



The Growth of Generative AI: Two Sets of Legal Issues Arising

G. ELIAS

Introduction

In recent years, the field of artificial intelligence (“AI”) has witnessed remarkable growth, one that has given rise to a new era of technological innovation and creativity. At the forefront of this revolution is Generative AI, a branch of artificial intelligence that can generate human-like text, images, and even entire pieces of content autonomously. This in turn has opened a world of possibilities across various industries. However, with the rapid rise of Generative AI comes a host of complex legal issues.

This Article explores the intricate web of legal challenges that have emerged in the wake of Generative AI, shedding light on the delicate balance between innovation and regulation in the rapidly-evolving landscape of Generative AI. It is worth noting that the evolution of laws governing the use of emerging technologies like AI is expected to continue in the coming years, which is viewed positively. It is anticipated that the concept of fair use will offer robust protection in copyright infringement cases. Simultaneously, the protection of individual privacy will remain sacrosanct for data protection cases, and these developments are desirable trends.

Overview

Generative AI refers to computational techniques that are capable of generating seemingly new, meaningful content such as texts, images, or audio output from training data. These techniques are primarily based on generative modelling that is instantiated with a machine-learning architecture that can create new data samples based on learned patterns. They can also be used to assist humans as intelligent question-answering systems. The Generative AI applications consist of practical use cases and search engine optimization content-generation or code-generation¹.

Generative AI is not a brand-new innovation as it had been introduced in the 1960s in chatbots. The Eliza Chatbot created by Joseph Weizenbaum in the 1960s was one of the earliest examples of Generative AI. These early implementations used a rule-based approach that broke easily due to a limited vocabulary, lack of context, and overreliance on patterns, among other shortcomings². They were also difficult to customize and extend. However, the field saw a resurgence in the wake of advances in neural networks and deep learning in 2010 that enabled Generative AI to automatically learn to parse existing text, classify image elements, and transcribe audio.

In 2014, Ian Goodfellow introduced the Generative Adversarial Networks (“GANs”) – a machine learning algorithm that could create convincingly authentic images, videos, and audio output of real people.³ GANs were a deep learning technique that provided a novel approach for organizing competing neural networks to generate and then rate content variations that could generate realistic people, voices, music, and text. Since then, progress in other neural network techniques and architectures has helped expand Generative AI capabilities.

¹ Stefan Feuerriegel et al, “Generative AI” (ResearchGate, May 2023) <[PDF Generative AI \(researchgate.net\)](#)> accessed 14 August 2023.

² George Lawton, “What is Generative AI” (TechTarget, July 2023) <<https://www.techtarget.com/searchenterpriseai/definition/generative-AI>> accessed 14 August 2023.

³ *Supra*.

Various Generative AI tools exist for different modalities, and they include image generation tools – DALL-E 2, Midjourney, and Stable Diffusion; music generation tools – Amper, Dadabots, and MuseNet; text generation tools – GPT, Jasper, AI-Writer, and Lex; code generation tools – CodeStarter, Codex, Github Copilot and Tabine; voice synthesis tools – Descript and, Listnr and Podcast AI.

Potential Opportunities and Issues

Generative AI is a general-purpose technology that can be applied extensively across many industries. It can easily interpret and understand existing content, automatically create new content, automate the manual process of writing content, summarize complex information into a coherent narrative, and simplify the process of creating content in a particular style.

For instance, law firms can use Generative AI to draft and interpret contracts, analyze evidence, and suggest arguments; film and media companies can use Generative AI to produce content and translate it into other languages with the actors' voices; architectural firms can use it to design and adapt prototypes more quickly. In the gaming sector, Generative AI can be used to design game content and levels.

Despite the many advantages of Generative AI, concerns exist as to the misuse and abuse of the technology and the potential legal issues it creates. For instance, Generative AI can provide inaccurate and misleading information, promote new kinds of plagiarism that ignore the rights of content creators and artists of original content, can be used to generate fake news, and even impersonate people for cyber-attacks.

In this article, we have classified most of these concerns into two broad classes of legal issues: intellectual property exploitation and data protection and privacy issues. Below, we look more closely at these two classes of issues and at other issues that do not fall neatly into either of the two classes.

Intellectual property issues:

Image-generator tools such as Stable Diffusion or Midjourney can produce great visuals in styles while text generators can write essays, poems, power points, and summaries. While it may appear that these AI tools can conjure new material from almost nothing, that's usually not the case. Generative AI platforms are trained on data lakes and question snippets – billions of parameters that are constructed by software processing huge archives of images and text⁴. Generative AI recovers patterns and relationships, which they then use to create rules, and then make judgments and predictions when responding to a prompt. This process comes with legal risks, including intellectual potential property infringement as seen in the American case of **Andersen et al v Stability AI Ltd.** (the "**Andersen Suit**")⁵.

