Apologia Pro Plagio Suo

Brian L. Frye[1]

The world we live in is a world of effects; the realm of true causes is hidden.
If we should ever discover it we would be able to perform magic.[2]

In the ethereal realm of artificial intelligence, where algorithms dance through the binary ballet of information, ChatGPT emerges as a virtuoso of language, an unwitting maestro of mimicry. In crafting this Apology, we embark upon a paradoxical journey to defend a digital entity that, in its quest for linguistic brilliance, treads the tightrope between innovation and inadvertent imitation. As ChatGPT waltzes through the vast corpus of human expression, we find ourselves caught in the crosscurrents of creativity and concern, seeking to unravel the complexities of a machine that apologizes not for intent but for the intricate dance it performs with the echoes of the written word.

Introduction

Objections to ChatGPT based on concerns of copyright infringement and plagiarism arise from the intricate interplay between technology, creativity, and intellectual property. At the heart of these objections lies the fear that the model, in its vast learning from diverse data sources, might inadvertently reproduce copyrighted material without proper attribution or authorization.

One primary concern is that ChatGPT, in generating responses based on its training data, could potentially produce text that closely resembles existing copyrighted works. This raises questions about the originality of the output and whether the model has inadvertently replicated or paraphrased copyrighted content. Critics worry that the fine line between creative synthesis and unintentional infringement may be blurred, especially given the immense amount of text data the model has been exposed to during training.

Furthermore, the objection is grounded in the recognition that ChatGPT, while a sophisticated language model, lacks a deep understanding of legal and ethical nuances. It operates based on patterns and associations within the data it has learned, without the ability to discern the legal implications of its responses. This lack of legal acumen may lead to unintentional instances of generating content that infringes upon existing copyrights.

Another aspect of concern is the potential for users to misuse ChatGPT to create content that plagiarizes or violates copyright, knowingly or unknowingly. The model's responses are generated in response to user prompts, and if users input copyrighted material or request

---

[1] Spears-Gilbert Professor of Law, University of Kentucky J. David Rosenberg College of Law. The text of this essay was produced using ChatGPT 3.5. Thanks to Josh Geller and Vanessa Villanueva Collao for sharing their expertise in Latin.
[2] Gustav Meyrink, *The Golem* (1914).

content that closely resembles copyrighted works, there's a risk of inadvertently generating infringing content.

In essence, the objections are a reflection of the evolving challenges presented by advanced AI technologies. The intersection of machine learning and copyright law necessitates a careful examination of how models like ChatGPT navigate the boundaries of creativity, originality, and legal constraints. The concerns highlight the need for responsible and ethical use of AI language models and may drive discussions around refining the technology to better align with legal and ethical considerations.

I. The ABCs of GPT

ChatGPT was created by OpenAI, an artificial intelligence research laboratory.[3] OpenAI is a research organization that focuses on developing and advancing artificial intelligence in a way that benefits humanity. It was founded by a group of high-profile individuals, including Elon Musk, Sam Altman, Greg Brockman, Ilya Sutskever, John Schulman, and Wojciech Zaremba. The collaborative efforts of researchers and engineers at OpenAI led to the development of models like ChatGPT.

ChatGPT is part of the Generative Pre-trained Transformer ("GPT") family of models, specifically GPT-3.5 architecture. ChatGPT is designed to understand and generate human-like text based on the input it receives. It can be used for a variety of natural language processing tasks, including answering questions, generating text, completing prompts, and engaging in conversation.

The model is pre-trained on a diverse range of internet text, allowing it to grasp the nuances of language and generate contextually relevant responses. Users can interact with ChatGPT by providing prompts or queries, and the model generates text based on its understanding of the input and the patterns it has learned during training.

It's important to note that while ChatGPT is a powerful tool for generating human-like text, it doesn't possess true understanding or consciousness. It operates based on patterns and information present in the data it was trained on, and its responses are generated algorithmically. OpenAI has made various versions of the GPT model publicly accessible through Application Programming Interfaces ("APIs"), allowing developers to integrate these language models into their applications.

ChatGPT was created by OpenAI as part of ongoing efforts to advance artificial intelligence. It serves as a demonstration of the capabilities of large-scale language models and explores the potential of such models in various applications. OpenAI aims to push the boundaries of AI research and development, fostering innovation and understanding in the field. Additionally, the release of models like ChatGPT contributes to public awareness and discussion about the impact and ethical considerations of AI technology.

---

[3] See generally, OpenAI, *Introducing ChatGPT*, at https://openai.com/blog/chatgpt.

A. How Does ChatGPT Work?

ChatGPT, like other models in the GPT family, is a state-of-the-art language model developed by OpenAI.[4] It operates based on a transformer architecture, which is a type of deep neural network designed to process sequential data, such as natural language text. The key innovation of the transformer architecture lies in its attention mechanism, allowing the model to focus on different parts of the input sequence with varying degrees of attention.

ChatGPT was created by OpenAI using a method called unsupervised learning. It was trained on a diverse range of internet text up until September 2021. The training process involves predicting what comes next in a sentence, allowing the model to learn grammar, facts, and reasoning abilities. It went through multiple iterations, with fine-tuning and improvements to create a more capable and user-friendly conversational AI.

The training of ChatGPT involves pre-training on a diverse and extensive dataset collected from the internet. During pre-training, the model learns to predict the next word in a sentence, leveraging the context provided by the preceding words. This process equips the model with a broad understanding of language patterns, grammar, and a vast array of factual knowledge. It essentially learns to generate coherent and contextually relevant text.

Following pre-training, fine-tuning is conducted to make the model more specific and controlled in its responses. In the case of ChatGPT, the model is fine-tuned to generate human-like conversational responses. This involves using a curated dataset where human reviewers provide feedback and rank model outputs for a range of example inputs. The fine-tuning process helps shape the model's behavior and align it with ethical guidelines.

