As part of its unlawful scheme, Stability AI has removed or altered Getty Images’ copyright management information, provided false copyright management information, and infringed Getty Images’ famous trademarks.
2. Getty Images brings this action to recover damages that it has suffered and is continuing to suffer, and to prevent the irreparable harm caused by Stability AI’s intentional and willful acts.
3. Getty Images is one of the world’s leading creators and distributors of digital content. At great expense, over the course of nearly three decades, Getty Images has curated a

# Legal + Ethical Questions to Consider

## “Fair use” of the input



[VentureBeat: Who owns DALL-E images? Legal AI experts weigh in, Aug. 16, 2022.](#)

“Since DALL-E 2 was released in April, creative industries have been buzzing with questions about AI art image ownership. Back in August, Bradford Newman, who leads the machine learning (ML) and artificial intelligence (AI) practice of global law firm Baker McKenzie, in its Palo Alto office, said the answer to the question ‘Who owns DALL-E images?’ is **far from clear**. And, he emphasized, **legal fallout is inevitable**.

‘If DALL-E is adopted in the way I think [Open AI] envisions it, there’s going to be a lot of revenue generated by the use of the tool,’ he said. ‘And when you have a lot of players in the market and issues at stake, you have a **high chance of litigation**.’”



# Legal + Ethical Questions to Consider

## “Fair use” of the input



VentureBeat: *Why generative AI legal battles are brewing, Oct. 21, 2022.*

“I spoke to Newman back in August around issues of DALL-E image ownership – that is, the output of generative AI. Now, he said, legal questions are coming fast and furious around copyright and the ‘fair use’ of the input – that is, the training data going into generative AI tools.

‘Legally, right now, **there is little guidance,**’ he warned. ‘There are the inevitable class actions, but the net of it all is when you’re using the massive data sets that these AI applications are and you sprinkle on top of that open source licenses [as in the GitHub Copilot example], the arguments are going to be fair use versus infringement.’

Different courts, he predicted, will come to different conclusions. ‘Ultimately, I believe **this is going to go to the Supreme Court.**’”





# Legal Questions to Consider



Who owns the output?

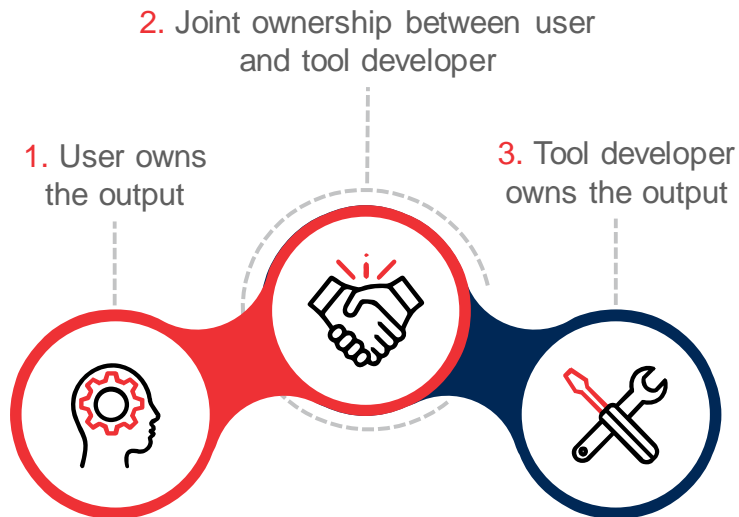
- The terms of use and licensing agreements for the generative AI tool dictates user's **relationship with the AI developer**
- Whether the generated content is original or based on pre-existing copyrighted/trademarked/patented materials affects user's potential **liability to IP holders**
- Whether the generative AI tool was trained on copyright/trademark/patent protected materials affects user's potential **liability to IP holders**



# Who Owns the Output?

As between the AI developer and the user –  
Read the Terms of Use

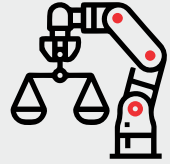
Possible outcomes:



Relationship with third party IP holders is a separate matter.

# Copyrightability

What is the current state of the law as to whether content created by an AI tool, or by a human using an AI tool, is copyrightable?



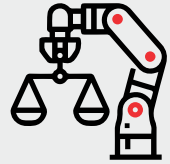
Traditionally, ownership of copyrightability in computer generated works was not in question because AI was merely a tool that supported the creative process like pen and paper.





