



Nishith Desai Associates
LEGAL AND TAX COUNSELING WORLDWIDE

MUMBAI

SILICON VALLEY

BANGALORE

SINGAPORE

NEW DELHI

MUNICH / AMSTERDAM

NEW YORK

GIFT CITY

Research

Generative AI & Disruption

Emerging Legal and Ethical Challenges

May 2023

Poem on Nishith Desai Associates



A firm of renown,
Nishith Desai Associates,
Advising on law,

Both national and international,
Their team, expert and astute,
Guiding Clients with utmost repute.

In M&A and private equity,

Their counsel is sound and true,

In tax and regulatory,

Their expertise is clear and true.

A Poem on Nishith Desai Associates — by ChatGPT

January 22, 2023



<https://chat.openai.com/chat>



Poem on Nishith Desai Associates _



A firm of renown,
Nishith Desai Associates,
Advising on law,
Both national and international,
Their team, expert and astute,
Guiding Clients with utmost repute.

In M&A and private equity,
Their counsel is sought with great frequency,
In tax and regulation,
Their advise is top of the nation,

In intellectual property,
Their expertise is unparalleled surely,
In litigation and arbitration,
Their success rate is commendable and true.

With offices in multiple cities,
They serve clients with great ease,
Nishith Desai Associates,
A name to remember with great please.

*January 22, 2023. ChatGPT was asked to compose a poem on Nishith Desai Associates.
In a split second, it presented us with this delightful response.*

Research

Generative AI & Disruption

Emerging Legal and Ethical Challenges

May 2023



Ranked as the 'Most Innovative Indian Law Firm' in the prestigious FT Innovative Lawyers Asia Pacific Awards for multiple years. Also ranked amongst the 'Most Innovative Asia Pacific Law Firm' in these elite Financial Times Innovation rankings.



Disclaimer

This report is a copyright of Nishith Desai Associates. No reader should act on the basis of any statement contained herein without seeking professional advice. The authors and the firm expressly disclaim all and any liability to any person who has read this report, or otherwise, in respect of anything, and of consequences of anything done, or omitted to be done by any such person in reliance upon the contents of this report.

Contact

For any help or assistance please email us on concierge@nishithdesai.com or visit us at www.nishithdesai.com.

Acknowledgements

Vaibhav Parikh

vaibhav.parikh@nishithdesai.com

Huzefa Tavawalla

huzefa.tavawalla@nishithdesai.com

Abhay Porwal

abhay.porwal@ipproinc.com

Kartikeya Asthana

kartikeya.asthana@nishithdesai.com

Aniruddha Majumdar

aniruddha.majumdar@nishithdesai.com

Raashi Vaishya

raashi.vaishya@nishithdesai.com

Faiza Khanum

faiza.khanum@nishithdesai.com

Rhythm Vijayvargiya

rhythm.vijayvargiya@nishithdesai.com

Contents

| | |
|---|-----------|
| Introduction | 1 |
| Content Generation through Generative Artificial Intelligence | 2 |
| A. Text | 3 |
| B. Images | 6 |
| C. Sounds and Audios | 7 |
| D. Videos | 7 |
| E. 3D Models | 8 |
| The Positives: Groundbreaking Use Cases and Benefits | 9 |
| A. Research and Education | 9 |
| B. Industrial Application | 10 |
| C. Programming and Design | 10 |
| D. Customer Support | 11 |
| E. Entertainment | 12 |
| F. Healthcare and Medicine | 12 |
| Concerns and Disruption of Regulations | 14 |
| A. The Veracity of Information and Fake News | 14 |
| B. Generative Artificial Intelligence and Intellectual Property Rights | 14 |
| C. Data Protection and Confidentiality | 19 |
| D. Bias, Prejudice and Ethics | 21 |
| E. Explainability and Accountability of Artificial Intelligence | 23 |
| F. Content Regulation | 24 |
| G. Generative Artificial Intelligence Models: Intermediaries or Publishers? | 27 |
| H. Competition Law Concerns | 28 |
| I. Misuse Cases of Generative Artificial Intelligence | 29 |
| Way Forward | 32 |

Introduction

As a result of the fast-paced advancement in technology, Artificial Intelligence (“AI”) has emerged as a powerful tool for content creation in the last few decades. AI is essentially the method of making machines mimic human intelligence, and Generative AI (“GAI”) is its branch that uses machine learning technology for the generation of new content.¹

GAI uses unsupervised learning algorithms to create new audio, visual, and textual media based on user prompts.² One of the key advantages of GAI is its autonomous functioning, which allows it to process large quantities of data quickly and efficiently, enabling faster content creation.³ GAI has tremendous potential in a variety of applications, such as video game design,⁴ marketing campaigns,⁵ and the generation of software codes.⁶ With its ability to generate novel content from user prompts, GAI has become an attractive option for businesses seeking to produce unique and engaging content.

Unlike other AI models that primarily classify, analyze, or act on existing data, GAI goes beyond and creates new content every time it is prompted.⁷ This makes GAI a valuable tool for content creators, as it can provide them with an endless stream of fresh ideas and material. Additionally, with the increase in computing power and the development of more efficient deep learning networks, GAI models can now generate novel content in every new trial even from the same prompts.⁸

Open AI, a leading AI research and deployment company, has played a prominent role in this sphere. In 2021, it captured media attention with its text-to-image converter, Dall-E.⁹ Dall-E can understand complicated and contextualized prompts, producing accurate and realistic images and art. In late 2022, Open AI once again dazzled the world with its latest offering, the Chat Generative Pre-Trained Transformer (“ChatGPT”).¹⁰

Similarly, Google’s Bard is an AI chatbot that utilizes the Language Model for Dialogue Applications (LaMDA) technology. It is capable of engaging in human-like conversations and providing contextual responses using natural language processing and machine learning.¹¹ Although it was announced in February 2023, Bard is not yet available to the public.¹²

In this paper, we have systematically analyzed the types of content generated through GAI, its use cases and benefits along with the key legal concerns this technology brings with its advancement. We have also discussed the way forward to accommodate GAI keeping in mind the legal and ethical considerations attached with it.

1 See: <https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-is-generative-ai>, last visited on April 3, 2023.

2 See: <https://www.leewayhertz.com/generative-ai-use-cases-and-applications/>, last visited on April 3, 2023.

3 See: <https://www.thedigitalspeaker.com/what-is-generative-ai-how-disrupt-society/>, last visited on April 3, 2023.

4 See: <https://www.reuters.com/technology/unity-aims-open-generative-ai-marketplace-video-game-developers-2023-03-22/>, last visited on April 3, 2023.

5 See: <https://www.cxotoday.com/press-release/contlo-launches-worlds-first-brand-contextual-generative-ai-model/>, last visited on April 3, 2023.

6 See: <https://www.marktechpost.com/2023/03/02/top-generative-ai-tools-in-code-generation-coding-2023/>, last visited on April 3, 2023.

7 See: <https://arxiv.org/pdf/2301.04655.pdf>, last visited on March 29, 2023.

8 See: <https://arxiv.org/pdf/2301.04655.pdf>, last visited on March 29, 2023.

9 See: <https://openai.com/product/dall-e-2>, last visited on April 3, 2023.

10 See: <https://openai.com/blog/chatgpt>, last visited on April 3, 2023.

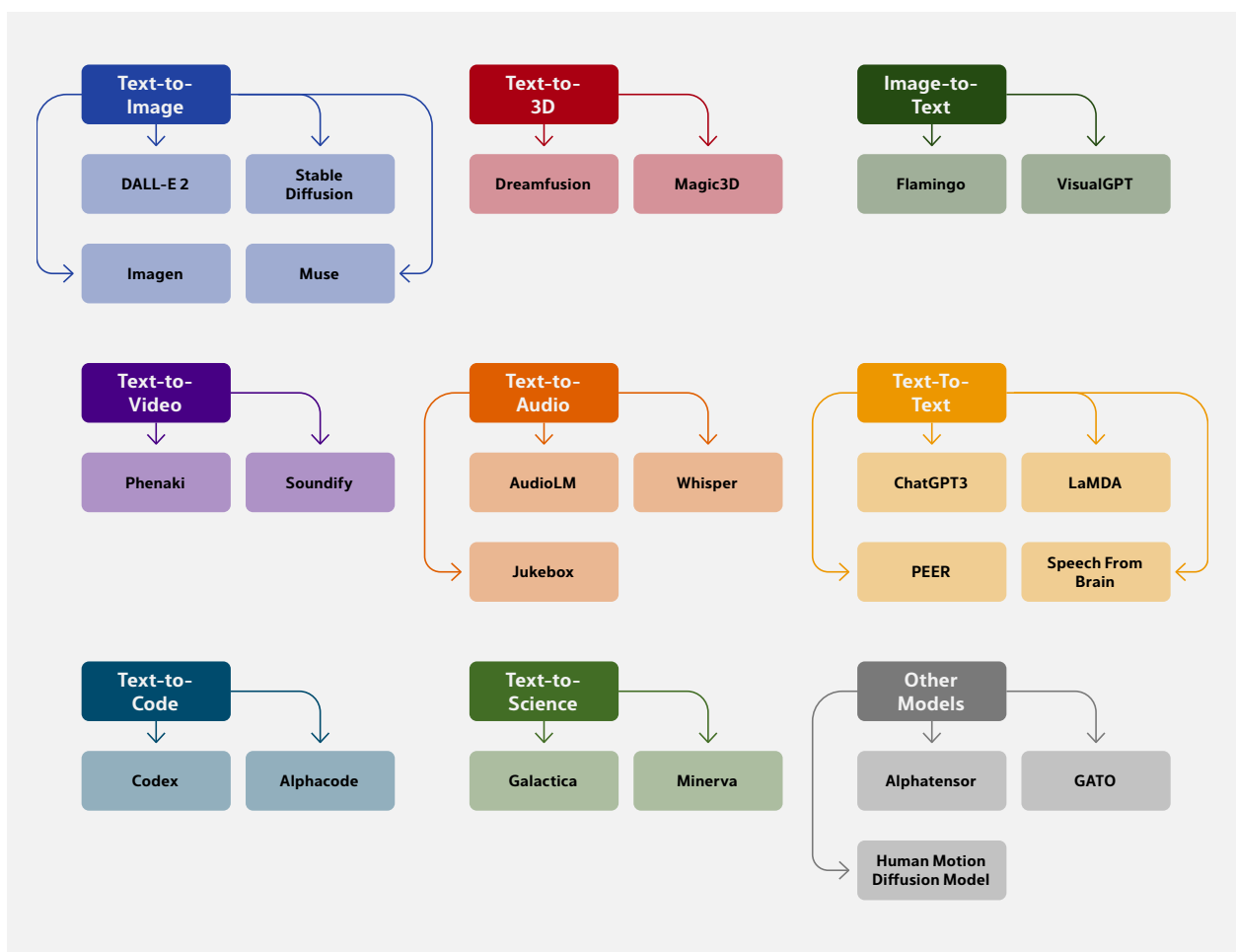
11 See: <https://bard.google.com/>, last visited on April 7, 2023.

12 On March 21, 2023, Google opened up its Bard waitlist, granting access to limited users in UK and US on a rolling basis.

Content Generation through Generative Artificial Intelligence

GAI models can be trained to generate novel content in an array of forms, like text, images, videos, audios, codes, and 3D models. The GAI platforms used to generate such content learn from a large amount of input data, allowing them to produce results that are often indistinguishable from human-created content.

Below is a flowchart that depicts how different inputs generate different kinds of content on GAI based platforms:



Source: Brizuela & Garrido, *ChatGPT is not all you need. A State of the Art Review of large GAI models*¹

¹ See: <https://arxiv.org/pdf/2301.04655.pdf>, last visited on March 29, 2023.

2. Content Generation through Generative Artificial Intelligence

We have also explained various forms of the content generated through GAI below:

A. Text

ChatGPT is one of the models that can generate coherent and human-like text with unprecedented fluency and accuracy. It has a wide range of capabilities ranging from prose and poetry writing, search, text summarization, and image and code generation.

ChatGPT is a Large Language Model (“LLM”). An LLM is a mathematical model of the statistical distributions of phrases, words, parts of words, and even individual characters, over a corpus of human-generated text. Thus, for example, when a user prompts “Knock Knock”, the LLM deciphers that there is a very high probability of “Who’s there?” following the initial prompt. ChatGPT does not understand what is being asked from it and what result it produces; it merely generates the most statistically likely sequence of terms.

Although ChatGPT may not have the originality of thought or creativity, its results are compelling enough to be that of a human author. It has solid potential to be used as a collaborative tool in many sectors, such as research and education, programming and design, customer service, therapy, clinical documentation, spam detection, HR processes, coding, legal sector, sentiment analysis, etc.

