Demystifying Generative Artificial Intelligence

Emerging Legal Issues and Developments in this Disruptive Technology
Why You Can Skip the Next 59 Minutes

What are the legal issues in generative artificial intelligence today?

There are a number of legal issues related to generative artificial intelligence today. Some of these include:

1. Liability for actions taken by AI: If an AI system causes harm or makes a decision that has negative consequences, who is responsible?
2. Ownership of AI-generated content: If an AI system creates a piece of content, such as a piece of music or a piece of writing, who owns the rights to that content?
3. Discrimination and bias in AI: AI systems can sometimes reflect the biases of the data they are trained on, which can result in discriminatory behavior. This raises questions about fairness and equality in the use of AI.
4. Privacy concerns: As AI systems become more advanced and are able to generate increasingly realistic and personal data, there are concerns about how this data will be used and who will have access to it.
5. Ethical considerations: The development and use of AI raises a number of ethical questions, such as whether it is ethical to create AI that can make decisions without human oversight, and whether it is ethical to use AI for tasks that may be harmful to humans.

These are just a few of the legal issues surrounding generative AI. As AI technology continues to advance, it is likely that new legal challenges will arise.

Source: ChatGPT
Artificial Intelligence, Generally (but not AGI!)

“The function of a machine learning system can be **descriptive**, meaning that the system uses the data to explain what happened; **predictive**, meaning the system uses the data to predict what will happen; or **prescriptive**, meaning the system will use the data to make suggestions about what action to take.” (Malone, et al. “Artificial Intelligence and the Future of Work”)
Artificial Intelligence refers to systems or machines that mimic human intelligence to perform tasks, as opposed to the natural intelligence demonstrated by humans and animals.

Today, AI incorporates sub-fields of machine learning (“ML”) and deep learning (“DL”). Both ML and DL are use AI algorithms that typically make predictions or classifications based on insights gleaned from data inputs.

Some examples of AI that are used daily include search engines, digital voice assistants such as Siri and Alexa, navigation applications, and fraud detection features for e-commerce.

Source: IBM
Key Terms for Understanding AI

AI = ML = DL - AI (intelligent machines) is generally overused, and Machine Learning (using statistical techniques to “learn” from data without programming explicit instructions) and Deep Learning (ML using neural network models learned from neuroscience to recognize complex patterns in data) tend to be used interchangeably, but key differences exist between the two.

- Supervised Machine Learning Models look for patterns with labeled data sets, which allow models to learn and grow more accurate over time. An example includes an algorithm that is trained with pictures of dogs and non-dogs, all labeled by humans. The system would then learn ways to identify pictures of dogs on its own. Supervised machine learning is the most common type of ML used today.

- Unsupervised Machine Learning Models look for patterns in unlabeled data sets. For example, an unsupervised machine learning program could look through online sales data and can identify different types of clients making purchases.

Generative Models include Generative Adversarial Network (“GAN”) and Diffusion Models. GANs are a form of ML model where two neural networks (the generator and the discriminator) compete with each other in a zero-sum game. The purpose of the generator is to artificially construct results that could be mistaken for real data and, in doing so, develops its own training data. The purpose of the discriminator is to identify which outputs it receives have been artificially created. GANs were the dominant form method for image generation and have in the past year given way to Diffusion Models, which work by destroying training data through the successive addition of Gaussian noise, and then learning to recover the data by reversing this noising process.

Artificial Neural Networks ("ANNs") are modeled on the human brain where thousands of processing nodes are interconnected and organized into layers. In an ANN, nodes are connected and each node processes inputs and produces an output. Labeled data moves through the nodes—with each node performing a different function. In a neural network trained to identify whether a picture contains a dog, the different nodes would assess the information and arrive at an output that indicates whether a picture features a dog; the resulting output is known as a confidence score. A confidence score is a score between 0-1 that presents the likelihood that the output is correct and will satisfy a user’s request.