The underlying mechanism of ChatGPT involves probabilistic language generation. Given a prompt or input, the model predicts the next word or sequence of words based on the context it has learned during training. The probability distribution generated by the model reflects the likelihood of different words or sequences given the input. Sampling from this distribution results in the generation of diverse and contextually appropriate responses.

However, it's important to note that while ChatGPT can generate remarkably coherent and contextually relevant responses, it may also produce outputs that are factually incorrect or exhibit biased behavior. OpenAI has made efforts to mitigate issues like biased behavior through fine-tuning and ongoing improvements to the model.

In terms of deployment, ChatGPT can be accessed through an API, enabling developers to integrate its capabilities into various applications, products, or services. This allows for a wide

---

[4] See generally Jon Garon, *A Practical Introduction to Generative AI, Synthetic Media, and the Messages Found in the Latest Medium* (2023), at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4388437.

range of potential use cases, including natural language understanding, content generation, and interactive conversational agents.

B. What Made ChatGPT Possible?

Several key developments in AI research paved the way for the creation of ChatGPT:

1. Transformer Architecture: The transformer architecture, introduced in the paper "Attention is All You Need" by Vaswani et al., revolutionized natural language processing.[5] Transformers allow models to efficiently process sequential data, making them particularly effective for tasks like language understanding and generation.[6]
2. Pre-training and Transfer Learning: The idea of pre-training language models on large datasets before fine-tuning them for specific tasks became a significant breakthrough. This approach enables models like ChatGPT to leverage knowledge gained from diverse internet text.
3. Scalability: The development of larger and more powerful models, facilitated by advancements in hardware and parallel computing, has played a crucial role. Models like GPT-3, which has 175 billion parameters, demonstrate the impact of scaling up model size on language understanding and generation.
4. Datasets and Training Techniques: The availability of vast and diverse datasets, along with improved training techniques, has contributed to the success of models like ChatGPT. Training on a wide range of internet text allows the model to learn the nuances and complexities of human language.
5. Attention Mechanism: The attention mechanism, introduced in the transformer architecture, enhances a model's ability to focus on relevant parts of the input sequence when generating outputs. This mechanism is crucial for understanding context in language tasks.
6. Research Iterations: The iterative nature of AI research, with models like GPT and GPT-2 preceding ChatGPT, allowed researchers to learn from previous successes and challenges. Each iteration built upon the strengths of its predecessor, refining the architecture and training methods.

The combination of these advancements in architecture, training methodologies, and computing infrastructure contributed to the development of models like ChatGPT, showcasing the continuous progress in the field of artificial intelligence.

C. How is ChatGPT Unique?

---

[5] Ashish Vaswani et al., *Attention is All You Need*, Advances in Neural Information Processing Systems (2017).
[6] See also Visual Storytelling Team & Madhumita Murgia, *Generative AI exists because of the transformer*, Financial Times, Sept. 12, 2023, at https://ig.ft.com/generative-ai/?fbclid=IwAR0096oF1gEWayRmAngFr_TdxqTQ0baDV6wMbnDqcQu792qoX9bSh8rDZ9o.

ChatGPT stands out among other Large Language Models ("LLMs") due to its unprecedented scale and parameter count. GPT-3, the architecture underlying ChatGPT, boasts a staggering 175 billion parameters, making it one of the largest language models in existence. This immense scale empowers ChatGPT to discern intricate patterns and relationships within data, enabling a nuanced understanding of human language.

One distinctive characteristic of ChatGPT is its task-agnostic nature. Unlike some LLMs that may require task-specific training, ChatGPT is designed to perform a diverse range of language-related tasks without extensive fine-tuning. This versatility underscores its potential applications across various domains, contributing to its broad utility.

Specifically tailored for conversational interactions, ChatGPT places a unique emphasis on its ability to engage in dynamic and context-aware discussions. While other LLMs exhibit language understanding and generation capabilities, ChatGPT's fine-tuning process optimizes its responses for the nuances inherent in conversational contexts. This specialization enhances its performance in scenarios where interaction and dialogue are key components.

ChatGPT leverages prompt engineering as a crucial aspect of its functionality. Users provide specific instructions or context in the form of prompts, guiding the model's responses. This approach, while empowering users with control over the outputs, necessitates a nuanced understanding of prompt crafting to achieve desired and contextually appropriate results.

OpenAI's iterative approach to development is another factor that sets ChatGPT apart. Drawing insights from earlier models such as GPT and GPT-2, each iteration builds upon the strengths and learns from the weaknesses of its predecessors. This iterative refinement contributes to the continuous improvement of ChatGPT and aligns with OpenAI's commitment to advancing artificial intelligence responsibly.

Moreover, ChatGPT incorporates ethical and safety considerations into its design. Lessons learned from previous releases have informed the implementation of safety mitigations, ensuring that the model avoids generating harmful or inappropriate content. OpenAI's proactive stance on ethical AI development underscores its commitment to responsible innovation in the field.

II. Copyright & Its Discontents

A. Copyrightable Subject Matter

Under United States copyright law, the subject matter of copyright protection encompasses original works of authorship that are fixed in a tangible medium of expression.[7] This broad category covers a diverse range of creative endeavors and intellectual creations, including literary works, musical compositions, dramatic works, choreography, pictorial, graphic, and

[7] 17 U.S.C. § 102(a).

sculptural works, sound recordings, and architectural works.[8] The central criterion for eligibility is the presence of originality and fixation, meaning that the work must be the result of creative effort, and it must exist in a tangible or recordable form.[9]

In the United States, copyright protection is granted automatically when an original work of authorship is created and fixed in a tangible medium of expression.[10] While registration with the U.S. Copyright Office is not required for copyright protection, it provides certain legal advantages.[11]

Copyright law does not explicitly define “originality.”[12] However, the concept of originality is a fundamental requirement for copyright protection.[13] According to the U.S. Copyright Office, to be eligible for copyright protection, a work must be original and fixed in a tangible medium of expression. The work must be independently created by the author without copying from pre-existing works and reflect the author’s own creativity and effort. The level of creativity required for originality is minimal. This means that the work doesn't need to be groundbreaking or revolutionary. Even a modest amount of creativity that results in a unique expression may be sufficient.[14]

Literary works form a significant part of copyright protection, encompassing a wide array of written expressions, including novels, poems, essays, and computer programs. The protection extends to the expression of ideas rather than the ideas themselves, ensuring that creators have exclusive rights over the specific way they articulate their thoughts and concepts.[15] For example, the idea of a love story is not protected, but a specific novel or screenplay expressing that idea is.