⁴ Gil Appel et al "Generative AI has an Intellectual Property Problem" (HBR, 07 April 2023) <<https://hbr.org/2023/04/generative-ai-has-an-intellectual-property-problem>> accessed 15 August 2023.

⁵ Case No. 3:23-cv-00201-WHO Andersen et al v Stability AI Ltd (2023) not yet reported.

The case was instituted in the US in January 2023 by three artists on behalf of themselves and a class of similarly-situated individuals against the AI image generator companies Stable Diffusion, Midjourney, and Deviant Art (the “**AI Companies**”). The Artists alleged that the AI Companies were misusing their work to train AI systems. One of the Claimants – Andersen – claimed that the AI Companies directly infringed copyrights she had registered in several of her works. Andersen further alleged that Stability AI scrapped billions of images from the internet to train its Stable Diffusion text-to-image system.

The plaintiffs claimed that Stability used their work without permission thereby infringing on their copyrights. The defendants on the other hand filed a motion to dismiss the claims of the plaintiffs arguing that the plaintiffs failed to specify the exact registered copyright works which were infringed. In July 2023, the Motion to dismiss the claims filed by the defendants was heard and a ruling is expected in the coming months. Also, in the case of *Getty Images (“Getty”), Inc v Stability AI*,⁶ Getty in February 2023 sued Stability AI for alleged brazen infringement of its intellectual property, stated that the AI Company copied more than 12 million images and their associated metadata without permission. Getty states that it has filed for copyright registrations on thousands of the images that it licences and has a copyright registration on its database of images. The database includes all the image metadata, containing information like the alt-text or description of the image, that tells the AI machine what it is looking at. Getty claimed that Stability AI used its images and database to train its AI machine because of examples of the output.

For instance, the AI-generated images by Stability AI contained the Getty Watermark. In response, Stability AI filed a motion to dismiss the suit for lack of personal jurisdiction, inability to join a necessary party, and failure to state a claim. The defendant argued that the courts of the district of Delaware where the matter was instituted lacked personal jurisdiction over Stability UK, which had no ties to Delaware except that its parent, Stability US was incorporated there. The defendant further argued that all the alleged acts giving rise to the complaint occurred outside Delaware and that the plaintiff only amended its complaint to include Stability UK after Stability US informed Getty Images’ counsel that Stability AI intended to move to dismiss for failure to join a necessary party.

The defendant thus argued that the amended complaint is defective because it improperly lumped together allegations against Stability US and Stability UK under the collective term “Stability AI”. Stability AI in the alternative requested that the suit be transferred to the Northern District of California, where a class action raising similar allegations against Stability AI is pending, to serve judicial economy and avoid the risk of inconsistent judgments. The plaintiff filed for jurisdictional discovery in response to the defendant's motion to dismiss. On September 22, 2023, the Court ordered the parties to meet and confer to determine the appropriate scope of jurisdictional discovery and that, if such agreement is not reached, the parties may be ordered to provide more letter briefing⁷.

⁶ Case 1:23-cv-00135-UNA Getty Images v Stability AI (Filed 02 March 2023) < <https://docs.justia.com/cases/federal/district-courts/delaware/dedce/1:2023cv00135/81407/1>> last accessed 06 September 2023

⁷ This refers to the process of submitting additional written arguments or briefs to court. It typically occurs when the parties involved in a legal case cannot agree on the scope of jurisdictional discovery and need to provide the court with further information and arguments to help the court make a decision.

Data Privacy issues

As stated earlier in the Overview, Generative AI tools operate by generating responses to prompts provided by users, using their training and algorithms to produce contextually-relevant and coherent text, images, or other output. These tools often require access to data, which may include personal or sensitive information which if not protected could pose a risk to data privacy in several ways⁸. At least 5 of these may be isolated here.

- i. **Data Breaches:** Where proper security measures are not put in place, Generative AI developers may be liable for data breaches that result from the unauthorized access or disclosure of sensitive user information. This can lead to privacy violations and potential misuse of personal data.
- ii. **Data Retention and Deletion Process:** If Generative AI tools retain user data for longer than necessary or fail to properly delete data upon request, there may be a risk of unauthorized access or misuse of personal information.
- iii. **Data Sharing:** Often Generative AI tools may share user data with third parties without explicit consent or for purposes beyond what was initially communicated. This can lead to unauthorized data sharing and potential privacy breaches.
- iv. **Promotion of Biases and Discrimination:** Generative AI tools may inadvertently perpetuate biases present in the training data. Where the training contains discriminatory patterns or biased information, the generated output will reflect and amplify those biases, thus further perpetuating discrimination against disadvantaged groups.
- v. **Inadequate Anonymization:** In instances where Generative AI requires access to personal or sensitive data for training or generating output if the anonymization techniques used are insufficient, there is a risk of re-identification. Where individuals can be identified from the generated data, it compromises their privacy.