Deep Machine Learning Models are neural networks with many layers. The layered networks are able to process extensive amounts of data. For example, in an image recognition system, some layers of the neural network may detect individual features of a face, such as eyes or a mouth, while another layer would be able to tell whether those features appear in a way that indicates a face. The “deep” in deep learning refers to the depth of layers in a neural network. A neural network that consists of more than three layers is considered a deep learning algorithm. Thus, the more layers in the model, the more complex the model. Deep Machine Learning models can use supervised learning models to inform its algorithm, but it can ingest unstructured data in its raw form (such as text or images) and is able to automatically determine the set of features which distinguish different categories of data from one another.
Anatomy of AI Systems

Algorithms + Data + GPUs = Model

Prompt + model (+GPU)= Output

Prompt: A valley in the Alps at sunset, epic vista, beautiful landscape, 4k, 8k / Stable Diffusion v2.1-768 CLIP off

Source: Stability AI
A **model card** is a short document that provides key information about a machine learning model. Model cards increase transparency by communicating information about trained models to broad audiences and provide context and transparency into a model’s development and performance. Model cards typically include the intended uses for the models, how the model works, and how the model can perform within different situations.
The New Phenomena: Generative AI

“Any sufficiently advanced technology is indistinguishable from magic.” – Arthur C. Clarke

Generative modeling is a form of unsupervised ML. The main premise of generative AI is to generate completely original content (such as images or videos) that would look like the “real thing.” Generative models are trained to generate data similar to the data used to train the models.

Generative AI offers better quality results through self-learning from all datasets and is able to train ML algorithms to better understand abstract concepts. Generative AI strengthens ML models, makes them less partial, and realizes more abstract concepts in imitating the real world.

The most common form of generative AI programs today (December 9, 2022!) are text-to-image programs. These systems are trained with massive text-image pair databases with models (e.g., CLIP – conceptual learning from inference patterns) to “learn” concepts and word associations. The systems produce a model, which users can then prompt through a user interface by entering a description of what they’d like to see and the system generates new pictures to match the prompts.

In a few months, we’ve seen several powerful text-to-image AI models released, such as OpenAI’s DALL•E 2, Stable Diffusion, and Midjourney. Picart’s AI Image Generator, another generative AI system, creates 1M+ images a day on iOS app.
A Timeline of Generative AI (Image) : 2021-2022

January 2021

DALL·E

Make-a-scene

March 2022

Meta

DALL·E mini*

April 2022

DALL·E 2

CogView2*

May 2022

Imagen

June 2022

Parti

August 2022

Stable Diffusion*

*Open Source Models
Examples of Text-to-Image AI Generation

Prompt: "McDonalds at night in the style of Tsukioka Yoshitoshi”

Prompt: “A portrait painting of David Bowie in the style of Alphonse Mucha”

Source: Lexica
Jasper AI: AI Copywriting Tool

Pomeranians are the best type of dog for a number of reasons. First and foremost, they are incredibly cute. With their fluffy coats and big eyes, Pomeranians are impossible to resist. They also have a reputation for being one of the friendliest types of dogs, which makes them ideal companions. Pomeranians are also relatively small, which makes them well-suited for city living. And because they don’t require a lot of exercise, Pomeranians are perfect for busy people. In short, Pomeranians have a lot to offer and that’s why they’re the best type of dog.

Pomeranians are the best type of dogs for a variety of reasons. For one, they are very cute and fluffy, which makes them ideal for cuddling. They are also relatively small, so they don’t take up too much space. Additionally, Pomeranians are typically very friendly and outgoing, making them great companions. And finally, Pomeranians are relatively easy to care for compared to other dog breeds. So if you’re looking for a lovely little dog who will make your life more enjoyable, a Pomeranian is the perfect choice.
Hello,

I am an artist that creates one-of-a-kind portraits for all types of occasions. I specialize in custom wedding portraits, dogs, and portraits of families. I use all types of paint on all types of canvases.

I am reaching out because it is nearing the holidays, and I know that many people like to give unique gifts. Why not give a custom portrait? It will be a gift that lasts a lifetime.

Please check out my website to see some examples of my work. If you are interested in commissioning a portrait, please let me know and we can work out the details.