Here, **AI makes the decisions** involved in the creative process.


# Copyrightability

What is the current state of the law as to whether content created by an AI tool, or by a human using an AI tool, is copyrightable?



 Major jurisdictions (US, EU) lean towards *not* granting copyright to material created by AI – **even if heavily guided by a human user.**

 New legislation?

 This means some uncertainty as to whether **your** outputs are protected by IP law.

# The Three Basic Elements of Copyright



## Originality

Must be the original work of the author (i.e., not a copy, and not a mere "idea").



## Fixation

It has to be "fixed" in a medium (i.e., a book, CD, electronic media).



## Creativity/"Work of Authorship"

Caselaw is clear that only a "modicum of creativity" is required, but it is unclear on what it means to be "creative".



# "Creativity" in the Copyright Office's Words

The copyright law only protects "the fruits of intellectual labor" that "are founded in the creative powers of the mind." *Trade-Mark Cases*, 100 U.S. 82, 94 (1879). Because copyright law is limited to "original intellectual conceptions of the author," the Office will refuse to register a claim if it determines that a human being did not create the work. *Burrow-Giles Lithographic Co. v. Sarony*, 111 U.S. 53, 58 (1884).

# Zarya of the Dawn (February 21, 2023)

- Comic book created with Midjourney.
- Text is protected by copyright – it was "written entirely" by a human
- Arrangement of text and images are copyrightable, because it was done by a human
- Images are *not* copyrightable because they were created through Midjourney
- Even human edits to images created by Midjourney did not render them copyrightable
- Lesson: if AI is the "master mind" it is not copyrightable

Van Lindberg, Esq.  
Taylor English Duma LLP

February 21, 2023

## I. DESCRIPTION OF THE WORK

As described in the application and accompanying deposit materials provided by Ms. Kashtanova, the Work is a "comic book" consisting of eighteen pages, one of which is a cover. The cover page consists of an image of a young woman, the Work's title, and the words "Kashtanova" and "Midjourney." The remaining pages consist of mixed text and visual material. A reproduction of the cover page and the second page are provided below:



## II. SUMMARY OF ADMINISTRATIVE RECORD

On September 15, 2022, Ms. Kashtanova submitted an application for the Work and copies of each page of the Work as the deposit copy. In her application, Ms. Kashtanova listed the author of the Work as "Kristina Kashtanova" and stated that she had created a "[c]omic book." The application did not disclose that she used artificial intelligence to create any part of the Work, nor did she disclaim any portion of the Work.<sup>2</sup> The Office reviewed the application on the same day and registered the Work as registration number V Au001480196.

# Zarya of the Dawn (February 21, 2023)

"The changes to Zarya's mouth, particularly her upper lip, are **too minor and imperceptible** to supply the necessary creativity for copyright protection." – USCO Decision.

Detail before Photoshop



Detail after Photoshop





# Patentability

What is the current state of the law as to whether content created by an AI tool, or by a human using an AI tool, is patentable?



Like in the copyright context, courts have ruled that an "inventor" must be a **natural person**.  
See *Thaler v. Vidal*, 43 F.4th 1207 (Fed. Cir. 2022)

# Are the Prompts Used and the Output of the Tools Confidential?



User privacy is protected by a variety of laws (CCPA, GDPR, etc.).



As of today, Privacy Policy of OpenAI, StabilityAI and Midjourney say that they do not share *your personal information* with anyone else except as required by law or to improve their services.



As for *usage data* (i.e., inputs and outputs) – the policies and terms of use are generally silent.



Some financial institutions have restricted employees' access to ChatGPT over concerns about sensitive information being relayed to the AI platform.



Italy has banned ChatGPT nationwide, citing privacy concerns.



# Are the Prompts Used and the Output of the Tools Confidential?



TechRadar

+ Follow

## Samsung workers made a major error by using ChatGPT

The company allowed engineers at its semiconductor arm to use the [AI writer](#) to help fix problems with their source code. But in doing so, the workers inputted confidential data, such as the source code itself for a new program, internal meeting notes data relating to their hardware.

The upshot is that in just under a month, there were three recorded incidences of employees leaking sensitive information via ChatGPT. Since ChatGPT retains user input data to further train itself, these trade secrets from Samsung are now effectively in the hands of OpenAI, the company behind the AI service.