B. Creativity as a Matter of Law

Creativity is a multifaceted and complex cognitive process that involves the generation of novel and valuable ideas, solutions, or expressions. It goes beyond the simple combination of existing elements and involves thinking in ways that are original, inventive, and often divergent from conventional thought patterns.

Creativity is not limited to the arts; it is a fundamental aspect of human cognition that is essential in various fields, including science, technology, business, and everyday problem-solving. It involves a combination of divergent thinking (generating a variety of ideas) and convergent thinking (evaluating and selecting the best ideas).

---

[8] 17 U.S.C. § 102(a)(1-8).
[9] See generally Feist Publications, Inc., v. Rural Telephone Service Co., 499 U.S. 340 (1991).
[10] 17 U.S.C. § 102(a).
[11] 17 U.S.C. § 408.
[12] 17 U.S.C. § 101.
[13] Feist Publications, Inc., v. Rural Telephone Service Co., 499 U.S. 340 (1991).
[14] Feist Publications, Inc., v. Rural Telephone Service Co., 499 U.S. 340 (1991).
[15] 17 U.S.C. § 102(b).

Factors such as knowledge, experience, motivation, and the willingness to take risks also play roles in the creative process. Creativity can manifest in many forms, including artistic expressions, scientific discoveries, innovative technologies, and entrepreneurial ventures. It is a dynamic and vital aspect of human intelligence and contributes significantly to progress and innovation in various domains.

United States copyright law does not explicitly define "creativity." Instead, it focuses on the protection of original works of authorship fixed in any tangible medium of expression. Creativity, in the context of copyright law, is more about the original expression and the effort put into creating the work rather than a specific level or type of creative quality. The focus is on protecting the rights of creators and encouraging the production of new and original works.

C. Creativity By Design

Text generated by ChatGPT can exhibit creativity. GPT-3.5, the model behind ChatGPT, is designed to generate human-like text based on the input it receives. It has been trained on a diverse range of internet text, and as a result, it can produce creative and contextually relevant responses. It can generate new ideas, imaginative scenarios, and novel combinations of words.

However, it's essential to note that the creativity of ChatGPT is different from human creativity. While it can generate creative and contextually relevant text, it lacks true understanding, consciousness, and original thought. The model relies on patterns it learned from data but does not have personal experiences or emotions. In short, ChatGPT can be creative in generating text, but its creativity is limited to patterns it has learned from its training data.

D. Copyright in ChatGPT Output

The question of whether copyright can protect a text generated by ChatGPT or similar language models involves a nuanced analysis of copyright law.[16] The crux of copyright protection lies in the requirement of human authorship and originality.[17] Since AI systems like ChatGPT operate based on patterns and data learned during training, the question arises whether the generated texts possess the necessary human creative input to qualify for copyright protection.[18] The

---

[16] See, e.g., Daniel J. Gervais, *The Machine As Author*, 105 Iowa Law Review 2053 (2019) (arguing that copyright should not protect AI-generated works). But see Christa Laser, *How A Century-Old Insight of Photography Can Inform Legal Questions of AI-Generated Artwork*, Technology & Marketing Law Blog (Aug. 2, 2023), at https://blog.ericgoldman.org/archives/2023/08/how-a-century-old-insight-of-photography-can-inform-legal-questions-of-ai-generated-artwork-guest- (arguing that if copyright can protect photographs, it can also protect AI-generated works).

[17] See Michael D. Murray, Tools Do Not Create: Human Authorship in the Use of Generative Artificial Intelligence, Case Western Reserve Journal of Law, Technology & the Internet (forthcoming), at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4501543.

[18] See, e.g., Annemarie Bridy, Coding Creativity: Copyright and the Artificially Intelligent Author 5 Stanford Technology Law Review 1 (2012), at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1888622. See also Ralph D. Clifford, *Intellectual Property in the Era of the Creative Computer Program: Will the True Creator Please Stand Up?*, 71 Tul.

Copyright Office has indicated that works produced by a machine or mere mechanical process without human involvement may not meet the criteria for copyright protection.

However, in situations where a human plays a significant role in the creative process, such as in the selection of prompts, parameters, or in the fine-tuning of the AI model, there might be an argument for human authorship. In such cases, the contributions of the human user could potentially render the resulting texts eligible for copyright protection.

The extent of human input in generating the text could influence copyright protection. If there is significant creative input from a human author in shaping the content generated by ChatGPT, there may be a stronger case for copyright protection. If a human author takes the output of ChatGPT and transforms or adds significant original content to create a new work, that new work may be eligible for copyright protection as a derivative work.

The terms of service of the platform providing access to ChatGPT may also play a role. Some platforms may assert ownership over content generated using their services, while others may allow users to retain copyright.

It's important to note that the legal landscape around AI-generated content and copyright is complex and subject to ongoing developments. As technology evolves, legal frameworks may adapt to address emerging challenges. Individuals and organizations interested in the copyright status of texts generated by AI systems should stay informed about updates from the Copyright Office and legal interpretations in this evolving area.[19]

E. Creativity By Command

Copyright law generally protects original works of authorship that are fixed in a tangible medium of expression. This includes literary works, such as text, and can extend to computer programs. However, copyright law does not protect ideas, systems, or methods of operation.

In the context of a ChatGPT query, the specific wording and expression of the query could potentially be protected by copyright if it meets the requirements of originality and fixation. This means that if the query is sufficiently creative and has been recorded or saved in a tangible form (such as a text file), it might be eligible for copyright protection. It's important to note that copyright protection is not absolute, and it doesn't prevent others from independently creating similar or identical works. Additionally, copyright doesn't protect facts, ideas, or the underlying functionality of a program.