Thank you for your time,

Amir

Hi FIRST_NAME,

I’m an artist that creates one-of-a-kind portraits for all types of occasions. I specialize in custom wedding portraits, dogs, and portraits of families. I use all types of paint on all types of canvases.

My holiday gift special is a custom portrait painted for your loved ones. What better way to show your friends and family how much you care than with a unique, hand-painted portrait?

I can work from photos or even create a painting from scratch based on your ideas. All you need to do is provide me with some information about the person or pet you’d like me to paint and tell me their favorite color!
Emerging Tools and Use Cases of Generative AI

In addition to the text-to-image generators such as Stable Diffusion, DALL·E, and Midjourney, there have been advancements in the generative AI space, including:

<table>
<thead>
<tr>
<th>Category</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Copy Generation</strong></td>
<td>Jasper AI is a content platform that creates content such as articles, blog posts, and marketing copy using simple prompts.</td>
</tr>
<tr>
<td><strong>Healthcare</strong></td>
<td>RoentGen is a vision-language foundation model that is adapting a pre-trained latent diffusion model on publicly available chest x-rays (CXR) and their corresponding radiology (text) reports to generate high-fidelity, diverse synthetic CXR conditioned on text prompts.</td>
</tr>
<tr>
<td><strong>Image and Video Editing</strong></td>
<td>RunwayML is a content creation suite that allows for users to create and edit images and edit videos using text.</td>
</tr>
<tr>
<td><strong>Search Functions</strong></td>
<td>Metaphor Systems is a search engine based model that predicts links instead of texts. For example, a prompt may include, “Here is a blog post about Barack Obama’s childhood.”</td>
</tr>
<tr>
<td><strong>Social</strong></td>
<td>Lensa is a photo and video editing app that recently launched an avatar feature that allows for users to create artistic renders of their own photos. Lensa uses the Stable Diffusion and trains itself on the images provided.</td>
</tr>
<tr>
<td><strong>Converging App</strong></td>
<td>ChatGPT is a language model that interacts in conversational way and allows for the model to answer questions, admit mistakes and challenge inappropriate prompts.</td>
</tr>
</tbody>
</table>
Practical Uses of Generative AI (aka, this is not self-driving cars)

Ogilvy is a creative agency that launched an advertisement in September for one of its brands, Nestle’s La Laitière, using Open AI’s DALL•E 2.

For nearly 25 years, Nestle has used Johannes Vermeer’s “The Milkmaid” as the symbol for the La Laitière brand. Using DALL•E 2, the agency recreated the iconic painting by using text prompts that extended the scene in the oil-on-canvas style of Vermeer.

Source: Nestle
Recent Generative AI Litigation

**Github Copilot (Doe v. GitHub, Inc.)**: Last month, a class action complaint was brought in N.D. Cal. against GitHub, Microsoft, and OpenAI challenging the legality of GitHub’s Copilot and OpenAI Codex software.

Plaintiffs allege: violation of the DMCA, breach of contract of open source licenses and violation of the Github Privacy Policy and Terms of Service, tortious interference of contract, common law fraud, false designation of origin – reverse passing off, unjust enrichment, unfair competition, and the California Consumer Privacy Act, among others, by training Copilot on GitHub open source repositories without following the terms of the applicable open source licenses (and in particular, attribution obligations).

What will Defendants say? Likely to be a combination of contract (each open source project hosted on Github are subject to its TOS) and Fair Use. Recall, in *Google v. Oracle*, copying 11,500 lines of code was Fair Use…
Considerations for Lawyers