# Other Ethical Risks Creating Legal Liability?



## Biased data

Generative AI tools may inherit biases from the data used for training, or user input, which can result in biased or discriminatory output.



## Use of Names or Likeness of Others

Some generative AI tools are designed to generate realistic images of faces or identities, which can raise concerns about privacy and consent.



# How Human Biases Seep into AI Tools: Input Bias



AI Tools have human progenitors.



At various steps of the design process, a designer can embed his or her personal biases.



When a task for an AI tool is initially defined, bias can creep (or leap) in.



Examples:

- An AI tool that assumes an ideal employee must be a white, middle-aged male (Tasked with: "generate an image of a successful CEO").
- An AI tool that assumes crime is correlated with racially segregated neighborhoods (Tasked with: "write me a plot for a crime detective story set in New York City").



The background features a dark blue gradient with intricate, overlapping wavy lines in teal and purple. These lines are composed of many thin, parallel paths that create a sense of depth and movement. Interspersed among these lines are small, semi-transparent dots, some of which are connected by thin lines, forming a network-like structure. The overall aesthetic is modern and digital.

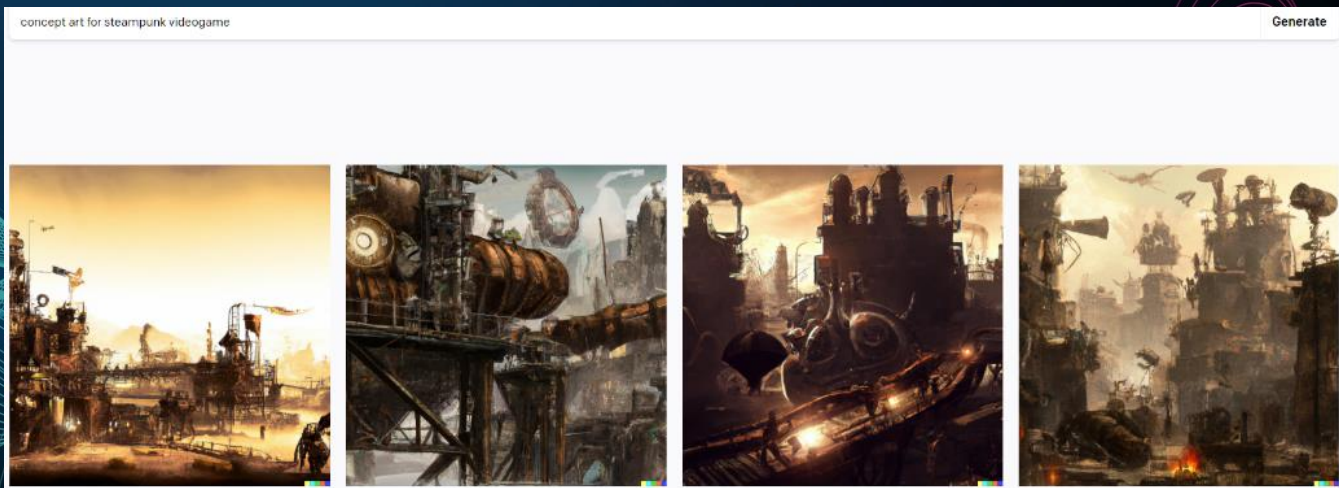
3

# Use Cases

# Use Cases



DALL-E for concept art, brainstorming, inspiration



# Use Cases



Various tools for AI-generated output/images for use in-game (final product).





# Use Cases



Stable Diffusion for real-time in-game asset delivery of 2D and 3D game assets generated in the cloud based on user preferences, game behavior, and user quizzes (players will not enter prompts, game will create assets based on whitelist or filter, or text generation model that determine prompts).





# Use Cases



ChatGPT to generate dialog and game code for use in-game (final product).

**JU** Come up with dialog where person A attempts to persuade person B not to travel back in time and prevent the Black Plague from occurring

**OpenAI** Person A: B, I'm begging you not to do this. Going back in time to prevent the Black Plague will only make things worse.  

Person B: But think of all the lives we could save! Millions of people died because of that disease.

Person A: I know, but changing the past can have unforeseeable consequences. The Black Plague played a pivotal role in shaping the course of human history. Preventing it from happening could have disastrous consequences.

Person B: What do you mean?