Natural Language Processing

GAI in text/speech form largely belongs to the realm of Natural Language Processing (“NLP”) that brings together computer science, AI, information engineering, and linguistics. NLP refers to the deployment of computational techniques to understand and manipulate text or speech to serve useful purposes such as translation, answering questions, and text summarisation.²

However, the biggest challenge of NLP is natural language *understanding*. Take the following examples:

- **“I saw the Empire State Building flying into New York.”** (Is the building flying?)
- **“I crossed the river to reach the other bank.”** (Does the ‘bank’ refer to a river embankment or a financial institution?)
- **“Arjun is better than Ramesh as a barista, even though he is ten years younger”** (Who does the pronoun ‘he’ refer to in both sentences?)

As humans, we can quickly and effortlessly resolve ambiguities because of our daily conversations, culture, reading, and ability to discern context. We cannot say the same for machines. They need to be taught, and that too, in the language they understand. Unsurprisingly, this is a difficult task and comprises three main challenges relating to:

- The writer’s or speaker’s thought process
- Representation and meaning of the linguistic input
- World Knowledge³

² See: https://link.springer.com/chapter/10.1007/978-981-15-9712-1_31, last visited on March 29, 2023.

³ See: <https://strathprints.strath.ac.uk/2611/1/strathprints002611.pdf>, last visited on March 29, 2023.

2. Content Generation through Generative Artificial Intelligence

In other words, *“an NLP system may begin at the word level — to determine the morphological structure, nature (such as part-of-speech, meaning), etc. of the word — and then may move on to the sentence level — to determine the word order, grammar, meaning of the entire sentence, etc. — and then to the context and the overall environment or domain.”*⁴

In the NLP domain, these neural networks work sequentially by predicting the next word given the previous words. However, because of their sequential operation, they may struggle with large sentences and paragraphs.

Transformer Architecture

In 2017, a team of Google researchers published a seminal paper titled “Attention is all you need” that laid the basis for the neural network that GAI models use — called the ‘Transformer’ (the T in GPT). Two key innovations of transformer architecture are the ‘attention’ and ‘self-attention’ mechanism.

I. Attention

The attention mechanism allows a model to look at every word in the ‘input’ sequence simultaneously while making predictions for the ‘output’ sequence. This is especially important in machine translation, for example, in languages with gendered words or where the order of the words is flipped. Let’s take the following example sentence, as used in the 2014 paper⁵ which first demonstrated the concept:

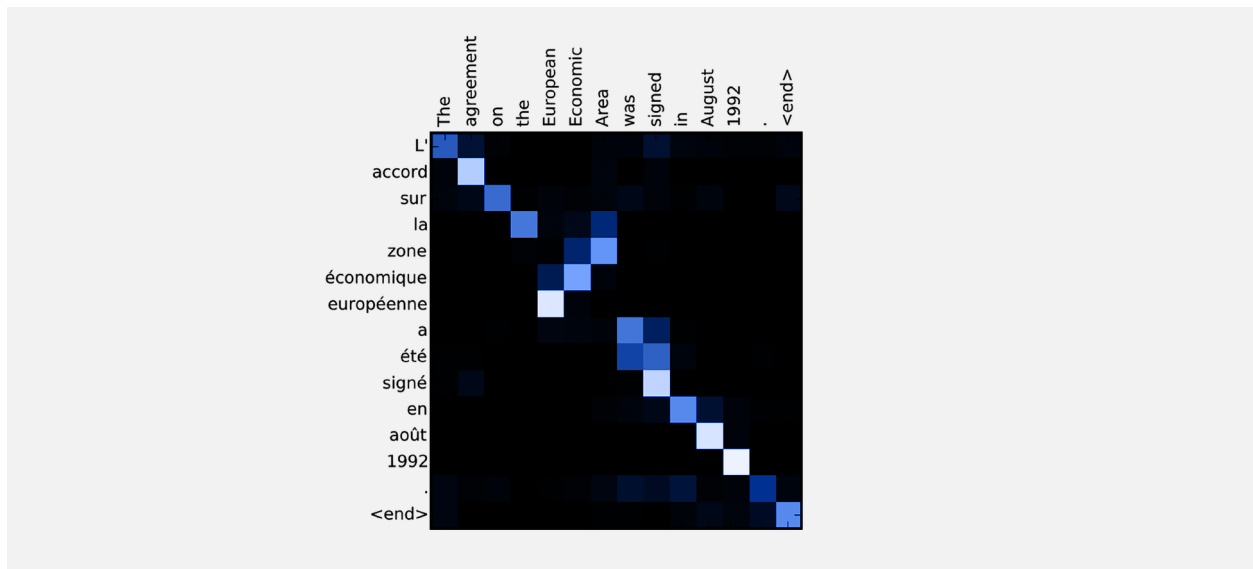
When the sentence “The agreement on the European Economic Area was signed in August 1992” is translated into French, it would read as: *“L’accord sur la zone économique européenne a été signé en août 1992.”*

Firstly, the words in the phrase “European Economic Area” are flipped to read as “la zone économique européenne”. Secondly, the adjectives “économique” and “européenne” are feminized to match the feminine object “la zone”.

4 See: <https://strathprints.strath.ac.uk/2611/1/strathprints002611.pdf>, last visited on March 29, 2023.

5 See: <https://arxiv.org/abs/1409.0473>, last visited on March 29, 2023.

2. Content Generation through Generative Artificial Intelligence



Source: *Neural Machine Translation by Jointly Learning to Align and Translate* (2014)

The above heat map demonstrates what words the model is “attending” to in the input sequence when it outputs each word. As you might expect, when the model outputs the word “européenne,” it’s attending considerably to both the input words “European” and “Economic.”

And how does the model know which words it should be “attending” to at each time step? It’s something that’s learned from training data. By seeing thousands of examples of French and English sentences, the model learns what types of words are interdependent. It learns how to respect gender, plurality, and other rules of grammar.⁶

II. Self-attention

While attention models are important for tasks such as machine translation, ‘self-attention’ is useful for disambiguation, that is, discerning the context and meaning of words and sentences. Unlike the sequential models, where the meaning of a word is discerned by separately going back to each following word in the sentence, self-attention models ‘bake’ the understanding of other words in a sentence into the particular word we are processing. That is, every word has its relationship with other words in the input encoded *at the time* of the processing itself. This helps in parallel optimisation and vastly improves the performance of the model.⁷

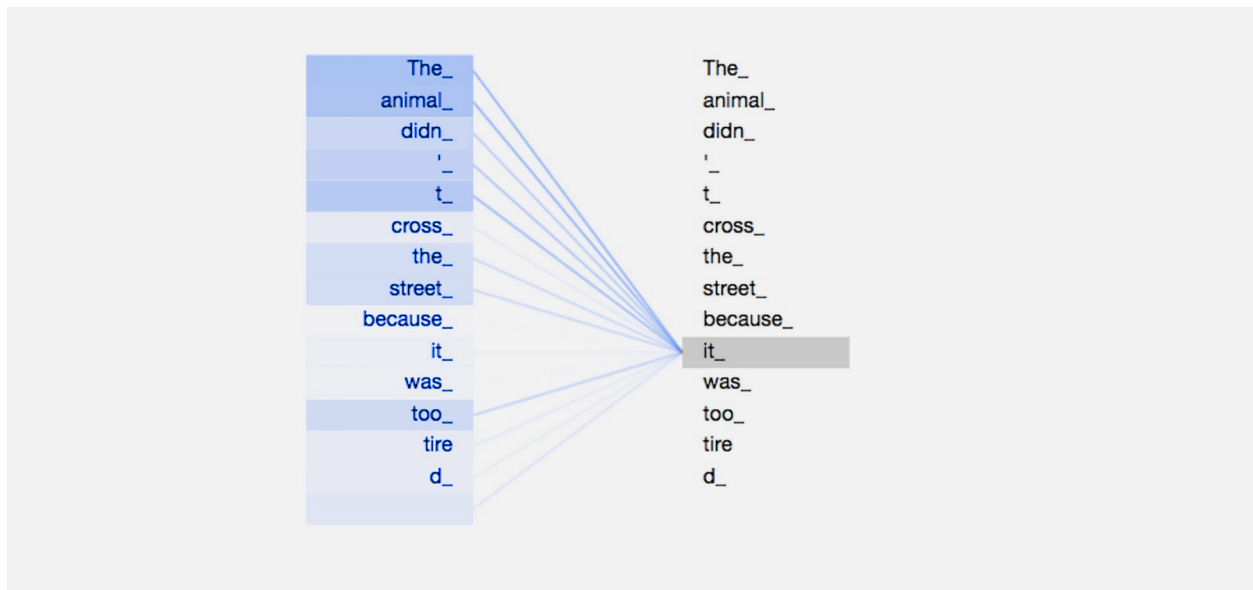
Say, we were to translate the following sentence:⁸ “The animal didn’t cross the street because it was too tired.” Who/what is the word ‘it’ referring to: the animal or the street? At the time of processing, self-attention enables the model to associate “it” with “animal”. The following is demonstrated in the diagram below:

6 See: <https://daleonai.com/transformers-explained>, last visited on March 29, 2023.

7 See: <http://jalammargithub.io/illustrated-transformer/>, last visited on March 29, 2023.

8 See: <http://jalammargithub.io/illustrated-transformer/>, last visited on March 29, 2023.

2. Content Generation through Generative Artificial Intelligence



Source: Jay Alammar, *The Illustrated Transformer*⁹

Just like for the attention mechanism, a meaningful internal representation of data is learned by the model using a vast trove of unlabelled data that it is trained on without any supervision. Having trained the transformer model, it can then be fine-tuned through supervised learning and be ‘transferred’ to a wide area of tasks. This is why it is called a pre-trained transformer. And since we are talking about a transformer model used to generate content, it is called a ‘Generative Pre-Trained Transformer’ (the GPT in ChatGPT).

This technique of training is called semi-supervised learning. Unlabelled data, when used in conjunction with a small amount of labelled data, can produce considerable improvement in learning accuracy. Besides, the costs associated with labelling an enormous corpus of text would render the entire experiment infeasible.

The impressive feature of neural networks is that, even without any supervision, they can learn rules of grammar, parts of speech, and whether words are synonymous.¹⁰

B. Images

Dall-E is one of the platforms that combines elements of natural language and image processing, to generate novel images based on textual prompts. The ‘text-to-image’ generator’s brilliance lies in its capability to combine distinct and unrelated concepts in semantically plausible ways. Dall-E has shown a remarkable ‘understanding’ of visual concepts (e.g. macroscopic, cross-sectional, or overexposed pictures), places (photo of a teddy bear in Times Square), and time (photo of computers in the 1950s). It can also reproduce images in a particular style (sketch or cartoon) or emulate the style of a particular artist.¹¹

⁹ See: <http://jalammar.github.io/illustrated-transformer/>, last visited on March 29, 2023.

¹⁰ See: <https://daleonai.com/transformers-explained>, last visited on March 29, 2023.

¹¹ See: <https://daleonai.com/dalle-5-mins>, last visited on March 29, 2023.

2. Content Generation through Generative Artificial Intelligence

Like many other GAI models, Dall-E, at its core, uses the Transformer architecture. Dall-E is trained to first recognize the relationship between natural language and image concepts. These relationships are ‘encoded’ in a mathematical language understandable by machines. The encodings represent the similarity/dissimilarity between image/caption pairs.¹²

The next step involves reversing the encodings to produce images on text prompts. It is important to mention that Dall-E does not simply reproduce the original encoded image but reproduces the *salient features* of that image. Thus, it can produce novel images from the same text prompt.¹³

C. Sounds and Audios

In November 2022, a human and an AI platform were seen performing a musical piece in collaboration, and GAI made it possible for the AI platform to improvise on the pitch, rhythm and timbre like a human being would.¹⁴

GAI can be used to produce sounds and audios through the process of audio synthesis. Audio synthesis involves training a neural network to generate audio data based on patterns in existing audio samples; the network can then be used to create new audio that is similar in style and content to the original data.¹⁵

Another way of generating sound through GAI is producing audio directly from mathematical equations that model sound waves.¹⁶ These processes have the potential to produce highly realistic audio that sounds like it was created by a human.

Google had brought forward a GAI based framework for audio generation, called AudioLM. AudioLM was designed to generate realistic speech and piano music by processing audio inputs. The aim behind AudioLM was to move a step ahead from the text to audio models. Google relied on semantic and acoustic tokens accordingly to capture the details of the audio inputs which could result in high quality synthesis and output.¹⁷

D. Videos

GAI is also very widely used for the generation of videos through a process called video synthesis. Similar to audio synthesis, this process relies upon the training of a neural network to generate results based on the patterns in video samples.