From the foregoing, AI developers should ensure that they comply with the relevant laws on acquiring data used to train their models. While there are no specific laws guiding automated website scraping, AI developers need to pay close attention to intellectual property and data protection laws in order to avoid scraping copyrighted data and/or personal data. This is because there is no way to obtain consent for such data and the use may not fall under either any of the lawful bases for processing personal data, or the terms and conditions on various websites permitting or disallowing automated the scraping of data. When it concerns a potential copyright infringement, AI developers may rely on the doctrine of fair use as a defence based on a consideration of the following factors: the (i) purpose and character of its usage; (ii) nature of the work; (iii) amount and substantiality of the portion used in relation to the work as a whole; (iv) effect of the use upon the potential market or value of the work.⁹

AI developers may also obtain the necessary permissions and compensate those persons who own the intellectual property they seek to add to their training data. Compensation could be in

⁸ Vlere Hyseni "Generative AI and Data Privacy" (PECB, 7 June 2023) <<https://pecb.com/article/generative-ai-and-data-privacy>> accessed 15 August 2023.

⁹ Section 20(1), Copyright Act, 2022.

the form of sharing the revenue generated by the AI tool. It is prudent for AI companies to explore non-court resolution in cases of copyright infringement. Customers of AI tools, on the other hand, should also do their due diligence by reviewing the terms of service and privacy policies issued by the suppliers of the tools.

AI developers should also adopt practices that minimize the collection and retention of personal data such as informing users to refrain from disclosing personal and sensitive data while using the AI tool. Where personal data is required for the training of Generative AI models, such personal data should be adequately anonymized through the use of techniques such as data aggregation, masking or perturbation, and differential privacy. Also, developers of AI tools should comply with relevant data protection laws as they relate to obtaining necessary consent and ensuring proper data handling practices.¹⁰

Altogether, it is important for AI companies to band together and embrace self-regulation as a form of risk management to avoid being swamped in a plethora of legal battles that can derail them from their innovation and subsequently lead to bankruptcy.

Other Generative AI Copyright Infringement Claims

Two other Generative AI copyright infringement claims have become notorious and bear further scrutiny. We briefly discuss some of these cases below in addition to the ones already mentioned.

In one of them, ***Programmers’ Class Action v GitHub***, in November 2022, a group of programmers filed a class action against Microsoft, GitHub, and OpenAI alleging a breach of the Digital Millennium Copyright Act due to their unlicensed use of their software code to develop their different AI machines. These AI machines were trained with collections of publicly-accessible codes. The plaintiffs stated that their materials/codes were utilized without complying with the open-source licensing terms.¹¹ The defendants, in January, filed motions to dismiss based on the lack of the legal standing of the plaintiffs to bring the case because they did not state any specific damages suffered. Also, the defendants argued that the plaintiff failed to mention specific works that were infringed upon.

In ***Thomson Reuters Enterprise Centre v. Ross Intelligence Inc***¹², the plaintiffs, owners of Westlaw, a legal search platform, instituted a suit against Ross Intelligence Inc. – developers of a new AI-enabled legal search platform – for copyright infringement and tortious interference without a contract. In developing their AI model, the defendant partnered with LegalEase who allegedly “used a bot...to download and store mass quantities of proprietary information” belonging to the plaintiff.¹³

The defendant filed a motion to dismiss which was rejected by the court. Subsequently, they filed a motion for summary judgment basing their defence on the concept of fair use. The defendant,

¹⁰ In Nigeria, this is the Nigerian Data Protection Act 2023, Nigeria Data Protection Regulation 2019 (“NDPR”) and the NDPR Implementation Framework.

¹¹ Case 3:22-cv-06823 J.Doe 1 and J.Doe 2. v GitHub and Ors (GitHub Copilot Litigation, 11 March 2022) <https://githubcopilotlitigation.com/pdf/06823/1-0-github_complaint.pdf>last accessed 18 August 2023

¹² Case 1:2020-cv-00613 (Justia US Law 26 April 262022).