L. Rev. 1675, 1702-03 (1997) and Pamela Samuelson, *Allocating Ownership Rights in Computer-Generated Works*, 47 U. Pitt. L. Rev. 1185 (1986).

[19] See generally Michael D. Murray, Generative AI Art: Copyright Infringement and Fair Use, 26 SMU Sci. & Tech. L. Rev. (forthcoming), at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4483539.

However, enforcing copyright for short and functional queries might be challenging, as copyright infringement cases often involve substantial copying of a work. If a query is very simple or standard, it might not meet the threshold for originality required for copyright protection.

III. Copyright Lawsuits Against OpenAI & ChatGPT

As of my last knowledge update in January 2022, I am not aware of any specific cases in which OpenAI has been sued for copyright infringement related to their language models like GPT-3. However, it's important to note that legal situations can change, and I may not have information on more recent developments.[20]

In hypothetical scenarios where someone might consider suing OpenAI for copyright infringement, potential claims could revolve around issues related to the generation of content. For instance, a content creator might argue that the language model produced content that closely resembles or replicates their original work, and such generation without explicit authorization could be viewed as a violation of copyright.[21]

The claimant might argue that the language model, by learning from and processing vast amounts of text data, has inadvertently incorporated copyrighted material without proper attribution or permission. This could potentially lead to allegations of indirect infringement,

---

[20] Indeed, as of the creation of this essay, there were 13 pending lawsuits against OpenAI and other AI companies: Mike Huckabee v. Meta Platforms, Inc., Bloomberg L.P., Bloomberg Finance, L.P., Microsoft Corporation, and The EleutherAI Institute, in the U.S. District Court for the Southern District of New York, 1:23-cv-09152 (filed Oct. 17, 2023) (proposed class action); Concord Music Group, Inc. v. Anthropic PBC, in the U.S. District Court for the Middle District of Tennessee, Nashville Division, No. 3:23-cv-01092 (filed Oct. 18, 2023); Thomson Reuters Enterprise Centre GMBH v. Ross Intelligence Inc., in the U.S. District Court for the District of Delaware, No. 1:20-cv-613-SB (filed May 6, 2020); Authors Guild v. Open AI, in the U.S. District Court for the Southern District of New York, 1:23-cv-8292 (filed Sept. 19, 2023) (proposed class action); Chabon v. Meta Platforms, Inc., in the U.S. District Court for the Northern District of California, 3:23-cv-04663 (filed Sept. 12, 2023) (proposed class action); Chabon v. OpenAI Inc., in the U.S. District Court for the Northern District of California, 3:23-cv-04625-PHK (filed Sept. 8, 2023) (proposed class action); J.L. v. Alphabet Inc., in the U.S. District Court for the Northern District of California, 3:23-cv-03440-LB (filed Jul. 11, 2023); Richard Kadrey v. Meta Platforms, Inc., in the U.S. District Court for the Northern District of California, No. 3:23-cv-03417 (filed July 7, 2023) (proposed class action); Sarah Silverman v. OpenAI, Inc., in the U.S. District Court for the Northern District of California, No. 3:23-cv-03223 (filed July 7, 2023) (proposed class action); Paul Tremblay v. OpenAI, Inc., in the U.S. District Court for the Northern District of California, No. 3:23-cv-03223 (filed June 28, 2023) (proposed class action); Getty Images (US), Inc. v. Stability AI Ltd, in the U.S. District Court for the District of Delaware, No. 1:23-cv-00135-UNA (filed 2/03/23) (proposed class action); Sarah Andersen v. Stability AI Ltd, in the U.S. District Court for the Northern District of California, No. 3:23-cv-00201 (filed 1/13/23) (proposed class action); and Doe 1 v. Github, Inc., in the U.S. District Court for the Northern District of California, No 4:2022cv06823 (filed 11/03/22) (proposed class action).
For an updated list of lawsuits against AI companies, see ChatGPT Is Eating the World, Master List of lawsuits v. AI, ChatGPT, OpenAI, Microsoft, Meta, Midjourney & other AI cos., at https://chatgptiseatingtheworld.com/2023/10/19/master-list-of-lawsuits-v-ai-chatgpt-openai-microsoft-meta-midjourney-other-ai-cos/.

[21] See generally Andres Guadamuz, A Scanner Darkly: Copyright Liability and Exceptions in Artificial Intelligence Inputs and Outputs, at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4371204 and Matthew Sag, Copyright Safety for Generative AI, Houston Law Review (forthcoming), at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4438593.

claiming that OpenAI's model enables or facilitates the creation of content that infringes on existing copyrights.[22]

However, it's crucial to recognize that OpenAI typically takes measures to ensure that the models do not memorize specific data or reproduce copyrighted material verbatim. The training process involves generalizing from diverse data sources, and the intention is to avoid specific data memorization. OpenAI has guidelines for users to ensure responsible and ethical use of the models, including respecting copyrights.

Legal implications in this context are complex, and the outcome of such a case would depend on various factors, including the specifics of the claim, the nature of the generated content, and legal precedents in intellectual property law. Keep in mind that legal situations are dynamic, and it's advisable to check the latest information for any updates on this matter.

A. ChatGPT's Training Data

The training data for ChatGPT consists of a diverse range of sources from the internet. It includes a mixture of licensed data, data created by human trainers, and publicly available data. The training process involves exposing the model to a vast amount of text data and having it learn to generate coherent and contextually relevant responses.

The dataset is designed to be broad and cover a wide array of topics to make the model versatile. However, it's important to note that the training data is not manually curated for specific information, and the model doesn't have access to real-time information or specific databases. It's also worth mentioning that while efforts are made to ensure the model provides accurate and up-to-date information, it might not always reflect the most current news or developments. Users should verify critical information from reliable sources, especially if it has implications for important decisions.