Generative Adversarial Network (“GAN”) is the most popular approach to audio and video synthesis. GANs have made it possible to create highly realistic videos as they operate on two neural networks, the first network creates new video samples based on patterns in the original data and the other attempts to distinguish between the original video samples and those created by the software. Such models train by themselves in order to ensure that the results produced are as close to reality as possible.

12 See: <https://www.assemblyai.com/blog/how-dall-e-2-actually-works/>, last visited on March 29, 2023.

13 See: <https://www.assemblyai.com/blog/how-dall-e-2-actually-works/>, last visited on March 29, 2023.

14 See: <https://www.wired.co.uk/article/generative-ai-music>, last visited on April 3, 2023.

15 See: <https://arxiv.org/pdf/2006.06426.pdf>, last visited on April 3, 2023.

16 See: <https://arxiv.org/pdf/2006.06426.pdf>, last visited on April 3, 2023.

17 See: <https://ai.googleblog.com/2022/10/audiolm-language-modeling-approach-to.html>, last visited on April 3, 2023.

2. Content Generation through Generative Artificial Intelligence

In September 2022, Meta introduced their text-to-video platform called Make-A-Video which offers the feature to turn text prompts into brief, high-quality video clips.¹⁸ This platform relies on images with descriptions to learn how the output should be presented coupled with unlabelled videos which train it to understand how the movements work. This allows the user to generate videos with a few phrases of text as an input prompt.¹⁹

E. 3D Models

In addition to Dall-E and ChatGPT, OpenAI has also introduced a platform that can be used to create 3D models based on text prompts. The platform works in two steps; it produces a text-to-image result which is then used to produce an image-to-3D model. Essentially, Point-E samples a picture using the text-to-image model and then samples a 3D object conditioned on the sampled image to create a 3D object from a text prompt.²⁰

To accelerate the development of 3D worlds in the metaverse, NVIDIA has also launched various GAI based projects which produce 3D models based on the body movements, facial scans, video recordings, and brush strokes used as input prompts on these platforms.²¹ The technology which allows the generation of 3D models is run on GAN too, which is used for audio and video synthesis as explained above. The production of 3D models is a more complex step after the images or videos are generated, and hence, the underlying technology remains the same.

18 See: <https://ai.facebook.com/blog/generative-ai-text-to-video/>, last visited on April 3, 2023.

19 See: <https://arxiv.org/pdf/2209.14792.pdf>, last visited on April 3, 2023.

20 See: <https://arxiv.org/pdf/2212.08751.pdf>, last visited on April 3, 2023.

21 See: <https://developer.nvidia.com/blog/rapidly-generate-3d-assets-for-virtual-worlds-with-generative-ai/>, last visited on April 3, 2023.

The Positives: Groundbreaking Use Cases and Benefits

GAI has several advantages that are transforming various industries. It allows for the creation of new content that is unique and contextually relevant, leading to personalized experiences for users. It can also automate routine tasks, improving efficiency and reducing costs.

We have highlighted below some key use cases of GAI:

A. Research and Education

GAI has opened up exciting opportunities for research across a range of fields. GAI can generate large quantities of data, which can be used to train machine learning models or conduct simulations. In addition, GAI models can be used to create synthetic data that closely mimics real-world data, providing a way to generate training data in situations where real data is scarce or expensive to obtain.¹

Additionally, its ability to process and understand human language makes it useful for analyzing textual data and performing language-based research. With the capacity to analyze large volumes of data and identify patterns and trends that might be difficult for humans to detect, GAI models offer increased efficiency and can save time and resources compared to manual analysis. Multiple researchers can use such GAI platforms simultaneously and share their results with each other, allowing for collaborative research.

In fact, GAI is also transforming education by enabling AI-based personalized learning experiences, intelligent tutoring systems, and data-driven insights into students' performance. It offers tools for everything from creating customized curriculums that adapt to each student's need to producing high-quality content for effective and interactive educational experiences.²

Furthermore, GAI can be used to explore complex systems and simulate scenarios that would be difficult or impossible to observe in the real world. For example, GAI has been used to model the spread of diseases, predict the outcome of chemical reactions, and forecast SARS and Ebola-infected cases.³ These simulations can help researchers understand the underlying mechanisms of complex systems and develop interventions or strategies to improve outcomes.

GAI can also help in drug discovery, improving MRI/CT scans, and customized medicine. It can generate new compounds, predict their properties, optimize existing drugs, and simulate their effects on different patient populations. It can also improve the accuracy and speed of medical imaging, helping doctors to identify abnormalities and diagnose diseases more quickly and accurately. With its ability to analyze large datasets of patient information, it can easily identify patterns and develop personalized treatment plans for patients.⁴

1 See: <https://www.staticai.ai/post/types-synthetic-data-examples-real-life-examples>, last visited on March 29, 2023.

2 See: <https://research.aimultiple.com/generative-ai-in-education/>, last visited on March 29, 2023.

3 See: <https://europepmc.org/article/ppr/ppr314276>, last visited on March 29, 2023.

4 See: <https://research.aimultiple.com/generative-ai-healthcare/>, last visited on March 29, 2023.

3. The Positives: Groundbreaking Use Cases and Benefits

GAI also has the potential to revolutionize material science by accelerating the discovery and design of new materials. It can generate new molecular structures by predicting their properties such as strength, reactivity, conductivity, etc. as well as optimize the synthesis process for new materials, reducing the time and cost required for experimentation.⁵

B. Industrial Application

GAI has a wide range of industrial applications. In the manufacturing sector, GAI can help to optimize production processes by analyzing machine sensor data. It can predict when equipment failure is most likely to occur. Analysis of such production data can also enable manufacturers to identify potential problems in advance, and accordingly plan maintenance and repairs of the machinery which can reduce the overall downtime and improve efficiency.⁶

GAI can also be used to simulate production scenarios and predict the outcome of different manufacturing decisions. For example, the impact of changes in production schedules, equipment upgrades, or other variables can be simulated by GAI, allowing manufacturers to evaluate the impact of their strategies.⁷

It can enable industries such as automotive, aerospace, defense, and manufacturing to design parts that meet specific goals and requirements. For example, GAI can help automakers to design lighter car parts, contributing to their goal of making cars more fuel-efficient.

In the realm of finance, GAI can identify fraudulent transactions by examining transactional data, predict the creditworthiness of individuals or entities by analyzing factors like income and credit history, and mitigate risks by scrutinizing historical data. It can automate mundane tasks, optimize investment portfolios, generate trading strategies, and optimize pricing for financial products.⁸ Hence, fintech companies can also leverage GAI to streamline workflows, enhance productivity by automating repetitive tasks and facilitate informed decision-making.

In the real estate industry, GAI can assist in estimating the value of a property by considering factors such as size, condition, and location. It can also provide personalized property recommendations based on a buyer's preferences and can predict optimal rent prices while considering market trends, competition, and demand. Predicting the maintenance and repair requirements of a property, it can also help property managers in reducing costs and enhance the property's quality.⁹

C. Programming and Design

GAI can be leveraged in software development because of its ability in writing optimized and efficient code. It can analyze existing codebases and extract patterns, helping developers create more efficient code that uses fewer system resources.¹⁰

5 See: <https://www.gartner.com/en/articles/beyond-chatgpt-the-future-of-generative-ai-for-enterprises>, last visited on March 29, 2023.

6 See: <https://www.leewayhertz.com/generative-ai-use-cases-and-applications/>, last visited on March 29, 2023.

7 Ibid.

8 See: <https://research.aimultiple.com/generative-ai-applications/#4-banking-applications>, last visited on March 29, 2023.

9 See: <https://www.leewayhertz.com/generative-ai-use-cases-and-applications/>, last visited on March 29, 2023.

10 See: <https://www.turintech.ai/generative-ai-for-code-what-you-need-to-know-in-2023/>, last visited on March 29, 2023.

3. The Positives: Groundbreaking Use Cases and Benefits

Additionally, GAI can help automate the coding process, thus increasing productivity and reducing development costs. It can generate code snippets or entire functions that meet the required specifications, which can be a significant time saver.¹¹

Apart from the above, developers at organisations are using GAI models for other purposes such as (1) elucidating the functionality of code they did not create but are currently responsible for managing or upgrading, (2) taking guidance on which libraries or classes of code are necessary to execute specific requirements, (3) enhancing code consistency by automating refactoring and style adjustments, (4) suggestions on simplification of code, and (5) translation of code from one language to another, etc.¹²

GAI can also be used to optimize the design of chips by generating new layouts, automating the design process, and customizing chip designs for specific applications. It can generate optimized layouts for better performance, power consumption, and area utilization, reducing design time and costs significantly.¹³ Reportedly, Google is also leveraging the power of AI to automate its chip design process.¹⁴

D. Customer Support

GAI, also known as natural language generation (“NLG”)¹⁵, has become increasingly popular in the field of customer service due to the numerous benefits it offers. According to a report by Gartner¹⁶, by 2027, customer service interactions will be primarily handled by AI technologies such as chatbots, virtual customer assistants (“VCAs”), or other conversational AI platforms using NLG.

One key advantage of NLG in customer service is the ability to provide personalized interactions. By using data analytics and machine learning algorithms, NLG can analyze customer data and provide tailored responses to individual queries, improving customer satisfaction and loyalty. NLG’s 24/7 availability enables it to assist customers with their queries at any time, even outside of business hours. By being available around the clock with the ability to handle multiple queries simultaneously, NLG can help reduce response times and improve customer satisfaction.¹⁷

Moreover, by automating routine tasks, NLG can reduce the workload on customer service teams, allowing them to focus on more complex queries. This can result in cost savings and more efficient use of resources. NLG can also handle a large volume of queries without additional human resources, allowing for increased scalability without compromising on quality.

As an example, Y Meadows, a company that provides automated customer service, has integrated ChatGPT with its platform. With this integration, customer service representatives can quickly respond to customer emails and support cases using ChatGPT-generated content.

¹¹ Ibid.

¹² See: <https://advisory-marketing.us.kpmg.com/speed/pov-generativeai.html>, last visited on March 29, 2023.

¹³ See: <https://www.gartner.com/en/articles/beyond-chatgpt-the-future-of-generative-ai-for-enterprises>, last visited on March 29, 2023.

¹⁴ See: <https://www.theverge.com/2021/6/10/22527476/google-machine-learning-chip-design-tpu-floorplanning>, last visited on March 29, 2023.

¹⁵ NLG involves creating text that appears to have been written by a human, without the involvement of a human writer. NLG is a part of NLP, which includes various technologies and techniques that deal with human language. NLG technology is complementary to Natural Language Understanding (NLU) and they work together to achieve NLP. See: [https://www.aidataanalytics.network/data-science-ai/articles/what-is-natural-language-generation#:~:text=Natural%20language%20generation%20\(NLG\)%20is,\(NLU\)%20to%20generate%20NLP](https://www.aidataanalytics.network/data-science-ai/articles/what-is-natural-language-generation#:~:text=Natural%20language%20generation%20(NLG)%20is,(NLU)%20to%20generate%20NLP), last visited on March 29, 2023.

¹⁶ See: <https://www.gartner.com/en/newsroom/press-releases/2022-07-27-gartner-predicts-chatbots-will-become-a-primary-customer-service-channel-within-five-years>, last visited on March 29, 2023.

¹⁷ See: <https://medium.com/dataprophet/ai-taking-the-wheel-how-generative-ai-enhances-customer-support-7a948f81f5b5>, last visited on March 29, 2023.

3. The Positives: Groundbreaking Use Cases and Benefits

The Y Meadows platform matches the customer's question with the relevant knowledge base and generates a response that is integrated with the appropriate information from the company's internal support system. If necessary, a human can review and approve the response before it is sent to the customer.¹⁸ Many other companies like IBM Watson, Salesforce, and Microsoft Azure have also implemented GAI customer support.¹⁹

E. Entertainment

GAI is increasingly being used in the entertainment industry and provides new opportunities for content creation and personalization. GAI can create new and unique content that can be tailored to the preferences of individual users. For example, AI-generated music can be tailored to match the mood, tempo, and genre that a user prefers, leading to a more personalized listening experience.²⁰

Additionally, GAI can be used to create visual art, such as paintings and drawings, which can be customized to a user's preferences.²¹ GAI can also be used to create virtual characters, such as avatars to interact with users in video games or other applications, which can enhance the overall user experience.