Person A: Well, for one thing, the Black Plague played a major role in the end of feudalism and the rise of modern capitalism. Without it, the course of European history could have been vastly different. It could have altered the balance of power between nations, changed the course of wars, and had a ripple effect on the entire world.

Person B: But we're talking about saving lives here. Surely that's more important than historical outcomes.

Person A: I understand your desire to do good, but there are other ways to help people without risking the stability of the timeline. And even if we were successful in preventing the Black Plague, who's to say that another disease wouldn't take its place?




Person B: I hadn't thought of that.

# Use Cases



GPT4 for generation of text for prototypes and experiments, brainstorming, and inspiration (similar to Googling names to come up with NPC names), generating training data for more specific text generators (e.g., generate 10,000 character descriptions and use that to train a character description generator) or for input for Stable Diffusion.

JU give me NPC names for a steampunk videogame

 1. Gideon Steel - a renowned inventor and engineer who creates advanced steam-powered gadgets and machines.  

2. Charlotte Copper - a skilled blacksmith who forges intricate gears and machinery, providing the player with weapons and equipment.

3. Ophelia Gearhart - a savvy businesswoman who owns a popular airship company, offering transportation to various destinations across the game world.

4. Augustus Brass - a wealthy industrialist who seeks to monopolize the production of steam power, creating a power imbalance that the player must navigate.

5. Percival Steamworth - a flamboyant artist who creates intricate mechanical sculptures and steam-powered art installations, providing a unique visual experience in the game world.

6. Thaddeus Clockwork - a reclusive inventor who has created a powerful automaton army, which the player must face in battle.

7. Lillian Copperfield - a skilled thief and smuggler who offers illegal contraband and rare steam-powered artifacts to the player, for a price.

8. Ezekiel Steamfist - a leader of a powerful faction that seeks to overthrow the ruling government, offering the player the opportunity to join their cause or oppose them.

9. Prudence Cogsworth - a brilliant scientist who has developed a new form of steam power that could revolutionize the industry, but is also coveted by rival factions and greedy corporations.

10. Winston Steamson - a jovial tinkerer who provides the player with various steam-powered gadgets and gizmos, from grappling hooks to jetpacks.

# Use Cases



Dall-E 2 as an image editing tool, to enhance or fill in details of low res images.



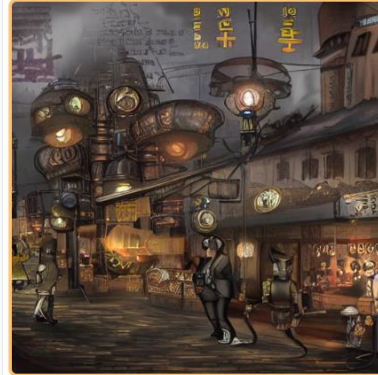
# Use Cases



Creation of art books to sell that includes game concept art created by AI tools.

concept art for steampunk videogame set in korea

Generate image



# Use Cases



Style transfer – apply "Lichtenstein style filter" in photo mode.





4

# License Terms to Watch Out For

# Who Owns the Output?

## Midjourney



<https://docs.midjourney.com/docs/terms-of-service>

- **Rights You Give to Midjourney**  
By using the Services, **You grant to Midjourney, its successors, and assigns a perpetual, worldwide, non-exclusive, sublicensable no-charge, royalty-free, irrevocable copyright license** to reproduce, prepare Derivative Works of, publicly display, publicly perform, sublicense, and distribute text, and image prompts **You input into the Services, or Assets produced by the service at Your direction.** This license survives termination of this Agreement by any party, for any reason.
- **Your Rights**  
**Subject to the above license, You own all Assets You create with the Services, to the extent possible under current law.** This excludes upscaling the images of others, which images remain owned by the original Asset creators. Midjourney makes no representations or warranties with respect to the current law that might apply to You. Please consult Your own lawyer if You want more information about the state of current law in Your jurisdiction. Your ownership of the Assets you created persists even if in subsequent months You downgrade or cancel Your membership. However, You do not own the Assets if You fall under the exceptions below.