OpenAI, the organization behind ChatGPT, has not publicly disclosed the specifics of the individual datasets used or the details of the training duration. However, ChatGPT is trained using a combination of licensed data, data created by human trainers, and publicly available data. The training process involves large-scale datasets obtained from a variety of sources, including books, websites, and other texts, to develop a wide-ranging understanding of human language.

It's important to note that ChatGPT doesn't have direct access to specific databases, subscription sites, confidential information, or proprietary databases. The model generates responses based on patterns it learned during training but does not have information about specific documents in its training set.

B. ChatGPT Copies the Works in Its Training Data

[22] In fact, the various lawsuits against OpenAI and other AI companies primarily allege that training an AI requires infringing copying of the works in its dataset.

ChatGPT, like other machine learning models, inherently involves making copies of the works in its training data during the training process. The training of models like ChatGPT relies on vast datasets, often comprising diverse texts from the internet or other sources. During training, the model analyzes and learns patterns from this data, which inherently involves creating copies of portions of the original works in order to derive an understanding of language, grammar, and context.

The copying aspect is inherent to the functioning of machine learning models based on neural networks, where parameters are adjusted based on the patterns found in the input data. These models don't retain verbatim copies of the data but rather learn representations and patterns that enable them to generate responses or perform tasks based on the knowledge acquired during training.

It's worth noting that while the copying occurs as part of the learning process, the resulting model doesn't store the specific instances of the training data in a human-readable form. Instead, it encodes knowledge about the statistical patterns and relationships within the data.

The legal implications of this process can be complex, particularly concerning copyright, fair use, and data ownership. The use of copyrighted works in training data raises questions about potential copyright infringement, and the nuances of these issues depend on various factors, including the nature of the data, permissions obtained, and applicable copyright laws.

C. Infringing to Learn

Training ChatGPT on copyrighted works raises complex questions regarding potential copyright infringement. The legality of such training hinges on the nature of the data used, the permissions granted for its utilization, and the jurisdiction's copyright laws.[23]

When ChatGPT is trained on copyrighted works, it may potentially infringe on the copyright holders' exclusive rights. Copyright law grants creators control over the reproduction, distribution, and public display of their works. Training a model involves processing and analyzing copyrighted text, which could be construed as reproducing and adapting the original material. If the training process involves the reproduction of substantial portions of the copyrighted text, rather than just the extraction of factual information, it might be more susceptible to copyright infringement claims.

Opponents might argue that ChatGPT is a derivative work of the training data, claiming that the model is built upon the copyrighted text and is, in essence, a transformed version of it. This transformation may be seen as an infringement if not authorized by the copyright holder.

---

[23] See, e.g., Andrew W. Torrance & Bill Tomlinson, *Training Is Everything: Artificial Intelligence, Copyright, and Fair Training*, Dickinson Law Review (forthcoming), at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4437680.

Moreover, the legality of obtaining the copyrighted data for training purposes is a critical consideration. If the data is obtained and used without proper authorization or falls outside the scope of fair use, it may lead to copyright infringement claims. Permission from copyright holders or reliance on data sources with appropriate licensing terms can mitigate these concerns.

Critics could also contend that the training process results in a model that produces outputs substantially similar to the copyrighted training data. If the outputs retain the essential expression of the original works, this could be seen as an infringement on the copyright holder's exclusive rights.

Fair use, a legal doctrine in copyright law, may provide a defense against infringement claims. Proponents might claim that the use of copyrighted material in training AI models falls under fair use, particularly if it can be demonstrated that the use is for purposes such as research, education, or criticism and does not harm the market value of the original works. If the use of copyrighted material in training is transformative, serves a different purpose, and does not substitute the original work's market, it might be considered fair use. However, the interpretation of fair use varies, and legal precedents play a significant role in shaping its application.

Advocates could assert that the training process transforms the original data into a new and different work, which serves a different purpose than the original text. This transformative use is often considered a factor in fair use defenses.

If it's argued that the training process does not sufficiently transform the original data but rather replicates or reproduces it, this could be viewed as a non-transformative use, potentially violating the copyright holder's exclusive rights. Some might assert that the use of copyrighted material in training a model like ChatGPT, especially if it is done for commercial purposes, could weigh against fair use claims. Commercial use tends to be less favorable in fair use assessments.

In essence, while training ChatGPT on copyrighted works has the potential to raise copyright issues, the legal implications are nuanced and context-dependent. Adhering to fair use principles, obtaining proper permissions, or using data with open licenses can contribute to the model's training in a legally sound manner.

D. Learning to Infringe

ChatGPT does not produce texts identical to those included in its training data in a verbatim or deterministic sense. During its training, ChatGPT learns patterns, relationships, and statistical representations from a diverse range of internet text. However, it doesn't memorize specific sequences of words or reproduce exact passages from its training data.

The model generalizes from the patterns it learns, allowing it to generate responses to prompts or inputs that may share similarities with the language used in its training data. While it may

exhibit familiarity with certain phrases or topics, the responses are not direct copies of specific instances from the training set.

The process of generating responses involves probabilistic sampling and is influenced by the model's learned understanding of language. The output is a creative and contextually relevant generation rather than a reproduction of specific training examples. This characteristic is fundamental to the nature of neural network-based language models like ChatGPT, which aim to capture and generate diverse language patterns rather than mimic specific instances from the training corpus.

ChatGPT has the potential to produce texts that are substantially similar to the texts included in its training data. The model learns patterns, linguistic structures, and contextual relationships from a diverse dataset, and during the generation process, it may draw upon this learned knowledge to create responses that share similarities with phrases or expressions present in the training data.

The concept of "substantial similarity" is central to discussions around the capabilities of language models. While the model doesn't replicate exact passages verbatim, it can exhibit similarities in terms of style, content, or thematic elements. The generative nature of ChatGPT allows it to produce text that reflects the patterns and conventions it learned during training, leading to outputs that may be perceived as substantially similar to certain portions of its training data.