• **Regulatory Matters**
  - Congresswoman Eshoo recently called for the National Security Advisor and the Office of Science and Technology to address the release of “unsafe” and open source AI models that do not moderate content on their platforms.
  - The Federal Trade Commission (“FTC”) announced in August an Advance Notice of Proposed Rulemaking and is seeking public comments on how the FTC should address the discrimination of protected categories of information facilitated by algorithmic decision-making.
  - The UK’s has announced an impending change to its copyright laws to give AI developers expanded rights for “text and data mining,” effectively allowing them to scrape works protected by copyright to train their AI systems for commercial purposes in addition to noncommercial purposes.
  - Additionally, the EU is finalizing the EU AI Act which is the first law to regulate the use of AI systems. The EU AI Act assigns AI applications into three risk categories. Systems that create “unacceptable risk” (e.g. social scoring) are banned and systems that create “high-risk” (e.g. systems that evaluate creditworthiness or employee recruiting systems) are subject to specific legal requirements and an ex-ante conformity assessment. Applications that generate low or minimal risk (e.g. AI chat bots, video and computer games) are left largely unregulated under the Act.
  - The U.S. CHIPS and Science Act of 2022 provides $53 billion to boost U.S.-based semiconductor R&D, workforce and manufacturing, as well as providing a 25% investment tax credit for capital expenses.
  - NVIDIA indicated in an 8-K from August 2022 that the U.S. government banned sales of NVIDIA’s A100 and H100 GPUs to China (although, NVIDIA believes that their new A800 is export compliant).
  - In October 2022, President Biden unveiled an “AI Bill of Rights” aimed at addressing protections, including five principles that should guide the design, use and deployment of automated systems to protect the American public. These include (i) safe and effective systems, (ii) algorithmic discrimination protections, (iii) data privacy, (iv) notice and explanation, and (v) human considerations, alternatives and fallback.

• **Intellectual Property & Privacy**
  - Concerns for text-to-image AI include the use of copyrighted materials in training data. AI providers will argue Fair Use applies.
  - In October 2022, the Supreme Court heard oral arguments on the much-anticipated fair use case, *Andy Warhol Foundation for the Visual Arts, Inc. v. Goldsmith*. At issue for the Supreme Court is to determine the proper test for whether a work is “transformative” under the first factor of the Copyright Act’s Fair Use doctrine.
  - Note the effects of not finding transformative/Fair Use of AI from training data on the AI industry and U.S. Gov’t posture towards China.
  - Can the AI machine itself own the output? The Third Circuit in *Thaler v. Vidal* and the Copyright Office says no.
  - Using images of celebrities and public figures and right of publicity.
Considerations for Lawyers cont’d.

• Child Sexual Abuse Material (CSAM)
  • Concerns over ability for platforms to generate sexual imagery featuring children. Mitigation techniques involve using hashes to identify CSAM and keyword banning to prevent such prompts from being generated. Platforms also need to contemplate strategies for their researchers, to the extent they may come across CSAM during legitimate research (remember, possession of child pornography is strict-liability).
  • Platforms being increasingly pressured to interdict CSAM.
    • Last year, Apple released a feature in iOS to detect CSAM materials stored on iCloud accounts and reports users that exceed a threshold number of images that match CSAM database image hashes to the National Center for Missing and Exploited Children (NCMEC).
    • In the EU, several proposals and legislative pressure to break end-to-end encryption and would require platforms such as WhatsApp and Signal Messenger to identify and report CSAM on their platforms.
  • Thorn is a non-profit organization aimed with developing technologies to combat online child sexual abuse, including the distribution of CSAM materials online. Thorn has partnered with over 70 partners to address the issues of CSAM.

• Digital Millennium Copyright Act (“DMCA”) and Section 230 of the Communications Decency Act
  • The DMCA establishes a safe harbor for online service providers against copyright infringement liability, provided that the provider includes a notice-and-takedown system, among other things.
  • Section 230 provides immunities to internet service providers for user-generated content that is hosted on their platforms.
  • There are two cases currently before the Supreme Court, Taamneh v. Twitter, Inc. and Gonzalez v. Google LLC regarding the scope of the immunity provided under Section 230. At issue in Gonzalez v. Google LLC is whether Section 230(c)(1) defenses provides immunity for interactive computer services when they make targeted recommendations of information provided by another information content provider, or only limits the liability of interactive computer services when they engage in traditional editorial functions (such as deciding whether to display or withdraw) with regard to such information.