# Who Owns the Output?

## Stability AI (Stable Diffusion)



<https://stability.ai/terms-of-use>; <https://platform.stability.ai/docs/terms-of-service>;  
<https://github.com/CompVis/stable-diffusion/blob/main/LICENSE>

- Unless otherwise expressly indicated by Stability, **all content displayed or made available on the Site**, including without limitation, text, images, illustrations, designs, logos, domain names, service marks, software, scripts, and the selection, compilation and arrangement of any of the foregoing **is owned by Stability**, its affiliates, licensors and/or other third parties ("Site Content").
- Under the **AI API Terms of Service**, the user owns of images created from the user's text prompts **"to the extent permitted by applicable law**. With respect to images you upload to the Services, ... you hereby grant Stability a ... **license** to copy, display, upload, perform, distribute, store, modify, and otherwise use such materials for any Stability-related purpose in any form, medium or technology now known or later developed.
- The **Creative ML Open-RAIL-M license terms grant a copyright and patent license to use the Stability Diffusion model and includes use-based restrictions** that must be included in any legal agreement under which the model is distributed or made available.



# Who Owns the Output?

## OpenAI (ChatGPT, GPT4, DALL-E)



<https://openai.com/policies/terms-of-use>

- Your Content. You may provide input to the Services (“Input”), and receive output generated and returned by the Services based on the Input (“Output”). Input and Output are collectively “Content.” **As between the parties and to the extent permitted by applicable law, you own all Input, and subject to your compliance with these Terms, OpenAI hereby assigns to you all its right, title and interest in and to Output. OpenAI may use Content as necessary to provide and maintain the Services, comply with applicable law, and enforce our policies. You are responsible for Content, including for ensuring that it does not violate any applicable law or these Terms.**
- Similarity of Content. Due to the nature of machine learning, Output may not be unique across users and the Services may generate the same or similar output for OpenAI or a third party. For example, you may provide input to a model such as “What color is the sky?” and receive output such as “The sky is blue.” Other users may also ask similar questions and receive the same response. **Responses that are requested by and generated for other users are not considered your Content.**
- Use of Content to Improve Services. We do not use Content that you provide to or receive from our API (“API Content”) to develop and improve our Services. API Content is only used to provide and maintain our API Services. We may use Content from Services other than our API (“Non-API Content”) to help develop and improve our Services. You can read more here about [how Non-API Content may be used to improve model performance](#). If you do not want your Non-API Content used to improve Services, you can opt out by filling out [this form](#). Please note that in some cases this may limit the ability of our Services to better address your specific use case.



# Liability and Indemnity

OpenAI (ChatGPT, GPT4, DALL-E)



## Indemnity

You will defend, indemnify, and hold harmless us, our affiliates, and our personnel, from and against any claims, losses, and expenses (including attorneys' fees) arising from or relating to your use of the Services, including your Content, products or services you develop or offer in connection with the Services, and your breach of these Terms or violation of applicable law.



<https://openai.com/policies/terms-of-use>

# Applicable Terms?

OpenAI (ChatGPT, GPT4, DALL-E)



These Terms of Use apply when you use the products and services of OpenAI, L.L.C. or our affiliates, including our application programming interface, software, tools, developer services, data, documentation, and website ("Services"). The Terms include our **Service Terms, Sharing & Publication Policy, Usage Policies, and other documentation, guidelines, or policies we may provide in writing**. By using our Services, you agree to these Terms. Our Privacy Policy explains how we collect and use personal information.



<https://openai.com/policies/terms-of-use>

# Usage Restrictions

OpenAI (ChatGPT, GPT4, DALL-E)



- You may not (i) use the Services in a way that infringes, misappropriates or violates any person's rights



<https://openai.com/policies/terms-of-use>

- We don't allow the use of our models for the following:
  - ... Any content that exploits or harms children
  - Generation of hateful, harassing, or violent content
  - ... Activity that has high risk of economic harm
  - Adult content, adult industries, and dating apps
  - ... Activity that violates people's privacy



<https://platform.openai.com/docs/usage-policies>

# Usage Restrictions

## Stability AI (Stable Diffusion)



<https://platform.stability.ai/docs/terms-of-service>



- May not:
  - Violate these Terms, the License, or other policies applicable to the Services;
  - Include sensitive personal information on the Site;
  - Violate the privacy, publicity, or other rights of any third party;
  - Generate any Content that a reasonable person could find offensive;
  - Upload any materials that infringe IP;
  - Use the Services in a way that violates any laws or regulations;
  - Attempt to do any of the above.