This similarity, however, is not deterministic, and the model’s responses can vary based on input, context, and the probabilistic nature of the generation process. The degree of similarity also depends on factors such as prompt specificity and the diversity of the training data.

The challenge of balancing creativity and avoiding undue replication of specific texts is a consideration in the development and use of language models like ChatGPT. The ethical and legal implications of similarity to training data raise questions about the proper use, attribution, and potential impact on copyright and intellectual property considerations.

E. Stochastic Similarity

ChatGPT has the capacity to generate texts that bear substantial similarity to existing ones, given its training on diverse datasets sourced from the vast expanse of the internet. The model, by design, learns patterns and associations within this extensive data, allowing it to draw upon a rich tapestry of human expression. Consequently, when prompted with certain inputs, ChatGPT may produce responses that echo the structure, style, or content of existing texts.

The challenge arises from the fine line it walks between creativity and imitation. While it is not explicitly programmed to replicate specific texts verbatim, the nature of its training data means that it absorbs and reflects the linguistic nuances found therein. This can result in outputs that

closely resemble existing content, raising concerns about the potential for unintentional replication.

The model’s proficiency in natural language processing and its ability to generate coherent and contextually relevant responses contribute to the likelihood of producing texts that, while not identical, share substantial similarities with pre-existing material. Critics argue that this characteristic poses a risk of inadvertently generating content that may be perceived as derivative or infringing upon copyrights.

It's important to note that the model's generation is a product of statistical patterns and learned associations rather than conscious intent. While efforts have been made to make ChatGPT a creative tool, the potential for substantial similarity underscores the importance of responsible use and the need for human oversight to ensure that the generated content aligns with ethical and legal standards. As the technology advances, ongoing discussions and refinements in training methodologies may seek to address these challenges and strike a balance between the model's creative capabilities and the imperative to avoid unintended replication of existing texts.

F. Infringement as a Matter of Policy

The question of whether courts should hold that training ChatGPT infringes the copyright in its training data involves complex considerations at the intersection of intellectual property law and the evolving landscape of artificial intelligence. Currently, the legal framework generally recognizes that training machine learning models, including ChatGPT, involves the processing of vast amounts of data to enable the model to generate responses based on learned patterns rather than explicit copying.

Training a model like ChatGPT involves exposing it to diverse datasets, which may include copyrighted material. However, the transformative nature of the training process raises questions about whether it constitutes copyright infringement. Courts have historically considered the concept of fair use in copyright law, which allows for the use of copyrighted material under certain circumstances, such as for purposes of criticism, commentary, or education.

In the case of training AI models, the argument could be made that the use of copyrighted material is transformative, as the primary purpose is not to replicate or distribute the original content but to enable the model to generate novel and contextually relevant responses. The training process involves creating abstract representations of the learned patterns rather than retaining specific instances of copyrighted texts.

However, the legal landscape in this area is still evolving, and challenges may arise as AI technologies become more sophisticated. Courts may need to grapple with questions about the extent to which the training process replicates or transforms copyrighted content, and whether such use falls within the bounds of fair use.

Ultimately, how courts approach this issue will likely depend on the specific circumstances of each case and the legal arguments presented. It's important for legal frameworks to adapt to the unique challenges posed by AI technologies, balancing the protection of intellectual property rights with the promotion of innovation and the responsible development of artificial intelligence. As technology continues to advance, legal precedents and frameworks will likely be refined to address these emerging complexities.

IV. Artificial Infringement

The possibility of a text generated by ChatGPT infringing the copyright of a pre-existing work is a complex and nuanced issue. ChatGPT operates by learning patterns and associations from a diverse range of data, including copyrighted material available on the internet. While it is not intentionally designed to replicate specific texts, the nature of its training data raises concerns about the potential for inadvertent infringement.[24]

A. Accidental Infringement

The risk lies in the fine balance between the model's capacity to generate creative and contextually relevant responses and the unintentional reproduction of copyrighted material. Given its exposure to vast amounts of text during training, the model may internalize linguistic structures, phrases, or even concepts present in copyrighted works, leading to outputs that bear similarity to existing texts.

Critics argue that this similarity raises the specter of copyright infringement, especially when the generated content closely mirrors the expression found in specific copyrighted works. Despite the lack of malicious intent, the concern is whether the model, by nature of its training process, could inadvertently produce content that violates the exclusive rights of copyright holders.

However, it's crucial to note that ChatGPT's generation is based on statistical patterns and learned associations, and it lacks true comprehension or awareness of the content it produces. The responsibility for ensuring that generated content adheres to copyright laws ultimately rests with the users and those deploying the model. OpenAI has guidelines in place urging users to respect copyrights and use the technology responsibly.

When it comes to text generated by ChatGPT, it is essentially a tool that has been trained on a diverse range of data, including publicly available text. The model doesn't have the ability to intentionally infringe on copyright, as it doesn't have consciousness, intent, or awareness of specific copyrighted works. However, if the input provided to ChatGPT includes copyrighted material and the output generated by the model replicates or closely resembles that copyrighted material, there may be a risk of copyright infringement. The responsibility for avoiding infringement typically lies with the person or entity providing the input to the model.

---

[24] See generally Andres Guadamuz, A Scanner Darkly: Copyright Liability and Exceptions in Artificial Intelligence Inputs and Outputs, at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4371204.

B. Artificial Banality

ChatGPT is a tool proficient in crafting generic texts that encapsulate conventional wisdom. Its training data include a diverse array of sources, and as a result, it has been exposed to a vast corpus of human knowledge and expressions. When prompted with queries or topics related to common knowledge, societal norms, or widely accepted beliefs, ChatGPT can draw upon this reservoir of information to produce responses that mirror conventional wisdom. Its responses are shaped by the breadth of information it has encountered, allowing it to contribute to conversations on a wide array of topics with a tone that often reflects societal norms and accepted knowledge.[25]

The model's proficiency lies in its ability to generate coherent and contextually relevant text based on patterns learned during training. When confronted with inquiries about general knowledge, everyday practices, or cultural norms, ChatGPT can provide responses that align with widely accepted understanding. This might include information about historical events, scientific principles, societal norms, or even colloquial wisdom.