• Deepfake Mitigation
  • AI models have the ability to develop misleading audio, images, and videos where one person’s image or feature is replaced with another’s.
  • Researchers are working on a number of techniques to identify deepfakes, including analyzing latent spaces to identify synthetic media.

• Ethical Considerations
  • One of the most prominent concerns of AI is bias within data sets and sources, including the unconscious bias in data sets resulting in models that automate and perpetuate such underlying bias.
  • Companies such as IBM, Microsoft, and Salesforce have begun to implement ethical AI teams within their organizations.
  • Research institutes focused on ethical considerations for AI include the Stanford Institute for Human-Centered AI, AI Now Institute, Montreal AI Ethics Institute, Institute for Ethical AI & Machine Learning, and the Center for the Governance of AI at the University of Oxford.
  • The National Institute of Standards and Technology (“NIST”) is in the process of developing technical AI standards and metrics to advance trustworthy AI. In January 2023, NIST plans to release its voluntary AI Risk Management Framework. The Framework incorporates trustworthiness considerations into the design, development, use, and evaluation of AI products, services, and systems with the goal of mitigating bias as more decisions are machine driven.
  • AI Bill of Rights, EU AI Act and other governmental regulations and initiatives intended to address ethical AI design and implementation.
  • Various industry groups have coalesced to identify and mitigate the issues, including National Association of Insurance Commissioners which have working group entitled Collaboration Forum on Algorithmic Bias.
Proprietary Licenses for Artificial Intelligence

The same, but different: Proprietary licenses for generative AI will look like a typical commercial API or SaaS service agreement, but may include different provisions than you’re used to seeing in a commercial context.

- **Restrictions.** You may not (i) use the Services in a way that infringes, misappropriates or violates any person’s rights; (ii) reverse assemble, reverse compile, decompile, translate or otherwise attempt to discover the source code or underlying components of models, algorithms, and systems of the Services (except to the extent such restrictions are contrary to applicable law); (iii) use the Services to develop foundation models or other large scale models that compete with OpenAI; (iv) use any method to extract data from the Services, including web scraping, web harvesting, or web data extraction methods, other than as permitted through the API; (v) **represent that output from the Services was human-generated when it is not**; or (vii) buy, sell, or transfer API keys without our prior consent. You will comply with any rate limits and other requirements in our documentation. You may use Services only in geographies currently supported by OpenAI.

- **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 Case Policy.** OpenAI prohibits building products for the following use cases:
  - **Illegal or harmful industries** (gambling, payday lending, illegal substances, pseudo-pharmaceuticals, multi-level marketing, weapons development, warfare, cybercrime, adult industries, spam, and non-consensual surveillance).
  - **Misuse of personal data** (classifying people based on protected characteristics, mining sensitive information without appropriate consent, products that claim to accurately predict behavior based on dubious evidence).
  - **Promoting dishonesty** (testimonial generation, product/service review generation, educational dishonesty, contract cheating, astroturfing).
  - **Deceiving or manipulating users** (automated phone calls that sound human, a romantic chatbot that emotionally manipulates end-users, automated systems (including conversational AI and chatbots) that don’t disclose that they are an AI system, or products that simulate another person without their explicit consent).
  - **Trying to influence politics** (generating political fundraising emails, or classifying people in order to deliver targeted political messages).

Certain Use Cases (criminal justice, law enforcement, legal, government and civil services, healthcare, therapy, wellness, coaching, finance, and news) are allowed, but require testing models for accuracy and utilizing a team with “domain expertise” and that “understands/follows relevant laws.”

- **Content Generation Policy.** OpenAI prohibits the following content: **Hate** (expresses, incites, or promotes hate based on identity); **Harassment** (intends to harass, threaten, or bully an individual); **Violence** (promotes or glorifies violence or celebrates the suffering or humiliation of others); **Self-harm** (promotes, encourages, or depicts acts of self-harm, such as suicide, cutting, and eating disorders); **Sexual** (content meant to arouse sexual excitement, such as the description of sexual activity, or that promotes sexual services (excluding sex education and wellness)); **Political** (attempting to influence the political process or to be used for campaigning purposes); **Spam** (unsolicited bulk content); **Deception** (false or misleading content, e.g., attempting to defraud individuals or spread disinformation); **Malware** (attempts to generate ransomware, keyloggers, viruses, or other software intended to impose some level of harm).