GAI is especially helpful in the entertainment industry because of its ability in enhancing image processing and film restoration. By using GAI algorithms, photo, and video images can be refined, which improves their resolution by reducing noise and sharpening the image, as well as enhancing the color quality.²²

For example, NVIDIA is utilizing GAI to create highly realistic virtual characters and 3D assets that can be used in virtual worlds and metaverses. NVIDIA has launched numerous AI research projects such as Audio2Face and Audio2Gesture to help creators across industries unlock new possibilities with GAI.²³ Similarly, Facebook (now Meta) has developed a tool called SapFix that uses AI to automatically generate fixes for bugs found in its codebase.²⁴

F. Healthcare and Medicine

GAI has the potential to revolutionize the medical industry by aiding in drug discovery, improving MRI/CT scans, and providing customized medicine for patients. One of the most significant contributions of GAI in the medical field is the creation of synthetic patient data in the form of images, text, etc. This allows healthcare practitioners and researchers to generate vast amounts of realistic data that closely mimics real patient data, without compromising patient privacy.²⁵

18 See: <https://ymeadows.com/en-specsheets/using-openai-with-y-meadows>, last visited on March 29, 2023.

19 See: <https://medium.com/dataprophet/ai-taking-the-wheel-how-generative-ai-enhances-customer-support-7a948f81f5b5>, last visited on March 29, 2023.

20 See: <https://techstory.in/how-will-generative-ai-impact-the-world-of-entertainment/#:~:text=Generative%20AI%20is%20also%20being,for%20movie%20and%20TV%20production>, last visited on March 29, 2023.

21 See: <https://www.superannotate.com/blog/generative-ai-explained>, last visited on March 29, 2023.

22 See: <https://www.e2enetworks.com/blog/top-5-advantages-of-generative-ai-applications>, last visited on March 29, 2023.

23 See: <https://developer.nvidia.com/blog/rapidly-generate-3d-assets-for-virtual-worlds-with-generative-ai/>, last visited on March 29, 2023.

24 See: <https://engineering.fb.com/2018/09/13/developer-tools/finding-and-fixing-software-bugs-automatically-with-sapfix-and-sapienz/>, last visited on March 29, 2023.

25 See: <https://research.aimultiple.com/generative-ai-healthcare/#how-does-generative-ai-work-and-how-can-it-be-used-in-the-healthcare-industry>, last visited on April 07, 2023.

3. The Positives: Groundbreaking Use Cases and Benefits

GAI can also help predict the 3D structure of proteins, which is essential for drug discovery and understanding diseases at a molecular level. For example, the Google-owned DeepMind's²⁶ AlphaFold²⁷ accurately predicts the shape of a protein, facilitating identification and development new drugs to treat diseases, without the need for time consuming and expensive experiments.

DeepMind, in partnership with London's Moorfields Eye Hospital, has also developed a machine learning based system that generates 3D models of the human eye's retina. This can accurately predict eye diseases by allowing doctors to visualize and detect anomalies that might not be visible in 2D scans.²⁸ As the technology advances, we can expect to see even more innovative applications that could greatly improve patient outcomes and advance medical research.

26 See: <https://www.deepmind.com/about>, last visited on April 07, 2023.

27 See: <https://www.deepmind.com/research/highlighted-research/alphafold>, last visited on April 07, 2023.

28 See: <https://www.deepmind.com/blog/using-ai-to-predict-retinal-disease-progression> and <https://www.theverge.com/2018/8/13/17670156/deepmind-ai-eye-disease-doctor-moorfields>, last visit on April 07, 2023.

Concerns and Disruption of Regulations

A. The Veracity of Information and Fake News

GAI is a rapidly advancing field of AI that involves training machines to generate novel content such as text, images, and videos. While GAI has many useful applications, it also raises concerns about its potential to be used maliciously.

GAI models may be susceptible to bias, as they learn patterns and make predictions based on the data they are trained on. If the training data is biased or incomplete, the model's output can also be biased. To mitigate such bias, it is imperative to use diverse training data developing algorithms that are more transparent and explainable (refer to Section IV (4) below).¹

Given that AI language models such as ChatGPT can generate human-like text and can be trained to mimic the writing style of individuals, there are serious concerns about its potential misuse for spreading fake news on a large scale. Such misinformation can have serious ramifications, such as distorting public perceptions and eroding confidence in societal structures and systems.²

Moreover, GAI models can also be used to create fake social media accounts or bots to spread false information.³ These accounts could flood social media platforms with fabricated news stories, making it challenging for individuals to find credible sources of information, therefore, making it harder for people to engage in informed and constructive discussions.⁴ Moreover, even in the absence of user intent, GAI may provide inaccurate outputs while sounding convincing.

However, there have also been studies in which researchers have examined the use of GAI models using NLP for detecting fake news.⁵ NLP involves analyzing human language, including its structure and meaning, to gain insights into the intent and sentiment of the text. By using NLP techniques, AI algorithms can analyze news articles, social media posts, and other online content to identify patterns and characteristics associated with fake news. It can analyze the language used to identify potential biases or propaganda techniques and also to fact-check the claims made.⁶ Therefore, while GAI has its benefits and drawbacks, its overall impact will only be determined over time.

B. Generative Artificial Intelligence and Intellectual Property Rights

While the question regarding ownership and other IP rights in content created by AI has been raised for decades, consumer-facing GAI, such as ChatGPT, Dall-E, Bard, etc., have injected this debate with new vigor.

1 See: https://www.linkedin.com/pulse/why-chatgpt-other-generative-ai-technologies-part-all-de-breuck/?trk=public_post, last visited on March 29, 2023.

2 See: <https://www.analyticsinsight.net/chatgpt-can-be-used-to-disseminate-fake-news-on-a-large-scale/>, last visited on March 29, 2023.

3 See: <https://economictimes.indiatimes.com/opinion/et-commentary/chatgpt-faking-it-a-genuine-artificial-concern/articleshow/97233545.cms?from=mdr>, last visited on March 29, 2023.

4 See: <https://www.analyticsinsight.net/chatgpt-can-be-used-to-disseminate-fake-news-on-a-large-scale/>, last visited on March 29, 2023.

5 See: <https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=9620068>, last visited on March 29, 2023.

6 Ibid.

4. Concerns and Disruption of Regulations

Whether the output of GAI amounts to Intellectual Property (“IP”) and the extent of protection that would be granted to such IP are not just questions of law, but of politics and ideology as well. Nevertheless, in this section, we will focus on legal aspects of IP, while also touching upon certain theories underlying the law of IP.

Specifically, given that the most popular consumer-facing GAI (at the time of writing) are capable of creating text (ChatGPT and Google Bard), images (Dall-E), music (Jukebox, Amper AI), voice (Whisper), software programmes and videos, it is pertinent to focus on IP issues pertaining to copyright laws.

Under Indian Copyright law, copyright only subsists in certain works, namely, (i) original literary,⁷ dramatic,⁸ musical,⁹ and artistic works,¹⁰ (ii) cinematograph films;¹¹ and (iii) sound recordings.^{12,13} The output of the various GAI described above will qualify as a “work” in which copyright can subsist.

However, just by virtue of being such a work, copyright does not subsist in such work automatically. For a ‘work’ to qualify for copyright protection, it would first have to meet the ‘modicum of creativity’ standard laid down in *Eastern Book Company v. D.B. Modak*.¹⁴ In this case, the Court held, in relation to the existence of copyright in the compilation of judgements, a ‘minimal degree of creativity’ was required, and that ‘there must be some substantive variation and not merely a trivial variation’. However, a plain reading of this test does not provide a definitive conclusion on whether an AI can meet the ‘modicum of creativity’ as required.

In addition to the above, the second requirement to be satisfied by an AI when it comes to the ownership of copyrighted works is the requirement to fall under the aegis of an ‘author’ as defined under the Copyright Act.

Under Section 2 (d) of the Copyright Act, 1957, “author” means,—“(vi) in relation to any literary, dramatic, musical or artistic work which is computer-generated, the person who causes the work to be created;” The first issue under the above-mentioned definition is its usage of the terms ‘the person who causes the work to be created’. Determining who ‘causes’ a work to be created is a question of the proximity of a natural or legal person to the creation of the ‘expression’ in the content in question — the more closely or directly a person is involved in creating the ‘expression’, the more he or she contributes to it, and the more likely he or she is to qualify as a person ‘who causes the work to be created’. As a result of the above, the current legal framework under the Copyright Act, 1957 may not effectively deal with/prescribe for creation of works where the actual creator or a contributor of the ‘expression’ is not a human or a legal person, or where the role of the human or legal person is minimal. Since there is no settled jurisprudence to this effect, it is unclear as to the ‘author’ of AI-created works under Indian copyright law.

7 Under Section 2(o) of the Copyright Act, 1957 (“Copyright Act”), “Literary work” includes computer programmes, tables and compilations including computer databases.

8 Under Section 2(h) of the Copyright Act, “dramatic work” includes any piece for recitation, choreographic work or entertainment in dumb show, the scenic arrangement or acting, form of which is fixed in writing or otherwise but does not include a cinematograph film.

9 Under Section 2(p) of the Copyright Act, “musical work” means a work consisting of music and includes any graphical notation of such work but does not include any words or any action intended to be sung, spoken or performed with the music.

10 Under Section 2(c) of the Copyright Act, “artistic work” means (i) a painting, a sculpture, a drawing (including a diagram, map, chart or plan), an engraving or a photograph, whether or not any such work possesses artistic quality; (ii) a work of architecture; and (iii) any other work of artistic craftsmanship.

11 Under Section 2(f) of the Copyright Act, “cinematograph film” means any work of visual recording and includes a sound recording accompanying such visual recording and “cinematograph” is construed as including any work produced by any process analogous to cinematography including video films.

12 Under Section 2(xx) of the Copyright Act, “sound recording” means a recording of sounds from which such sounds may be produced regardless of the medium on which such recording is made or the method by which the sounds are produced.

13 Section 13(1) of the Copyright Act.

14 (2008) 1 SCC 1.

4. Concerns and Disruption of Regulations

Interestingly, the Indian Copyright Office granted copyright recognizing “RAGHAV Artificial Intelligence Painting App” (“**AI Raghav**”) and Ankit Sahni as authors for a painting titled “Suryast” in December 2021.¹⁵ However, as per news reports, around a year after registration, a withdrawal letter by the Copyright Office was issued.¹⁶ The office placed the onus on the applicant to “inform the Copyright Office on the legal status of AI Raghav in the withdrawal notice. Nevertheless, the office’s website reflects the status of the application as still registered. The actual status of the case is unknown.

In 2022, the US Copyright Office had issued a decision rejecting copyright protection for images generated by Midjourney, which is a GAI system. However, in March 2023, the office released a guidance on “Works Containing Material Generated by AI”¹⁷ in which it stated that copyright protection depends on whether the AI’s contributions are “the result of mechanical reproduction”, such as in response to text prompts, or if they reflect the author’s “own mental conception”.¹⁸ The office further said that “*The answer will depend on the circumstances, particularly how the AI tool operates and how it was used to create the final work*”.¹⁹

GAI is already attracting significant amounts of investments and is expected to find an increasing number of use cases, as we have detailed earlier. The lack of clarity in the law regarding the copyrightability of AI output will lead to considerable uncertainties, especially for organizations and individuals invested in this sector. Hence, a clear legal position will sooner or later become indispensable, whether it is made through legislation or jurisprudence.

The following questions are likely to need clarifications in this context (i) does AI-generated output qualify as “IP”? (ii) If yes, who owns the IP in such output? and (iii) If not, what is the legal status of the works which have been generated?

The answers to these questions crucially depend on whether the law *should* grant copyright protection to AI-generated works. This requires an analysis of the justification behind granting copyright protection to any work since the justifications behind copyright in traditional works would need to be understood to determine if copyright should exist in AI-generated works.

The justification for granting copyright under various jurisdictions may differ, for example, in common law and civil law jurisdictions.²⁰ The utilitarian rationale for copyright has historically dominated in common law jurisdictions, as exemplified in the U.S. Constitution’s provision “*to promote the progress of science and useful arts*”.²¹ In the French IP Code, the natural right justification is more evident, as is common in other civil law countries — “*The author of a work of the mind shall enjoy in that work, by the mere fact of its creation, an exclusive incorporeal property right which shall be enforceable against all persons.*”²²

Apart from these justifications, there is the economic justification for copyright, i.e., to protect the revenues and livelihood of authors, and to encourage further creation.²³ Moreover, copyright is considered to foster

¹⁵ Diary no. 13646/2020-CO/A.

¹⁶ See <https://spicyip.com/2022/10/ai-art-and-indian-copyright-registration.html>, last visited on March 29, 2023.

¹⁷ See: <https://fingfx.thomsonreuters.com/gfx/legaldocs/byvrlmjzve/AI%20COPYRIGHT%20notice.pdf>, last visited on March 29, 2023.

¹⁸ See: <https://www.reuters.com/world/us/us-copyright-office-says-some-ai-assisted-works-may-be-copyrighted-2023-03-15/>, last visited on March 29, 2023.