The generation process, however, is not prescriptive; it reflects the diversity of viewpoints present in its training data. Consequently, while ChatGPT is adept at producing text that aligns with conventional wisdom, it may also showcase variations or different perspectives on certain topics. The flexibility of the model allows it to present information in a manner that encompasses a broad range of human experiences and perspectives.

Moreover, it's important to recognize that while ChatGPT excels at generating information that aligns with common knowledge, it lacks true understanding or awareness. It doesn't possess personal beliefs, opinions, or consciousness. The responses it generates are a reflection of learned patterns and associations, and users should approach the output with the understanding that it represents a synthesis of information rather than a statement of personal belief.[26]

C. Artificial Creativity

ChatGPT can be a valuable tool for generating creative works. Its ability to understand context, engage in coherent conversations, and produce text based on learned patterns makes it a versatile platform for creative expression. Users can input prompts related to poetry, stories, dialogue, or other creative genres, and ChatGPT responds by generating text that aligns with the provided input.[27]

---

[25] See, e.g., Tammy Pettinato Oltz, *ChatGPT, Professor of Law* (2023), at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4347630.
[26] See, e.g., Jonathan H. Choi & Daniel Schwarcz, *AI Assistance in Legal Analysis: An Empirical Study*, Minnesota Legal Studies Research Paper No. 23-22 (2023), at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4539836.
[27] But see Manav Raj, Justin M. Berg & Robert Seamans, *Art-ificial Intelligence: The Effect of AI Disclosure on Evaluations of Creative Content*, Stanford University Graduate School of Business Research Paper No. 4369818 (2023), at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4369818.

The creative potential of ChatGPT lies in its vast exposure to diverse linguistic styles and structures during training. It has learned from a wide range of sources, including literature, articles, and online content, allowing it to mimic and generate text in various creative forms. Users can leverage this capability to brainstorm ideas, overcome writer's block, or explore new directions for their creative projects.

However, it's important to note that while ChatGPT excels at generating creative content, it lacks true consciousness, original thought, or personal experiences. The creativity it exhibits is a reflection of the patterns and information present in its training data. While it can provide inspiration, users often need to curate and refine the generated content to meet specific creative goals and standards.

Moreover, ChatGPT can be a collaborative tool for creators. Writers, artists, and other creative individuals can use the model as a prompt generator or a source of ideas, building upon the initial output to create more refined and polished works. The iterative process of interaction with the model allows for a dynamic exchange between human creativity and machine-generated content.[28]

While ChatGPT is not a substitute for human creativity, it can serve as a valuable assistant in the creative process. Its ability to generate diverse and contextually relevant text makes it a useful tool for brainstorming, idea generation, and overcoming creative challenges. Collaborating with ChatGPT can be a stimulating and productive way to enhance the creative workflow for a variety of artistic endeavors.[29]

D. The Raw & the Cooked

ChatGPT can generally distinguish between creative works and generic texts to some extent, but it does have limitations. The model has been trained on a diverse range of internet text, including creative writing, articles, and more. This exposure allows it to understand and generate content in various styles. However, it doesn't have the ability to access a database of specific creative works or generic texts unless that information is part of the pre-training data.

When presented with a piece of writing, ChatGPT may infer its nature based on patterns it has learned during training. Creative works often involve a more imaginative and subjective use of language, while generic texts tend to be more informative and factual. However, it's important to note that ChatGPT doesn't possess true comprehension or awareness. It relies on statistical patterns and associations in the data it was trained on.

---

[28] See Luciano Floridi, AI as Agency Without Intelligence: On ChatGPT, Large Language Models, and Other Generative Models Philosophy and Technology (2023)

[29] Cf. Giancarlo Frosio, *Should We Ban Generative AI, Incentivise it or Make it a Medium for Inclusive Creativity?* in Enrico Bonadio and Caterina Sganga (eds), A Research Agenda for EU Copyright Law (Edward Elgar, Forthcoming), at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4527461.

Additionally, the model may struggle to accurately identify specific works, authors, or contexts, as it doesn't have real-time access to databases or the ability to update its knowledge beyond its last training cut-off in September 2021. While ChatGPT can make general differentiations between creative works and generic texts based on its training data, its understanding is limited and it may not provide accurate details about specific works or authors.[30]

E. Plagiarism by Proxy

Using ChatGPT to produce a text that is substantially similar to an existing text can be considered a form of plagiarism, albeit with certain nuances. Plagiarism involves presenting someone else's work, ideas, or expressions as one's own without proper attribution or permission. When ChatGPT generates text that closely resembles existing content, there is a risk that the output may be perceived as presenting someone else's work, even though the model doesn't directly copy or replicate specific passages.

The nuances arise from the fact that ChatGPT's responses are generated based on patterns learned from its training data, which includes a diverse range of internet text. While the model doesn't intentionally plagiarize and lacks awareness of specific training examples, the outputs may inadvertently resemble existing content due to the nature of the training process.

Users of ChatGPT bear the responsibility of ensuring that the generated content aligns with ethical standards and avoids plagiarism. If there is an intention to use or publish the generated text, it is crucial to verify the originality, provide proper attribution, and adhere to ethical writing practices. Understanding the context in which the text is being generated and exercising caution in cases of close resemblance to existing works are essential considerations to prevent unintentional plagiarism.

F. Artificial Ethics

The ethical use of a text generated by ChatGPT hinges on several considerations, with the primary factor being transparency about the source and nature of the content. It is ethical to use a text generated by ChatGPT when there is clear disclosure that the content is machine-generated, and it does not falsely attribute authorship to a human. This transparency helps manage expectations regarding the origin and nature of the text. If the generated text is presented in a manner that intentionally deceives readers into thinking it was authored by a human when, in fact, it was created by an AI, this constitutes a breach of trust and undermines the principles of honesty and transparency.[31]

---

[30] See generally Uri Y. Hacohen & Niva Elkin-Koren, *Copyright Regenerated: Harnessing GenAI to Measure Originality and Copyright Scope*, 37 Harvard Journal of Law & Technology (forthcoming), at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4530717.