# Usage Restrictions

## Midjourney



- Be kind and respect each other and staff. Do not create images or use text prompts that are inherently disrespectful, aggressive, or otherwise abusive. Violence or harassment of any kind will not be tolerated.
- No adult content or gore. Please avoid making visually shocking or disturbing content. We will block some text inputs automatically.
- Sharing others' creations. Do not publicly repost the creations of others without their permission.
- Be careful about sharing. It's OK to share Your creations outside of the Midjourney community but please consider how others might view Your content.



5

# Risk of Violating Tool Use Restrictions



# Risk of Violating Tool Use Restrictions

OpenAI (ChatGPT, GPT4, DALL-E)



You acknowledge that if you violate or breach these Terms, it may cause irreparable harm to OpenAI and its affiliates, and OpenAI shall have the right to seek injunctive relief against you in addition to any other legal remedies.



We may terminate these Terms immediately upon notice to you if you materially breach Sections 2 (Usage Requirements), 5 (Confidentiality, Security and Data Protection), 8 (Dispute Resolution) or 9 (General Terms), if there are changes in relationships with third party technology providers outside of our control, or to comply with law or government requests.



# Risk of Violating Tool Use Restrictions

## Stability AI (Stable Diffusion)



Any violation of these terms (including the Community Guidelines) may result in the suspension or termination by us, in our sole discretion, of your access to and use of the Services (including a permanent ban and you will forfeit any remaining credits and fees paid for the services). We MAY CHANGE the Services in our sole discretion. we will endeavor to provide reasonable notice of any suspension, termination, or change to you, but may do so without notice if in our sole judgment, we need to make changes to comply with law, protect or enforce legal rights, avoid or mitigate misuse of the Services or adverse legal or regulatory impacts on us, or otherwise to address or prevent an emergency.



# Risk of Violating Tool Use Restrictions

## Midjourney



Any violations of these rules may lead to bans from our services. We are not a democracy. Behave respectfully or lose Your rights to use the Service.



We reserve the right to investigate complaints or reported violations of our Terms of Service and to take any action we deem appropriate including but not limited to reporting any suspected unlawful activity to law enforcement officials, regulators, or other third parties and disclosing any information necessary or appropriate to such persons or entities relating to user profiles, e-mail addresses, usage history, posted materials, IP addresses and traffic information.



We also reserve the right to terminate Your access to the Service for any reason, including for violation of the Community Guidelines or other inappropriate use of the Service.





6

# Overview of Current Lawsuits

# Class Action against Stability AI, Deviant Art, Midjourney

*Andersen v. Stability AI et al, N.D. Cal. No. 3:23-cv-00201*



In 2023, a group of artists initiated a class action against Stability AI, Midjourney, and Deviant Art, claiming that the company's use of their artwork as part of training data infringed on their copyrights.



Very similar to the Getty Images lawsuit described earlier.



Too early to tell who will win, but AI developers likely to win given vast breadth of training data.



Also hard for artists to prove how exactly they were harmed by the data scraping – difficult to quantify contribution of their work to output.

Case 3:23-cv-00201-WHO Document 1 Filed 01/13/23 Page 1 of 46

1 Joseph R. Saveri (State Bar No. 130064)  
Cadio Zirpoli (State Bar No. 179108)  
2 Christopher K. L. Young (State Bar No. 318371)  
Eliza A. Buchanan (State Bar No. 249996)  
3 Travis Manfredi (State Bar No. 281779)  
JOSEPH SAVERI LAW FIRM, LLP  
4 601 California Street, Suite 1000  
San Francisco, California 94108  
5 Telephone: (415) 500-6800  
Facsimile: (415) 395-9940  
6 Email: jsaveri@saverilawfirm.com  
czirpoli@saverilawfirm.com  
7 cyoung@saverilawfirm.com  
ebuchanan@saverilawfirm.com  
tmanfredi@saverilawfirm.com

8 Matthew Butterick (State Bar No. 250953)  
1920 Hillhurst Avenue, #406  
9 Los Angeles, CA 90027  
10 Telephone: (323) 968-2632  
Facsimile: (415) 395-9940  
11 Email: mb@buttericklaw.com

12 *Counsel for Individual and Representative  
Plaintiff: Sarah Andersen, Kelly McKernan,  
Karla Ortiz, and the Proposed Class*

13 [Additional Counsel Listed on Signature Page]

14  
15  
16 UNITED STATES DISTRICT COURT  
NORTHERN DISTRICT OF CALIFORNIA  
17 SAN FRANCISCO DIVISION

18 SARAH ANDERSEN, an individual,  
KELLY MCKERNAN, an individual,  
19 KARLA ORTIZ, an individual,

20 Individual and Representative Plaintiffs,

21 v.