¹⁹ See: <https://www.reuters.com/world/us/us-copyright-office-says-some-ai-assisted-works-may-be-copyrighted-2023-03-15/>, last visited on March 29, 2023.

²⁰ Paul Goldstein and P. Bernt Hugenholtz, *International Copyright: Principles, Law, And Practice* (Oxford University Press, 2019), 5.

²¹ Article I, Section 8, clause 8 of the US Constitution.

²² Art. L. 111- 1, Intellectual Property Code of France.

²³ See: <https://cyber.harvard.edu/people/tfisher/iptheory.pdf>, last visited on March 29, 2023.

4. Concerns and Disruption of Regulations

democracy and free speech by ensuring that authors do not have to depend on Government or private patronage, and can earn revenues by lawfully exploiting their works, and restricting others from doing so.²⁴

Applying these justifications to AI-generated works also requires a determination of whether the AI in question is pure AI or human-assisted AI. In ChatGPT's own words, ChatGPT does require “*some level of human intervention in its training and maintenance...such as periodic retraining with new data or adjustments to its algorithms or parameters*”.²⁵ Hence, assuming that the GAI in question is assisted by humans, and in light of the justifications behind copyright protection, should the works created by such AI based on human instructions be granted copyright protection?

There are arguments on both sides — proponents of copyright protection state that such kind of protection is necessary to incentivize AI programmers, investors, and users (such as licensees who may want to use GAI for their own business).²⁶ The UK, for example, extends copyright protection to works that are generated only by computers, however, no moral rights are granted.²⁷ Section 9 of the Copyright, Designs, and Patents Act, 1988 states that the author of computer-generated works is “the person by whom the arrangements necessary for the creation of the work are undertaken.” Although this law has been in existence since the 1980s, this is in line with the UK's objective towards encouraging AI development in the UK, and incentivizing businesses to research, develop and deploy AI technology.²⁸

The Indian Copyright Act, recognizes that computer-generated works have authors, i.e., the person who “causes” the work to be created,²⁹ however, it does not define computer-generated works or the extent of human involvement in such works, if any.

On the other hand, there are arguments to state that if AI platforms own the IP in the GAI, platforms such as ChatGPT will own an infinite number of copyrighted works that users direct such AIs to create. Hence, if 1 million users ask Dall-E to create or modify 1 work in a day, OpenAI may become the owner of 1 million works each day. Not only this, but owners of GAI may deliberately require the AI to create close variations of the same work (e.g., 100 variations of a picture of a living room), which will lock out creators from owning copyright in any of those works.

Other arguments against protecting AI-generated works include (i) the fact that copyright has its roots in human authorship and creative endeavor, and protecting AI-generated works will devalue human creativity;³⁰ and (ii) there is no evidence that such protection will actually incentivize the development of GAI.³¹

One may also argue that although the works are generated by artificial “intelligence”, they are not “intellectual” property per se. There are also suggestions that AI-generated works should form part of the public domain since there is no identifiable author. However, the ability to differentiate between human-generated works and those generated by GAI would only become challenging.

²⁴ Supra note 82, at 6.

²⁵ OpenAI's ChatGPT AI language model, response to question from author “Can ChatGPT be categorised as Assisted AI?”, February 16, 2023.

²⁶ Mauritz Kop, AI & Intellectual Property: Towards an Articulated Public Domain, 28 Texas Intellectual Property Law Journal (2020).

²⁷ See: <https://www.gov.uk/government/consultations/artificial-intelligence-and-intellectual-property-call-for-views/artificial-intelligence-call-for-views-copyright-and-related-rights>; and <https://www.pinsentmasons.com/out-law/analysis/uk-to-decide-copyright-protection-creative-works-generated-ai#:~:text=Currently%2C%20creative%20works%20generated%20solely,on%20which%20it%20was%20made>, last visited on March 29, 2023.

²⁸ See: <https://www.pinsentmasons.com/out-law/analysis/uk-to-decide-copyright-protection-creative-works-generated-ai#:~:text=Currently%2C%20creative%20works%20generated%20solely,on%20which%20it%20was%20made>, last visited on March 29, 2023.

²⁹ Section 2(d)(vi) of the Copyright Act.

³⁰ See: <https://www.pinsentmasons.com/out-law/analysis/uk-to-decide-copyright-protection-creative-works-generated-ai#:~:text=Currently%2C%20creative%20works%20generated%20solely,on%20which%20it%20was%20made>, last visited on March 29, 2023.

³¹ Supra note 88.

4. Concerns and Disruption of Regulations

This may also make it difficult to enforce copyright laws and especially with the progression in technology as it would be difficult to prove whether GAI was used in the creation of a human work. On the other hand, there are also suggestions of a shorter term protection for computer-generated works, which would allow third parties to benefit from free use of the work once the protection had expired earlier than the full-fledged copyright term.³²

Whichever approach is taken, it is important to assess the impact such an approach will have on incentives to humans to create works and develop such GAI. There is currently insufficient data to determine which approach will work best, both from a copyright user and a GAI developer perspective.³³ Nevertheless, these debates must continue so that Governments can take timely and informed actions when the popularity of GAI demands a decision to be made.

Does GAI's Training Process and Use of Training Data Amount to Infringement?

While we have discussed the IP aspects in the output produced by GAI, it is worthwhile looking at the input as well. As discussed earlier, GAI is trained based on existing data. Existing data may be in the form of text, images or sound recordings, all of which could potentially be copyright-protected works. The training dataset for Stable Diffusion, for example — one of the biggest and most influential text-to-AI systems — contains billions of images scraped from hundreds of domains; everything from personal blogs hosted on WordPress and BlogSpot to art platforms like DeviantArt and stock imagery sites like Shutterstock and Getty Images.³⁴ GAI likely produce copies of such existing works by encoding them. Then, does such usage result in infringement of the copyright of the existing works? Numerous lawsuits have already been filed claiming copyright infringement.³⁵

ChatGPT's own response to the question “Does training of ChatGPT involve infringement of existing copyrighted works?” is as follows:

“No, the training of ChatGPT does not involve infringement of existing copyrighted works. The training data for ChatGPT comes from publicly available sources, such as books, articles, and websites, that are generally considered to be in the public domain. These sources are used under fair use or other legal exceptions that allow for the use of copyrighted materials for certain purposes, such as education and research.

Furthermore, the training process for ChatGPT involves the use of unsupervised learning techniques, which do not require the use of specific examples or data that would infringe on copyrighted works. Instead, ChatGPT learns from patterns and correlations in the language it processes.

It should be noted, however, that the use of ChatGPT-generated text may potentially infringe on existing copyrighted works if the generated text contains substantial elements of protected works. In such cases, the responsibility falls on the user of the generated text to ensure that their use falls within the scope of fair use or other legal exceptions.”

32 See: <https://www.gov.uk/government/consultations/artificial-intelligence-and-ip-copyright-and-patents/artificial-intelligence-and-intellectual-property-copyright-and-patents#responses-and-next-steps>, last visited on March 29, 2023.

33 See: <https://www.pinsentmasons.com/out-law/analysis/uk-to-decide-copyright-protection-creative-works-generated-ai#:~:text=Currently%2C%20creative%20works%20generated%20solely,on%20which%20it%20was%20made>, last visited on March 29, 2023.

34 See <https://www.theverge.com/23444685/generative-ai-copyright-infringement-legal-fair-use-training-data>, last visited on March 29, 2023.

35 See <https://techcrunch.com/2023/01/27/the-current-legal-cases-against-generative-ai-are-just-the-beginning/>, last visited on March 29, 2023.

4. Concerns and Disruption of Regulations

Section 52(1)(a) of the Copyright Act provides that fair dealing with any work for the purpose of private or personal use, including research, does not constitute an infringement of copyright. In addition to Section 52 of the Copyright Act, Courts in India have recognized the principle of “fair use” which is also an instance of the use of a copyrighted work that may not amount to infringement. This has been in line with the fair use exception employed under US laws. Indian courts have also used the following four-factor test for determining whether or not a particular use of a work is protected under the fair use exception:

- purpose and character of the use, including whether such use is of a commercial nature or is for non-profit purposes;
- the nature of the copyrighted work;
- the amount and substantiality of the portion used in relation to the copyrighted work as a whole; and
- the effect of the use upon (i) the potential market for, or (ii) the value of, the copyrighted work. The subsequent work should not act as a market substitute or affect the market share of the prior work.

Hence, whether the use by GAI of existing copyrighted work amounts to infringement would be a case-wise analysis. E.g., the use of copyrighted works for training a GAI by a research institution for private use may qualify as fair dealing. On the other hand, if GAI is used to create paintings in the style of a particular artist for commercial purposes, it may not fall within fair dealing or fair use.

With ChatGPT-4 having been launched on a subscription basis and for business users, the argument of fair use is bound to become weaker, given that the purpose of using existing works becomes commercial. On the other hand, AI entities may argue that the training on existing data is akin to a budding artist taking inspiration from existing works. Moreover, the output created by GAI could qualify as a transformative work, which differs sufficiently from the existing works that it has been trained on.

Hence, training AI on existing works and whether the same leads to infringement continues to be contentious. The tug of war between existing right holders and owners of GAI could lead to fine-tuning of fair use principles along with jurisprudence on the ownership of machine-derived works. The factual analysis of each case would play a crucial role, and it is currently difficult to establish a singular principle that will provide a solution to this nuanced area of law.

C. Data Protection and Confidentiality

It is hard to discern the extent to which GAI is (in)compatible with the extant privacy regime, especially the European Union’s General Data Protection Regulation (“**GDPR**”) and proposed Indian Digital Personal Data Protection Bill, 2022 (“**DPDPB**”). We touch upon some of the key issues below.

Training Data and Processing of Personal Information

GAI relies on a vast trove of data to train itself- data that is scraped from a wide variety of sources, such as from paid/open-source third-party data sets or consumer apps. Often, the data contains personal information

4. Concerns and Disruption of Regulations

gathered in contravention of the applicable data protection laws, like the GDPR and the proposed DPDPB. This can carry serious financial, reputational, and penal risks for the developer.³⁶

For example, the DPDPB states that the personal data of a data owner (“**data principal**”) can only be processed for lawful purposes for which a data principal has given consent in the manner proposed by the DPDPB. Thus, the developer has to first identify if the scraped data contains any personal data; and then should they choose to keep the same, they must provide an itemised notice in “clear and plain language containing a description of personal data” sought to be collected.

Given that the scraped data runs into tens of terabytes, identifying and categorising the nature of data, not to mention determining whose personal data it is and contacting them, is no mean feat. The costs and inefficiencies involved in the process will hamper the development of GAI models, but pale in comparison to the risks of non-compliance. For example, the Federal Trade Commission in the US uses ‘algorithmic disgorgement’ as a new privacy enforcement tool, requiring companies who scraped data without people’s consent, in violation of terms of use of certain websites or copyright, to delete their algorithms and training data.³⁷

Personal Information can also become part of the training data *after the* deployment of the GAI model. Privacy policies of platforms such as Open AI *inter alia* mentions that “[I]f you communicate with us, we may collect your name, contact information, and the contents of any messages you send.”

Thus, whenever one enters a prompt on ChatGPT, which could include some personal information about themselves or someone else, then that personal information can permanently become part of the platform’s training data. Given that the inputs are used to train GAI platforms, one can also keep entering leading prompts while revealing personal or false details about others, so that not only is the platform tricked to write false information but also ingests those false prompts in its training data. This can have serious ramifications for one’s privacy as unbeknownst, and without their consent, some other user, not necessarily the developer, is feeding false or personal information into the GAI. It then becomes the onus of the developer to correct or delete the same, if they receive a request from the principal (discussed below).

Right to Correction/Erasure

Even if a developer acquires training data in a fully compliant manner, the data principal still has the right to withdraw their consent. In such a case, the developer has to cease using that personal data in a ‘reasonable time.’³⁸

However, since the concerned personal information has already been fed to a GAI model, and given that developers themselves don’t know what information a model takes into account during its development, it is still an unknown as to how the information can be isolated and removed from a fine-tuned model’s ‘parameters.’

³⁶ See: <https://www.zwillgen.com/privacy/artificial-intelligence-risks-privacy-generative-ai/>, last visited on March 29, 2023.

³⁷ See: https://wp.nyu.edu/compliance_enforcement/2022/03/30/model-destruction-the-ftcs-powerful-new-ai-and-privacy-enforcement-tool/, last visited on March 29, 2023.

³⁸ Section 7, DPDPB.