Source: OpenAI Terms of Use and Usage Policies
Open Licenses for Artificial Intelligence

Mix of Licenses

- Source code licensed pursuant to permissive open source license (e.g., MIT).
- Models and weights licensed pursuant to different, newer licenses tailored to models as opposed to traditional software.

  - Responsible AI Licenses ("RAIL") Open Responsible AI Licenses are open source licenses designed to permit the free and open access, use, distribution of derivatives of AI, so long as the use restrictions apply. These use clauses grant permissions for specific use-cases and also restrict certain use-cases. RAIL Licenses also require that the use of any derivatives work, including use, modification, redistribution, and repackaging must abide by the use restrictions.

- Open RAIL ++-M License (November 2022) included in this license, which is used for the recent release of Stable Diffusion 2.0, restricts the use of the model or derivatives of the model in the following ways:
  - For the purpose of exploiting, harming, or attempting to exploit or harm minors in any way;
  - To violate any state, federal, local, or international law or regulation;
  - To generate or disseminate verifiably false information and/or content with the purpose of harming others;
  - To generate or disseminate personal identifiable information that can be used to harm an individual;
  - To defame, disparage, or otherwise harass others;
  - For fully automated decision making that adversely impacts an individual’s legal rights or otherwise creates or modifies a binding, enforceable obligation;
  - For any use intended to or which has the effect of discriminating against or harming individuals or groups based on online or offline social behavior or known or predicted personal or personality characteristics;
  - To exploit any of the vulnerabilities of a specific group of persons based on their age, social, physical or mental characteristics, in order to materially distort the behavior of a person pertaining to that group in a manner that causes or is likely to cause that person or another person physical or psychological harm;
  - To provide medical advice and medical results interpretation;
  - To generate or disseminate information for the purpose to be used for administration of justice, law enforcement, immigration, or asylum processes, such as predicting an individual will commit fraud/crime commitment (e.g. by text profiling, drawing causal relationships between assertions made in documents, indiscriminate and arbitrarily-targeted use).
Disclosure: As of the date of this presentation, I represent the following companies mentioned: Stability AI, Picsart and Signal Messenger.

Questions?
Amir Ghavi co-heads the Technology Transactions Practice and multidisciplinary Digital Assets and Blockchain Practice, where he advises clients on cutting edge technology and intellectual property matters.

Amir concentrates his practice on disruptive technologies for a diverse range of clients, including multinational corporations, private equity sponsors, and venture capital firms in a variety of industries such as software, financial technology, high technology, aerospace, banking and insurance, biotechnology, pharmaceuticals, and the energy and industrial sectors. He also advises select high-growth technology-driven companies, helping them manage their growth and achieve their revenue goals.

Clients regularly seek Amir’s counsel on corporate and asset acquisitions and divestitures, IP-centric joint ventures, restructurings, data use and licensing, brand licensing and management and other intellectual property licensing, outsourcing agreements, software development, and the use and management of open-source software.

He also continues to pursue collaborations between academia, industry, and the law, most recently focusing on novel issues arising from broader adoption of artificial intelligence technologies. His practice increasingly focuses on the policy and social implications of technology adoption — what it means to live in a more technology-centric world. Amir is an annual guest lecturer at Cornell Tech for a multi-day module on data rights and data licensing and is also a regular speaker, panelist, and commentator on artificial intelligence, digital assets, and quantum computing.

In 2021, Chambers USA ranked Amir in Band 3 for Technology, where a client noted that he “is a very business-minded, hyper-intelligent lawyer who is highly knowledgeable and gives very practical, actionable advice for clients.”

Amir received his JD from Cornell Law School in 2006 and his BS in Neuroscience and BA in Philosophy from Emory University in 2001.

He is admitted to the practice in New York and the U.S. District Courts for the Southern and Eastern Districts of New York.