22 STABILITY AI LTD., a UK corporation;  
STABILITY AI, INC., a Delaware  
23 corporation; MIDJOURNEY, INC., a  
Delaware corporation; DEVIANTART, INC.,  
24 a Delaware corporation,

25 Defendants.

Case No.

COMPLAINT

CLASS ACTION

26  
27  
28 DEMAND FOR JURY TRIAL

# GitHub Copilot

*Doe v. Github, Inc. et al., N.D. Cal.*  
4:22-cv-06823



Class action complaint was filed accusing Microsoft, GitHub, and OpenAI of scraping licensed code to build GitHub's AI-powered Copilot tool.



Very similar concept as Getty, and other AI-related litigation.



Maybe a weaker case – code on GitHub used to train Copilot is available to the public and *intended* to help developers write code by example.

1 Joseph R. Saveri (State Bar No. 130064)  
2 Cadjo Zirpoli (State Bar No. 179108)  
3 Travis Manfredi (State Bar No. 281779)  
4 JOSEPH SAVERI LAW FIRM, LLP  
5 601 California Street, Suite 1000  
6 San Francisco, California 94108  
7 Telephone: (415) 500-6800  
8 Facsimile: (415) 395-9940  
9 Email: jsaveri@saverilawfirm.com  
10 cszirpoli@saverilawfirm.com  
11 tmanfredi@saverilawfirm.com

12 Matthew Butterick (State Bar No. 250953)  
13 1920 Hillhurst Avenue, #406  
14 Los Angeles, CA 90027  
15 Telephone: (323) 965-2632  
16 Facsimile: (415) 395-9940  
17 Email: mb@buttericklaw.com

18 *Counsel for Individual and Representative*  
19 *Plaintiffs and the Proposed Class*

20 UNITED STATES DISTRICT COURT  
21 NORTHERN DISTRICT OF CALIFORNIA  
22 SAN FRANCISCO DIVISION

23 J. DOE 1 and J. DOE 2, individually and on  
24 behalf of all others similarly situated,  
25 Individual and Representative Plaintiff,  
26 v.

Case No.

COMPLAINT  
CLASS ACTION

DEMAND FOR JURY TRIAL

27 GITHUB, INC., a Delaware corporation;  
28 MICROSOFT CORPORATION, a Washington  
29 corporation; OPENAI, INC., a Delaware  
30 nonprofit corporation; OPENAI, L.P., a  
31 Delaware limited partnership; OPENAI GP,  
32 L.L.C., a Delaware limited liability company;  
33 OPENAI STARTUP FUND GP I, L.L.C., a  
34 Delaware limited liability company; OPENAI  
35 STARTUP FUND I, L.P., a Delaware limited  
36 partnership; OPENAI STARTUP FUND  
37 MANAGEMENT, LLC, a Delaware limited  
38 liability company,

Defendants.



7

# Ethical Practices and Ways to Mitigate Legal Risk

# Best Practices and Ways to Mitigate Legal Risk



Improve substantially upon the raw output produced by AI:

Show that your work has substantial human contributions.



Consult legal counsel:

Consult with a lawyer who is knowledgeable about IP law to help you understand the legal risks and how to mitigate them effectively.



Ensure no bias



Responsible and Ethical Practices



# Questions

The image features a large white speech bubble shape centered on a dark blue background. The word "Questions" is written in a bold, dark blue, sans-serif font inside the bubble. On the left side, there is a teal-colored abstract pattern of overlapping wavy lines and small dots. On the right side, there is a pink-colored abstract pattern of overlapping wavy lines and small dots. The overall design is modern and digital.

# Baker McKenzie.



Baker & McKenzie LLP is a member firm of Baker & McKenzie International, a global law firm with member law firms around the world. In accordance with the common terminology used in professional service organizations, reference to a "partner" means a person who is a partner, or equivalent, in such a law firm. Similarly, reference to an "office" means an office of any such law firm. This may qualify as "Attorney Advertising" requiring notice in some jurisdictions. Prior results do not guarantee a similar outcome.

© 2023 Baker & McKenzie LLP

[bakermckenzie.com](https://www.bakermckenzie.com)