¹³ C.A. No. 20-613-LPS Thomson Reuters Enterprise Centre and Ors v. Ross Intelligence (District of Delaware, US Court, 26 April 2022) <https://www.ded.uscourts.gov/sites/ded/files/opinions/20-613_0.pdf> last accessed 18 August 2023.

not denying the allegations, argued that the use of Westlaw's content was functional and transformative and that the number of copyrighted materials used was of "little weight", and that they did not compete in the market for the copyrighted materials. The plaintiff, in its opposition, argued that the legal research product produced by the defendant would compete and replace Westlaw without any transformative purpose and that they copied in bad faith, accessing the Westlaw content illegally after it was refused a licence.¹⁴

The two above-mentioned cases are as-yet undecided. They raise key issues that would shape the discourse around the applicability of intellectual property rights concerning Generative AI usage. It would be interesting to see the extent to which the defence of fair use can avail AI companies in these cases and how the courts would interpret and apply the concept. It is far from obvious that the plaintiffs will be victorious. Already, in the case of *Authors Guild v. Google*¹⁵, the court ruled that Google's use of data from millions of copyrighted books over the Internet without a licence, to launch its book search project, constitutes fair use¹⁶.

These decisions could also potentially influence the methodology for obtaining training data for Generative AI models in the future.

The potential for the development of regulation around the use of AI in Nigeria.

While Nigeria usually plays catch-up with the rest of the world on technology and innovation, the increasing use of artificial intelligence tools and the rise of AI-based startups in Nigeria signify a fast adoption of AI in Nigeria. Presently, the Nigeria Information Technology Development Agency (**NITDA**) is working on drafting a National Artificial Intelligence Policy and a Nigerian code of practice for AI¹⁷. This is a commendable move as it has become important for policies to be developed to guide AI companies and the use of their underlying technologies, in addition to extant laws like the Copyright Act, 2022, Nigeria Data Protection Act, 2023, Nigeria Data Protection Regulation, 2019 and the Nigeria Data Protection Regulation 2019: Implementation Framework, 2020.

The Copyright Act, for instance, has copious provisions on copyright protection, the assignment of copyright in works, infringement of copyright, and how fair dealing constitutes an exception to copyright infringement. For example, where the use of a copyrighted work is for public interest provided that no revenue is derived, and the use does not substantially affect the potential market value of the original work.¹⁸ On the other hand, the data protection laws provide the lawful basis for the processing of personal data and how an AI Company can deal with processing the personal data of children and other individuals.

¹⁴ Case 1:20-cv-00613-SB (District of Delaware, US Court, 02 June 2023) Copyright Alliance <<https://copyrightalliance.org/wp-content/uploads/2023/02/TR-Response.pdf>> last accessed 28 August 2023.

¹⁵ Case 13-4829-cv (Justia US Law, 16 October 2015) <<https://www.scotusblog.com/wp-content/uploads/2016/01/Authors-Guild-v.-Google.pdf>> last accessed 06 September 2023.

¹⁶ Kyle Wiggers 'The current legal cases against generative AI are just the beginning' (TechCrunch, 27 January 2023)<<https://techcrunch.com/2023/01/27/the-current-legal-cases-against-generative-ai-are-just-the-beginning/>> last accessed 18 August 2023.

¹⁷ Ojukwu Emmanuel 'NITDA Drafting the Nigerian Code of Practice For Artificial Intelligence Tools Such as ChatGPT and Others'(Tekedia, 14 June 2023) <https://www.tekedia.com/nitda-drafting-the-nigeria-code-of-practice-for-artificial-intelligence-tools-such-as-chatgpt-and-others/> last accessed 06 September 2023.

¹⁸ Section 20, Copyright Act, 2022.

Authors



Eberechukwu Ezike
Associate
eberechukwu.ezike@gelias.com



Ayomipo Adeniyi
Associate
ayomipo.adeniyi@gelias.com



Uchenna Ezeagu
Associate
uchenna.ezeagu@gelias.com

LOCATIONS

LAGOS OFFICE
6 Broad Street
Lagos, Nigeria

T: +234 (1) 460 7890
E: gelias@gelias.com

ABUJA OFFICE
2nd Floor, Abia House,
Plot 979, First Avenue,
Central Business District
F.C.T, Abuja.

T: +234 (1) 888 8881

Practices • Arbitration • Banking • Capital Markets • Competition • Compliance • Corporate • Data Protection • Derivatives • Employment • Fintech • Foreign Investment • Intellectual Property • Litigation • Mergers and Acquisitions • Tax • "White Collar" Sanctions •

Sectors • Agribusiness • Commercial Banks • Commodities • Construction • Distributors • Development Finance • Electric Power • Entertainment • External Trade • Fintech • Healthcare • Infrastructure • Insurance • Investment Banks • Manufacturing • Media • Mining • Oil and Gas • Pension Managers • Private Equity • Real Estate • Services • Technology • Telecommunications • Transport •

www.gelias.com