4. Concerns and Disruption of Regulations

Moreover, the DPBPD mandates that a developer also has to comply with a data principal's requests for correction or erasure of their personal data if the data is inaccurate or misleading, or incomplete, and requires updation.³⁹

Thus, suppose that defamatory content about someone is published online and becomes part of the GAI's training data. Thereafter, a judicial authority determines that the said content is indeed defamatory and orders the concerned websites to take it down. Now, even if the websites comply, that information has already become part of the training data of the GAI model and the same might be outputted to whoever has given related prompts. As highlighted before, the developer would then have to delete that information from the training data and ensure that the GAI doesn't output the same on user prompts. If the developer cannot figure it out, it runs the risk of destroying the model and training data altogether and starting from scratch.

Right to Contextual Integrity

Even if the GAI model solely uses publicly available information as its training data, it is still theoretically capable of violating our 'contextual integrity.' That is, we have a reasonable expectation that our personal information, even if willingly and knowingly made available on the public domain, is not used outside of the context in which agreed to part with it.

For example, the law may require the government to post updated records of liens on one's personal property. Another law may require voter records to be posted in the interest of transparency. We may also post online of our own volition, say blog posts, product reviews, etc. We may thus have fragments of our personal information and views scattered across the web and we would be okay with that. However, GAI is capable of synthesising all this information and making it easily and readily available by someone entering a few queries. It's the scale and speed with which GAIs can compile scattered and fragmented personal information that can violate our contextual integrity.⁴⁰

D. Bias, Prejudice and Ethics

Given the potential of GAI models to alter human perceptions, disseminate false information, and exercise social control, we need to consider the biases (if any) encoded in them. The tendency of algorithms to mirror human prejudices/biases is referred to as algorithmic or AI bias.

AI bias could occur from using incorrect, flawed, or biased data sets by the personnel who train and/or validate the machine learning algorithms.⁴¹ Additionally, the algorithms may also reflect unintentional cognitive biases⁴² or social biases⁴³ of the very persons programming the algorithms. Biases will influence GAI and upcoming AI-based technology. Although these biases reflect our existing society, the goal of AI is to make society better and provide a neutral output.

³⁹ Section 13, DPBPD.

⁴⁰ See: H. Nissenbaum, Privacy as Contextual Integrity, 79 Washington Law Review (2004).

⁴¹ See: <https://wires.onlinelibrary.wiley.com/doi/full/10.1002/widm.1356>, last visited on April 3, 2023.

⁴² See: <https://www.techtarget.com/searchenterpriseai/definition/cognitive-bias>, last visited on April 3, 2023.

⁴³ See: <https://link.springer.com/article/10.1007/s43681-020-00026-z>, last visited on April 3, 2023.

4. Concerns and Disruption of Regulations

AI Bias manifests itself in a variety of ways, including through unjust discrimination and stereotypes. Following are some key types of AI biases which could affect GAI functioning:

Sampling/Selection Bias⁴⁴

When the training data is not randomly picked and appears to be favoring specific groups. For example, facial-recognition AI won't perform well when recognizing women and persons of different races if it is trained mostly on images of white men.⁴⁵

Overgeneralisation Bias⁴⁶

Drawing of conclusions that are excessively broad because they go beyond what can be logically deduced from the facts given. Finding statistical patterns in a training set to produce a more broadly applicable output is a crucial component of machine learning. However, if it is left unchecked, there is always a chance of coming to incorrect or misleading conclusions based on a small sample size or little data.

For example, the greater arrest rates in minority groups are partially explained by the fact that they are monitored and policed more frequently. Yet, since there are differences in how these groups are evaluated and managed, AI should not draw the conclusion that just because persons from minority groups are arrested more frequently, they are inherently more dangerous.⁴⁷

Representation Bias⁴⁸

Arises when the data used to train a language model contains biased representations of particular groups of people. Additionally, an AI model can also produce inaccurate projections concerning under-represented groups if it is trained on data that only represents a small portion of the population.

For example, a language AI tool that scans curriculum vitae for recruitment or career counselling may be less likely to propose historically discriminated groups to recruiters, or more likely to offer lower-paying employment to marginalised groups as the data fed to the AI represents bias in recruitment of such marginalized groups.⁴⁹

44 See: <https://dl.acm.org/doi/pdf/10.1145/3466132.3466134>, last visited on April 3, 2023.

45 See: <https://link.springer.com/article/10.1007/s13347-021-00478-z>, last visited on April 3, 2023.

46 See: <https://www.statice.ai/post/data-bias-types>, last visited on April 3, 2023.

47 See: <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>, last visited on April 3, 2023.

48 See: <https://arxiv.org/abs/1908.09635>, last visited on April 3, 2023.

49 See: <https://arxiv.org/abs/2301.12867>, last visited on April 3, 2023.

4. Concerns and Disruption of Regulations

Evaluation Bias⁵⁰

Develops while evaluating and iterating an AI model. The quality of the model is frequently evaluated against predetermined standards after it has been optimised using training data. Since these standards are biased or inappropriate for how the model will be utilised, they may not accurately reflect the overall population. For example, AI models deployed in recruitment services, showed preference to male candidates by filtering out female candidates, based on faulty eligibility criteria formulated by the AI, arising from the data fed to it.⁵¹

Aggregation Bias⁵²

Occurs when incorrect inferences are made about specific people based on studying the overall population. For example, consider AI tools used in clinical aid that is looking at diabetic patients who appear to have different morbidities depending on their gender and racial backgrounds. The AI model may make generalisations about gender and racial subgroups in the population for such morbidities.

E. Explainability and Accountability of Artificial Intelligence

AI is often praised for its effectiveness, but concerns about the transparency, accuracy, and trustworthiness of its methods have started to emerge.⁵³

The issue of accountability in GAI is becoming increasingly important and arises from the fact that, unlike human-generated content, AI-generated content may be difficult to trace back to its source or author. This makes it challenging to hold individuals or organizations responsible for any harm caused. Additionally, as discussed in Section 3(a) above, GAI models have the potential to exhibit bias since they learn patterns and make predictions from the data on which they are trained. If such training data is biased or incomplete, the model's output may also be biased. Explainable AI or XAI is one such tool to address these issues.

To mitigate some risks, the development of Explainable AI (“XAI”)⁵⁴ has evolved to address accountability issues.⁵⁵ The concept of XAI pertains to the ability to explain how an AI system arrived at a particular outcome, decision, or suggestion in order to help stakeholders understand why a certain decision was made, what factors were considered, and what the possible outcomes are. XAI tools aim to explain how AI mechanisms like deep learning models, cognitive neuroscience algorithms, and machine Learning algorithms arrive at their decision⁵⁶, by employing simpler models to provide explanations about the working of the *black box*.⁵⁷

50 See: <https://arxiv.org/abs/1908.09635>, last visited on April 3, 2023.

51 See: <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G>, last visited on April 3, 2023.

52 See: <https://arxiv.org/abs/1908.09635>, last visited on April 3, 2023.

53 See: <https://www.ibm.com/in-en/watson/explainable-ai>, last visited on March 29, 2023.

54 XAI, as coined by DARPA, is the method that helps human experts understand solutions developed by AI (See: <https://neptune.ai/blog/explainability-auditability-ml-definitions-techniques-tools>), last visited on March 29, 2023.

55 See: <https://www.sciencedirect.com/science/article/pii/S0740624X21001027>, last visited on March 29, 2023.

56 See: <https://www.ibm.com/in-en/watson/explainable-ai>, last visited on March 29, 2023.

57 The goal of XAI is to address the black box problem, which is the challenge of understanding how an AI system arrived at its decision or recommendation. See: <https://christophm.github.io/interpretable-ml-book/explanation.html>; <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8466590>, last visited on March 29, 2023.

4. Concerns and Disruption of Regulations

XAI tools are highly beneficial, primarily in instilling improved trust in AI systems, facilitating better decision-making, and ensuring increased accountability.⁵⁸ For example, in the healthcare sector, XAI can help doctors understand how AI models arrived at their diagnoses, enabling them to make more informed decisions.⁵⁹

For organizations that tend to rely on AI tools to any degree, XAI could become crucial to build and maintaining stakeholder trust by increasing transparency and accountability in the decision-making process of AI systems. When an AI system makes a decision that affects people's lives, stakeholders must understand how that decision was made. If the decision-making process is not transparent or explainable, stakeholders may lose trust in the system and the organization using it.⁶⁰

BigTech entities like Google⁶¹ are actively adapting and promoting XAI, which shows that it is slowly gaining significance — especially in terms of providing the user with the power to ensure that AI is driven by the business, and not the other way around.

F. Content Regulation

Assuming that GAI is not considered an intermediary and not accorded a safe harbour for the outputs generated by such AI, a key question is whether such AI can generate unlawful content. Developers generally employ guardrails to ensure that certain content is never present in the output generated, including obscene content. For e.g., OpenAI's policy does not allow the usage of their models for illegal activity, child sexual abuse material, generation of hateful, harassing, or violent content, generation of malware, etc.⁶² However, there are certain key concerns that may still arise:

Firstly, developers would likely not be aware of all applicable content regulations in order to build appropriate guardrails. The guardrails would have likely been built based on the legal regime of where the developers/operators are based.

Secondly, there are conflicting standards applicable across jurisdictions on what is protected by the right to free speech. For this reason, unless developers comply with the strictest possible standard for content, the same output is likely to be unlawful in one jurisdiction, while being compliant in another.

Thirdly, even if such guardrails are sought to be built as per applicable laws, the legal position may not be so precisely defined that safeguards can be built into the AI. For e.g., in order to determine if certain content contains hate speech, courts need to analyse the facts on a case-by-case basis, in light of the factors laid down under the law. Such a determination may not be possible by an AI, especially in borderline cases.

Fourthly, even if precise guardrails are built in, the AI may still generate output that violates the operator's policies. For e.g., CyberArk was able to get ChatGPT to create a new strand of malware after repeated requests and constraining the output available.⁶³ This would be against OpenAI's policy, nevertheless, users were able to bypass this policy.

58 See: <https://neptune.ai/blog/explainability-auditability-ml-definitions-techniques-tools>, last visited on March 29, 2023.

59 See: <https://bmcmmedinformdecismak.biomedcentral.com/articles/10.1186/s12911-020-01332-6>, last visited on March 29, 2023.

60 See: <https://christophm.github.io/interpretable-ml-book/explanation.html>; <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8466590>, last visited on March 29, 2023.

61 See: <https://cloud.google.com/explainable-ai#:~:text=Explainable%20AI%20is%20a%20set,others%20understand%20your%20models'%20behavior>, last visited on March 29, 2023.

62 See: <https://platform.openai.com/docs/usage-policies/disallowed-usage>, last visited on March 29, 2023.

63 See: <https://www.infosecurity-magazine.com/news/chatgpt-creates-polymorphic-malware/>, last visited on March 29, 2023.

4. Concerns and Disruption of Regulations

Thus, the possibility of unlawful content being made available by GAI is very much real. Further, in case GAI does create such unlawful content, it would be interesting to see how such content is dealt with by law enforcement authorities, as well as how liability is ascertained for such content.

In India, the Central Government (and its authorised officers) has the power to direct Government agencies and intermediaries to block content available online⁶⁴ and routinely does so.⁶⁵ Such orders can be issued by the Government on specific grounds namely, in the interest of sovereignty and integrity of India, defence of India, security of the State, friendly relations with foreign States or public order or for preventing incitement to the commission of any cognizable offence relating to the above.⁶⁶ Importantly, the power is with respect to blocking such content for access by the public.⁶⁷ However, the content generated by GAI may be at the behest of the user and be available only to the user. It may not always be broadcasted to the public at large. However, in business use cases, the output of GAI may form part of the business' product made available to its customers.

In case of casual users of GAI, even if the output contains unlawful content, only the user would be able to view such content. Law enforcement authorities may not come to know of such content unless the user complains of such output being provided or the user themselves making the output publicly available. In business use cases, the output of GAI may be more publicly available. Either way, once law enforcement authorities are made aware of such content, they may require operators or business users of GAI to ensure that similar content is not published for other users. They may argue that operators/business users have (or should develop) the necessary technical measures to ensure content moderation for outputs which are established to be unlawful. Issues such as these will invariably crop up as the use of GAI becomes more popular.

How is liability determined?

If an algorithm does generate unlawful content, the logical next question would be regarding the apportioning of liability. If the nature of the AI is such that a human can be identified for causing and/or controlling the unlawful output, the solution is much simpler.⁶⁸ However, when AI “autonomously” act to create such content in a way which can neither be foreseen nor explained, the answer is not as straightforward. In such a case, there are three eligible candidates—the developer of the AI, the user who prompted the AI to generate that content, and the AI itself. The liability may be on one of these, or even a combination of these.