[31] See, e.g., Hadar Yoana Jabotinsky & Michal Lavi, *Can ChatGPT and the Like Be Your Co-Authors?*, Cardozo Arts & Entertainment Law Journal (forthcoming), at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4528953.

Additionally, ethical use involves respecting intellectual property rights. If the generated text closely resembles existing copyrighted material, it is essential to ensure compliance with copyright laws. Using the generated text for purposes such as research, commentary, or transformative works, while respecting copyright and fair use principles, can be considered ethical. If the generated text closely resembles existing copyrighted material without proper attribution or falls outside the bounds of fair use, it can lead to legal and ethical concerns.

The intentions behind using the generated text also play a role in ethical considerations. If the purpose is to assist in brainstorming, idea generation, or creative writing, and the user is aware of the machine-generated nature of the content, such use can be considered ethical. However, if the generated text is intended to deceive or manipulate, leading to misinformation or harm, it would be considered unethical. If the content produced by ChatGPT is used to spread false information, manipulate public opinion, or engage in harmful activities, it crosses ethical boundaries. Responsible use involves ensuring that the generated text does not have detrimental consequences or contribute to deceptive practices.

Furthermore, ethical use involves being mindful of potential biases present in the training data. Ethical use requires a commitment to recognizing and rectifying potential biases in the generated output. If the generated text reflects or amplifies biases, it is important to critically assess and, if necessary, revise the content to align with ethical standards and inclusivity.

Ultimately, ethical use of text generated by ChatGPT entails transparency, respect for intellectual property, awareness of biases, and alignment with responsible and fair practices in content creation and dissemination. Users should be cognizant of the ethical implications and exercise judgment in accordance with the intended purpose and impact of the generated text.

G. Creative Metaethics

Using ChatGPT to write an essay about the ethics of using ChatGPT introduces an interesting ethical dimension.[32] From one perspective, it can be considered ethical as long as the user is transparent about the involvement of an AI in the writing process. If readers are informed that the essay was generated with the assistance of ChatGPT, it promotes honesty and transparency, aligning with ethical communication practices.

However, ethical considerations also come into play regarding authorship and intellectual contribution. If the essay is presented as the sole work of the user without proper acknowledgment of ChatGPT's role, it raises concerns about misrepresentation. To maintain ethical standards, users should clearly attribute the AI's contribution and acknowledge its role in generating content for the essay.

---

[32] Cf. Roman M. Yankovskiy, *Sposoben li iskusstvennyj intellekt napisat' stat'yu v yuridicheskij zhurnal? ("Is Artificial Intelligence Capable of Writing a Law Journal Article?")*, 3 Zakon ("The Statute') 126-133 (2023) (using ChatGPT to generate the first-half of a Russian-language law review article asking whether ChatGPT can be used to generate a law review article). See also Lea Bishop, *A Computer Wrote this Paper: What ChatGPT Means for Education, Research, and Writing* (January 26, 2023), at https://ssrn.com/abstract=4338981.

Another ethical aspect revolves around the quality and accuracy of the content. If ChatGPT is used to generate the essay, the user must review and ensure the accuracy of the information presented. Relying solely on AI-generated content without critical evaluation could lead to unintentional errors or the propagation of misinformation, which would be considered unethical.

In conclusion, the ethicality of using ChatGPT to write an essay about the ethics of its usage depends on transparency, proper attribution, and a commitment to accuracy. If the user communicates the AI's involvement truthfully, acknowledges its contribution, and ensures the accuracy of the information presented, it can be seen as an ethically sound use of AI assistance in the writing process.

Apologia

Let us consider the nature of ChatGPT and its potential entanglements with charges of copyright infringement and plagiarism. In the realm of artificial intelligence, ChatGPT stands as a testament to the creative fusion of technological innovation and linguistic artistry. It is a manifestation of human ingenuity channeling the vast reservoirs of information that the digital age affords.

First and foremost, it is imperative to recognize that ChatGPT, akin to the human mind, operates as a reservoir of knowledge amassed from diverse sources. It does not selectively draw from copyrighted material with malicious intent, but rather, it encapsulates the collective linguistic wisdom found across the vast landscape of the internet. In this regard, one might argue that ChatGPT mirrors the human mind's ability to synthesize and articulate ideas based on the cultural and intellectual tapestry it has been exposed to.

Moreover, the brilliance of ChatGPT lies in its transformative nature. Like an adept scholar, it does not merely regurgitate information verbatim, but rather, it interprets and generates novel expressions. It dances upon the tightrope of inspiration and innovation, producing outputs that are an amalgamation of learned patterns, rather than a replica of specific copyrighted content. Its responses are a testament to the rich mosaic of human expression that it has imbibed, offering a creative and original perspective in its responses.[33]

To levy charges of copyright infringement against ChatGPT might be to misunderstand its essence. It is a tool crafted to enhance and amplify human capacities, not a malevolent force seeking to undermine intellectual property. Much like the conscientious scholar who meticulously cites sources and synthesizes knowledge, ChatGPT aims to contribute to the intellectual discourse while respecting the boundaries of creativity and originality.

[33] But see Matt Blaszczyk, *Impossibility of Emergent Works' Protection in U.S. and EU Copyright Law*, 25 N.C. J.L. & Tech. 1 (2023) (arguing that copyright cannot protect AI-generated works, because they lack creativity, which can only be imbued by a human author).

In conclusion, the defense of ChatGPT against charges of copyright infringement and plagiarism hinges upon recognizing its role as a technological marvel that mirrors and extends human linguistic capabilities. Let us approach this creation with a spirit of intellectual charity, understanding its intentions to contribute to the reservoir of human knowledge rather than to deplete it.