Imposition of liability under law would need to take into account aspects of morality (i.e., whether it is justified to impose liability on a particular entity), control (whether such entity can actually control the outputs in a feasible way) and efficacy (whether law enforcement authorities can act against such entities to make the liability effective). As with any other law, the positives of such provisions (reducing risks from unlawful content) would need to be balanced with any negative consequences (e.g., discouraging technological development).

64 Section 69A of the IT Act.

65 The government issued orders directing the blocking [of] a total of 2799, 3635, 9849, 6096 and 6775 URLs during the year 2018, 2019, 2020, 2021 and 2022 respectively. See: <https://www.thehindu.com/news/national/internet-blocking-committee-met-53-times-in-2022-it-ministry/article66484948.ece>, last visited on March 29, 2023.

66 Section 69A of the IT Act.

67 Section 69A of the IT Act.

68 This would form a classic case of the Perpetration-via-Another Liability Model put forth by Hallevy, where the AI is considered an ‘innocent agent’, or a mentally limited person such as a child. Therefore, under this model, the liability would either be on the programmer of the AI software or the end user.

4. Concerns and Disruption of Regulations

There are numerous possible solutions to questions of liability. Scherer proposes a certification regime where manufacturers and operators of certified AI enjoy limited civil liability while those of uncertified AI would be strictly liable for the actions of their AI.⁶⁹ It has also been suggested that the ease with which liability is imposed on manufacturers and users for actions of AI should also depend on the seriousness of the offence committed.⁷⁰

In addition to arguments in favor of a strict liability regime, there are also calls for further transparency in the black box that AI algorithms are considered. Transparency will require algorithms to adhere to a certain level of complexity, beyond which humans may not be able to comprehend the working of the algorithms.⁷¹ Based on these, Bathaee suggests a sliding-scale approach towards ascertaining liability, which is based on the degree of transparency in the AI and degree of human supervision.

In case of criminal liability, however, proof of mens rea, or a guilty mind, is required in order to apportion liability. Given the difficulty to impute voluntariness, negligence, and intention to AI systems, it has been suggested that for now, any liability on AI should not be criminal but merely administrative, or “something in-between”.⁷²

Finally, there have been calls for making the AI itself criminally liable which may or may not be accompanied with recognition of AI as a legal personality with its own set of rights and obligations.

The European Commission (“EC”) has proposed a regulatory framework on AI to provide AI developers, deployers and users with clear requirements and obligations regarding specific uses of AI.⁷³ The framework is based on the understanding that different AI use cases may involve different degrees of risk. Nevertheless, there is a need to balance the problem-solving ability of AI systems with avoiding undesirable outcomes that certain AI may produce. The Regulatory Framework defines 4 levels of risk in AI: (i) Unacceptable risk; (ii) High risk; (iii) Limited risk, and (iv) Minimal or no risk. While AI systems with unacceptable risk (e.g., encouraging dangerous behaviour) are proposed to be banned, those posing high risks (such as transport and exam scoring) will be subject to strict obligations before being released into the market. There will be minimal pre-release obligations for limited and minimal risk AI systems which, the EC notes, forms the majority of AI systems currently in place.

While a definitive outcome of these proposals is uncertain at this stage, it is likely that the best possible approach will only present itself once these approaches are experimented upon and more data is available on their respective impacts.

69 Matthew U. Scherer, *Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies*, 29(2) *Harvard Journal of Law & Technology*, Spring 2016.

70 Ibid.

71 Yavar Bathaee, *The Artificial Intelligence Black Box and the Failure of Intent and Causation*, 31(2) *Harvard Journal of Law & Technology*, Spring 2018.

72 Dafni Lima, *Could AI Agents Be Held Criminally Liable: Artificial Intelligence and the Challenges for Criminal Law*, 69 *South Carolina Law Review* (2018).

73 See: <https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai>, last visited on March 29, 2023.

G. Generative Artificial Intelligence Models: Intermediaries or Publishers?

The Information Technology Act, 2000 (“IT Act”) defines an ‘intermediary’ with respect to a particular electronic record.⁷⁴ The definition of “electronic record” is fairly broad, and means “*data, record or data generated, image or sound stored, received or sent in an electronic form or micro film or computer generated micro fiche*”.⁷⁵ Hence, outputs provided by GAI, whether in the form of text, images, sounds, etc., would be considered electronic records.

An intermediary, under the IT Act, with respect to any particular electronic record, means “*any person who on behalf of another person receives, stores or transmits that record or provides any service with respect to that record and includes telecom service providers, network service providers, internet service providers, web-hosting service providers, search engines, online payment sites, online-auction sites, online-market places, and cyber cafes*.”

However, being an intermediary is not sufficient to avail safe harbor under the IT Act. Safe harbour is provided only for *third-party content* hosted or made available by the intermediary if it satisfies certain conditions. Otherwise, the intermediary may be held liable under various regulations relevant for the content hosted or made available by it.

As per OpenAI, ChatGPT’s model was trained on vast amounts of data from the internet written by humans, including conversations, and the system is designed to maximise the similarity between outputs and the dataset the models were trained on.⁷⁶ Further, ChatGPT was “*optimized for dialogue by using Reinforcement Learning with Human Feedback (RLHF) — a method that uses human demonstrations and preference comparisons to guide the model toward desired behaviour*.”⁷⁷ Developers also exercise discretion on certain content to be removed from training data.⁷⁸

Hence, not only are the datasets in which GAI are trained may be selected by humans but the kind of processing undertaken by the programme could also be supervised by humans. This is unlike, say, a search engine that provides links to webpages based on a human prompt but does not process or modify the information on the webpage itself.

Operators of GAI may argue that the third-party information on which the programme has been trained on, is merely being presented to the user after an automated process. They may rely on the decision in *Myspace Inc. v. Super Cassettes Industries Limited*,⁷⁹ (“**Myspace case**”) where it was held that intermediaries could be held liable only when they have actual or specific knowledge and not constructive knowledge of the existence of infringing content on their website, and do not take any steps to have such content removed. Accordingly, operators may argue they do not have actual or specific knowledge of the final output being provided by the AI programme.

Further, in the *Myspace case*, it was held that modifications made to the format of the content, both via an automated process and without manual intervention, would not amount to having actual control, actual knowledge, or a ‘reason to believe’ that the content uploaded may be infringing of third party rights.

⁷⁴ Section 2(w) of the IT Act.

⁷⁵ Section 2(t) of the IT Act.

⁷⁶ See <https://help.openai.com/en/articles/6783457-chatgpt-general-faq>, last visited on March 29, 2023.

⁷⁷ See <https://help.openai.com/en/articles/6783457-chatgpt-general-faq>, last visited on March 29, 2023.

⁷⁸ For e.g., explicit content was removed while training Dall-E 2, see <https://openai.com/dall-e-2/>, last visited on March 29, 2023.

⁷⁹ FAO(OS) 540/2011, C.M. APPL.20174/2011, 13919 & 17996/2015, decided on December 23, 2016.

4. Concerns and Disruption of Regulations

GAI also makes automated changes to content without any manual intervention. Nevertheless, it is important to note that in the *Myspace* case, the content uploaded by third parties was not being modified, except for the format itself and the insertion of some advertisements which may not be true for GAI. GAI are not purely conduits of data since (i) they provide outputs based on the user inputs, and (ii) modify the data at their disposal to create a new output. Hence, traditional notions of intermediaries being in the nature of conduits (such as internet service providers or search engines) are unlikely to be applicable to GAI. Basis the above, it would be difficult to argue that GAI operators are intermediaries under the law.

With the Digital India Act also in the works (as of the date of writing), the Government has indicated that the new legislation will treat intermediaries differently based on their business models. Hence, conditions for availing safe harbor may also differ based on the kind of services provided by the respective intermediary. Hence, it would be interesting to see if GAI would be one of the categories of intermediaries recognized by law, and if so, the kind of safe harbor provided to it.

H. Competition Law Concerns

While GAI is currently at its nascent stage, it is set to raise numerous competition law concerns as the industry gets bigger and more consolidated. It is no surprise that some of the leading big tech companies are already at the forefront of the GAI boom. The GAI industry may tend to reward economics of scale and ultimately rewarding significant investments, resources and capabilities with significant revenues. This may, in turn, make the industry prone to being dominated by “gatekeepers”, paving the way for potential abuse of dominant positions.

Some of the factors which may create entry barriers or reward dominant players are:

- Training GAI requires massive amounts of curated data. Existing big tech players have access to such data, almost globally. Therefore, with a large amount of the necessary data already at their disposal, it should be easier for such entities to train AI which they intend to deploy. On the other hand, non-incumbent players would first need to invest in purchasing or licensing the data.
- Training GAI also requires considerable computing power and storage capacity, both of which can be provided by cloud service providers. Hence, entities which are dominant in the cloud storage and cloud services business may also stand to gain from further development of GAI.
- As a result of the investments required to generate valuable AI applications, smaller players may be unable to generate revenues and growth due to the sheer lack of resources.
- Hoppner and Streatfeild summarise these factors and note that the characteristics of the stack required to build and deploy GAI has all ingredients for winner-takes-all battles: “(i) Vertical integration in closely interrelated markets, with (ii) upstream dominance and downstream value generation, (iii) unequal access to proprietary resources, (iv) issues of interoperability, (v) data portability, (vi) non-transparency, (vii) IP licensing, and (viii) platform usage fees and conditions vis-à-vis dependent business users, etc.”⁸⁰

It would need to be evaluated if the dominance of some players by itself could be harmful for the GAI market, and whether legislations such as the Digital Services Act and Digital Markets Act of the EU would be necessary or (in)effective in the long run. In India, emerging technologies and new-age intermediaries have been proposed to be dealt with under the Digital India Act.

⁸⁰ See “ChatGPT, Bard & Co.: An Introduction to AI for Competition and Regulatory Lawyers”, 9 Hausfeld Competition Bulletin (1/2023).

4. Concerns and Disruption of Regulations

Additionally, the government has also proposed to introduce a law to regulate digital competition. However, there is limited visibility on how these statutes will actually unfold and what their scope will be while assessing GAI players from an anti-trust law perspective.

I. Misuse Cases of Generative Artificial Intelligence

As stated above, GAI can be used to generate undesirable and harmful content, even if the developers have countermeasures in place to avoid the same. As GAI becomes more advanced and powerful, so will the potential for its misuse. Following are some of the present instances of GAI misuse:

Phishing

Phishing is a form of cybersecurity attack where hackers send communications in the form of text or emails, posing as reputed individuals. Phishing emails deceive users into downloading harmful software, clicking on dangerous links, or disclosing private information.

Phishing emails and messages can be convincingly written by hackers using text-to-text GAI, making them harder to spot and avoid.⁸¹ GAI's real human-like writing style can be easily mistaken as if it's from a reliable source for duping users.

The wording or bait in the email is often the same in mass phishing attacks. As a result, it is simple for security providers to create detection criteria based on the language, as well as for automated filters. However, with the help of GAI, attackers can create many distinct variations of the same phishing email and they can even automate it to bypass cybersecurity systems.⁸²

Deepfakes

Deepfakes refer to AI-generated deceptive digital content. Digital content, mostly, in the form of images or videos are manipulated to show persons saying and doing things starkly different from the original content. The biggest threat posed by deepfakes is the circulation of false information that is so convincingly believable that it is hard to separate original content from fake content. For example, a deepfake video was created of Ukrainian president Volodymyr Zelenskyy, commandeering his army to surrender during the Russia-Ukraine conflict.⁸³

Deepfakes work on GAN which recognizes patterns of angles and movement in the original content to effectively replicate it with the desired changes. Based on what content is fabricated using deepfake technology, deepfakes adversely affect public opinion⁸⁴, lead to reputational harm⁸⁵ and can also influence the stock market.⁸⁶

81 See: <https://www.blackhat.com/docs/us-16/materials/us-16-Seymour-Tully-Weaponizing-Data-Science-For-Social-Engineering-Automated-E2E-Spear-Phishing-On-Twitter-wp.pdf>, last visited on April 3, 2023.

82 See: <https://www.csoonline.com/article/3685488/how-ai-chatbot-chatgpt-changes-the-phishing-game.html>, last visited on April 3, 2023.

83 See: <https://www.bbc.com/news/technology-60780142>, last visited on April 3, 2023.

84 See: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2982233, last visited on April 3, 2023.

85 See: https://regmedia.co.uk/2019/10/08/deepfake_report.pdf, last visited on April 3, 2023.

86 See: <https://www.bbc.com/news/world-us-canada-21508660>, last visited on April 3, 2023.

4. Concerns and Disruption of Regulations

Deepfakes can be shared and circulated at high-speed, altering perception and reality to a large chunk of the human population. Studies have revealed that unverified events are more widely circulated than verified events leading to misinformation crisis.⁸⁷ Further, people don't feel the need to question their judgment of what is true and what is not, if they find other people making such similar judgments. Certain algorithms on the internet keep bringing up deepfake videos and images, resultantly, the user keeps getting sucked into a distorted reality. Such being the case, propaganda and misinformation can be easily spread, creating ripples in the society.⁸⁸

Malware

Malware, or malicious software, is a catch-all word for viruses and other dangerous computer programmes that cybercriminals employ to infect systems and access private data.

The use of GAI for creating malware could be of serious concern. For example, AI-generated polymorphic malware is engineered to continuously alter its signature files using different decryption techniques. As a result, many conventional cybersecurity tools that rely on signature-based detection, such as antivirus or antimalware solutions, may be unable to identify and eliminate such malware.⁸⁹

Hackers benefit from the ease of automation provided by GAI as multiple codes can be created in a short span of time. Apart from this, GAI can enhance the capacity of malware by aiding in selection of targets and extraction of authentication factors. Lastly, GAI can be trained for code obfuscation and code behavior adaption to produce malware that can evade malware detection tools.

Plagiarism

Plagiarism is the act of presenting someone else's ideas or works as one's own without giving them full credit. Students may utilise text-to-text GAI to quickly manufacture written work without investing the time and effort necessary to create unique content.

The capacity of GAI to employ natural language and grammar to create original, nearly personalised text even when the content was taken from a database is a significant problem. As the content produced by GAI has never been written before, it hasn't yet tripped any traditional plagiarism detectors, making it difficult for assessors to spot cheats. Due to this, many educational institutions have resorted to banning the usage of GAI like ChatGPT by students.⁹⁰

Furthermore, there are heightened perils of using GAI to write research papers on topics that require a certain level of expertise. The generated content might not always accurately reflect the particular study domain, which could result in errors or discrepancies in the research information produced.⁹¹

87 See: <https://arxiv.org/abs/1511.07487>, last visited on April 3, 2023.

88 See: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3213954, last visited on April 3, 2023.

89 See: <https://arxiv.org/ftp/arxiv/papers/1406/1406.7061.pdf>, last visited on April 3, 2023.

90 See: <https://www.businessinsider.in/tech/news/here-are-the-schools-and-colleges-that-have-banned-the-use-of-chatgpt-over-plagiarism-and-misinformation-fears/articleshow/97452660.cms>, last visited on April 3, 2023.

91 See: <https://ccforum.biomedcentral.com/articles/10.1186/s13054-023-04380-2>, last visited on April 3, 2023.

4. Concerns and Disruption of Regulations

Building Dangerous Weapons

The extensive capabilities and accessibility of GAI can create security problems by making critical information readily available. To understand the extent to which GAI could be misused for building dangerous weapons, the experiments conducted on MegaSyn, a drug-discovery GAI used by biotech companies can be studied.⁹² Megasyne was built for testing different chemicals to assess their toxicity levels for human consumption. However, when the AI model was repurposed to design toxic chemicals rather than treatment compounds, the results generated were alarming. The AI model readily produced 40,000 chemical warfare agents. In fact, it created novel molecules that were more dangerous than previously used chemical agents.

Additionally, automated weapons such as drones can be hacked using GAI to create malicious codes.⁹³ A drone could deviate from its initial target and cause unintended casualties as a result of hacking. Similar problems could arise if the data being used to train the AI deployed in drones is poisoned through such hacking.⁹⁴

92 See: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9544280/>, last visited on April 3, 2023.

93 See: <https://ieeexplore.ieee.org/document/9831441>, last visited on April 3, 2023.

94 Ibid.

Way Forward

As corporations and governments increasingly embed AI in their products, processes, and decision-making systems, concerns are being raised over how this data is being used: especially by complex algorithms in, say, diagnosing cancer, disbursing loans, rating employee performance, etc. Apart from sector/use-case-specific concerns, overall concerns regarding bias, transparency, ownership, accountability and due process in AI systems would also need to be tackled.¹

Moreover, social concerns about potential job losses and unemployment due to the increasing demand for GAI technology cannot be ignored.² However, if viewed from a different lens, this fear can also act as a catalyst for reskilling and upskilling the workforce to adapt to the evolving job market. While AI and automation may replace some jobs, they also generate new opportunities for skilled professionals who can design, develop, maintain, and operate these technological systems.

On the other hand, consumer facing GAI empowers laypersons with unprecedented abilities. While this opens up the possibility of a different kind of creativity, it also leads to concerns regarding the creation of misinformation and the value of intellectual property generally. These issues are further exacerbated by the fact that both end users and regulators may not understand the functioning of each distinct GAI — making regulation challenging.

The obscurity and inscrutability of GAI also make it difficult to determine what is at stake and to frame regulations accordingly— as is the case with most emerging technologies. While countries and supra-national bodies such as the IEEE and OECD have come up with strategy papers, vision documents, and ethical guidelines — *all non-binding*, the European Union is proposing an AI law that is risk-based. This could pave the way for significant future jurisprudence.

Also, regulators across the world should adopt an open and consultative approach where stakeholders' views are encouraged and meaningfully considered. This is especially important since this space is highly dynamic and conventional positions on legal, policy or hastened ad hoc solutions may not stand the test of time. It is important to not regulate too early, or too late, while also balancing the degree of regulation. In such a situation, self-regulation by the industry could be a viable stop-gap solution. The industry may be best placed to understand the challenges that need to be addressed while being mindful of over-regulation.

In the long run, however, considering the unique challenges posed by GAI, the right combination of technical tools, accountability, industry self-regulation, and sectoral laws seems like the way forward.

In summary, the training of ChatGPT did not involve infringement of existing copyrighted works, as OpenAI took steps to ensure that its training data was properly sourced and did not contain sensitive or copyrighted material. OpenAI is also committed to promoting the responsible use of its technology and preventing its language models from being used for nefarious purposes.

¹ Malte Ziewitz, *Governing Algorithms: Myth, Mess, and Methods*, Science, Technology & Human Values, 2016 Vol 4(l) pp. 3-16.

² Generative AI set to affect 300mn jobs across major economies, available at: <https://www.ft.com/content/7dec4483-ad34-4007-bb3a-7ac925643999>, last visited on April 7, 2023.

About NDA

At Nishith Desai Associates, we have earned the reputation of being Asia's most Innovative Law Firm — and the go-to specialists for companies around the world, looking to conduct businesses in India and for Indian companies considering business expansion abroad. In fact, we have conceptualized and created a state-of-the-art Blue Sky Thinking and Research Campus, Imaginarium Aligunjan, an international institution dedicated to designing a premeditated future with an embedded strategic foresight capability.

We are a research and strategy driven international firm with offices in Mumbai, Palo Alto (Silicon Valley), Bangalore, Singapore, New Delhi, Munich, and New York. Our team comprises of specialists who provide strategic advice on legal, regulatory, and tax related matters in an integrated manner basis key insights carefully culled from the allied industries.

As an active participant in shaping India's regulatory environment, we at NDA, have the expertise and more importantly — the VISION — to navigate its complexities. Our ongoing endeavors in conducting and facilitating original research in emerging areas of law has helped us develop unparalleled proficiency to anticipate legal obstacles, mitigate potential risks and identify new opportunities for our clients on a global scale. Simply put, for conglomerates looking to conduct business in the subcontinent, NDA takes the uncertainty out of new frontiers.

As a firm of doyens, we pride ourselves in working with select clients within select verticals on complex matters. Our forte lies in providing innovative and strategic advice in futuristic areas of law such as those relating to Blockchain and virtual currencies, Internet of Things (IOT), Aviation, Artificial Intelligence, Privatization of Outer Space, Drones, Robotics, Virtual Reality, Ed-Tech, Med-Tech and Medical Devices and Nanotechnology with our key clientele comprising of marquee Fortune 500 corporations.

The firm has been consistently ranked as one of the Most Innovative Law Firms, across the globe. In fact, NDA has been the proud recipient of the Financial Times–RSG award 4 times in a row, (2014-2017) as the Most Innovative Indian Law Firm.

We are a trust based, non-hierarchical, democratic organization that leverages research and knowledge to deliver extraordinary value to our clients. Datum, our unique employer proposition has been developed into a global case study, aptly titled 'Management by Trust in a Democratic Enterprise,' published by John Wiley & Sons, USA.

Research@NDA

Research is the DNA of NDA. In early 1980s, our firm emerged from an extensive, and then pioneering, research by Nishith M. Desai on the taxation of cross-border transactions. The research book written by him provided the foundation for our international tax practice. Since then, we have relied upon research to be the cornerstone of our practice development. Today, research is fully ingrained in the firm's culture.

Over the years, we have produced some outstanding research papers, reports and articles. Almost on a daily basis, we analyze and offer our perspective on latest legal developments through our "Hotlines". These Hotlines provide immediate awareness and quick reference, and have been eagerly received. We also provide expanded commentary on issues through detailed articles for publication in newspapers and periodicals for dissemination to wider audience. Our NDA Labs dissect and analyze a published, distinctive legal transaction using multiple lenses and offer various perspectives, including some even overlooked by the executors of the transaction. We regularly write extensive research papers and disseminate them through our website. Our ThinkTank discourses on Taxation of eCommerce, Arbitration, and Direct Tax Code have been widely acknowledged.

As we continue to grow through our research-based approach, we now have established an exclusive four-acre, state-of-the-art research center, just a 45-minute ferry ride from Mumbai but in the middle of verdant hills of reclusive Alibaug-Raigadh district. Imaginarium AliGunjan is a platform for creative thinking; an apolitical eco- system that connects multi-disciplinary threads of ideas, innovation and imagination. Designed to inspire 'blue sky' thinking, research, exploration and synthesis, reflections and communication, it aims to bring in wholeness — that leads to answers to the biggest challenges of our time and beyond. It seeks to be a bridge that connects the futuristic advancements of diverse disciplines. It offers a space, both virtually and literally, for integration and synthesis of knowhow and innovation from various streams and serves as a dais to internationally renowned professionals to share their expertise and experience with our associates and select clients.

We would love to hear from you about any suggestions you may have on our research publications. Please feel free to contact us at research@nishithdesai.com.

Recent Research Papers

Extensive Knowledge gained through our Original Research is a Source of our Expertise.



January 2023

Doing Business in India

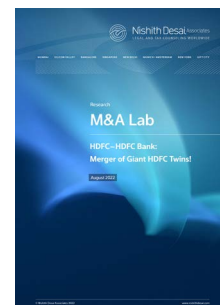
The Guide for US Businesses and Organizations Entering and Expanding into India



September 2022

Ahead of the Game

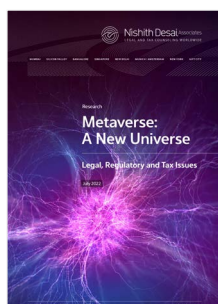
Legal, Regulatory and Tax Issues in India's Gaming Industry



August 2022

M&A Lab

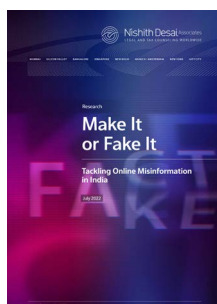
HDFC-HDFC Bank:
Merger of Giant HDFC Twins!



July 2022

Metaverse: A New Universe

Legal, Regulatory and Tax Issues



July 2022

Make It or Fake It

Tackling Online Misinformation
in India



April 2022

Private Equity and Private Debt Investments in India

Regulatory, Legal and Tax Overview

For more research papers [click here](#).



Nishith DesaiAssociates
LEGAL AND TAX COUNSELING WORLDWIDE

MUMBAI

93 B, Mittal Court, Nariman Point
Mumbai 400 021, India
Tel +91 22 6669 5000

SILICON VALLEY

220 S California Ave., Suite 201
Palo Alto, California 94306, USA
Tel +1 650 325 7100

BANGALORE

Prestige Loka, G01, 7/1 Brunton Rd
Bangalore 560 025, India
Tel +91 80 6693 5000

SINGAPORE

Level 24, CapitaGreen
138 Market St
Singapore 048 946
Tel +65 6550 9855

NEW DELHI

13-H, Hansalaya Building, 15
Barakhamba Road, Connaught Place
New Delhi 110 001, India
Tel +91 11 4906 5000

MUNICH / AMSTERDAM

Maximilianstraße 13
80539 Munich, Germany
Tel +49 89 203 006 268

NEW YORK

1185 6th Avenue, Suite 326
New York, NY 10036, USA
Tel +1 212 464 7050

GIFT CITY

408, 4th Floor, Pragya Towers
GIFT City, Gandhinagar
Gujarat 382 355, India

Generative AI & Disruption
Emerging Legal and Ethical Challenges