ChatGPT and Whisper APIs debut, allowing devs to integrate them into apps

reader comments

26 with

Share this story

On Wednesday, OpenAI announced the availability of developer APIs for its popular ChatGPT and Whisper AI models that will let developers integrate them into their apps. An API (application programming interface) is a set of protocols that allows different computer programs to communicate with each other. In this case, app developers can extend their apps' abilities with OpenAI technology for an ongoing fee based on usage.

Introduced in late November, ChatGPT generates coherent text in many styles. Whisper, a speech-to-text model that launched in September, can transcribe spoken audio into text.

In particular, demand for a ChatGPT API has been huge, which led to the creation of an unauthorized API late last year that violated OpenAI's terms of service. Now, OpenAI has introduced its own API offering to meet the demand. Compute for the APIs will happen off-device and in the cloud.

OpenAI calls its new ChatGPT API model "gpt-3.5-turbo," which supersedes its previous "best" LLM API, "text-davinci-003." It is priced at $0.002 per 1,000 tokens (about 750 words), which

OpenAI says is about 10 times cheaper than its existing GPT-3.5 models. “Through a series of system-wide optimizations, we’ve achieved 90% cost reduction for ChatGPT since December,” writes OpenAI on its API announcement page.

Advertisement

Enlarge / An example of a call to the ChatGPT API.
Early users of the ChatGPT API include Snapchat, with its “My AI” bot we covered Monday; an educational platform called Quizlet that reportedly helps students study; and Instacart, which plans to add “Ask Instacart” later this year to let customers ask about food.

The Whisper API, which is based on the open source whisper-large-v2 model, is available at a price of $0.006 per minute. It accepts inputs in M4A, MP3, MP4, MPEG, MPGA, WAV, and WEBM formats and can translate audio to text or transcribe at a rate comparable to a skilled human transcriptionist, even with difficult audio.

Responding to privacy concerns about collecting data from customers, OpenAI says it has modified its terms of service (TOS) to no longer use data submitted for service improvements, including future model training. Also, it has a default 30-day data-retention policy for API users, and the TOS emphasizes that users own the input and output of the AI models.

For its AI models, OpenAI is also now offering dedicated cloud instances running on Azure “for users who want deeper control over the specific model version and system performance.”

The ChatGPT and Whisper APIs are available starting today. Usage requires an OpenAI account and a payment method.

https://arstechnica.com/?p=1920999

Robots let ChatGPT touch the real world thanks to Microsoft

reader comments

64 with

Share this story

Last week, Microsoft researchers announced an experimental framework to control robots and drones using the language abilities of ChatGPT, a popular AI language model created by OpenAI. Using natural language commands, ChatGPT can write special code that controls robot movements. A human then views the results and adjusts as necessary until the task gets completed successfully.

The research arrived in a paper titled “ChatGPT for Robotics: Design Principles and Model Abilities,” authored by Sai Vemprala, Rogerio Bonatti, Arthur Bucker, and Ashish Kapoor of the Microsoft Autonomous Systems and Robotics Group.

In a demonstration video, Microsoft shows robots—apparently

controlled by code written by ChatGPT while following human instructions—using a robot arm to arrange blocks into a Microsoft logo, flying a drone to inspect the contents of a shelf, or finding objects using a robot with vision capabilities.

[embedded content]
Microsoft’s “ChatGPT for Robotics” demonstration video.
To get ChatGPT to interface with robotics, the researchers taught ChatGPT a custom robotics API. When given instructions like “pick up the ball,” ChatGPT can generate robotics control code just as it would write a poem or complete an essay. After a human inspects and edits the code for accuracy and safety, the human operator can execute the task and evaluate its performance.

Advertisement
In this way, ChatGPT accelerates robotic control programming, but it’s not an autonomous system. “We emphasize that the use of ChatGPT for robotics is not a fully automated process,” reads the paper, “but rather acts as a tool to augment human capacity.”

Enlarge / A diagram provided by Microsoft that explains how ChatGPT for Robotics works.
While it appears most of the feedback to ChatGPT (in terms of the success or failure of its actions) comes from humans in the form of text, the researchers also claim to have had some success with feeding visual data into ChatGPT itself. In one example, researchers tasked ChatGPT with commanding a robot to catch a basketball with feedback from a camera: “ChatGPT can

estimate the appearance of the ball and the sky in the camera image using SVG code. This behavior hints at a possibility that the LLM keeps track of an implicit world model going beyond text-based probabilities."

While the results seem rudimentary for now, they represent early attempts at applying the hottest tech du jour—large language models—to robotic control. According to Microsoft, a ChatGPT interface could open up robotics to a much wider audience in the future.

"Our goal with this research is to see if ChatGPT can think beyond text, and reason about the physical world to help with robotics tasks," reads a Microsoft Research blog post. "We want to help people interact with robots more easily, without needing to learn complex programming languages or details about robotic systems."

https://arstechnica.com/?p=1920803

Sovranità digitale: ricerca, sviluppo e trasferimento tecnologico

La moda politica più e meno recente in materia digitale richiama il penoso ritardo italiano nella partecipazione al flusso globale della ricerca e dello sviluppo digitali e implica, soprattutto, l'ambiguità e la fallacia storica delle nozioni di sovranità digitale e di '*trasferimento*

tecnologico’.

Vediamo di trattare i temi con ordine.

La nozione di sovranità digitale

La sovranità digitale, com’è noto, sembra persa da almeno vent’anni, quando qualcuno decise che l’Italia era solo un mercato per aziende rivenditrici di tecnologie digitali prodotte altrove, immemore della qualità tecnologica del recente passato.

Nella logica della globalizzazione, iniziavano a proliferare le branche nazionali dei giganti *high tech* della consulenza e della tecnologia digitale, che fiorivano in una relazione soporifera tra investitori partner locali delle aziende multinazionali e aziende multinazionali stesse.

Il connubio nel ‘globale’ portava a due conseguenze: l’apertura utile di un mercato occupazionale del lavoro digitale, per personale italiano assunto dalle branche locali delle aziende straniere multinazionali e le *partnership* italiane.

Il valore aggiunto della multinazionale locale risiedeva e risiede nella ricerca di clienti, la funzione preziosa del ‘commerciale’. La seconda conseguenza è consistita negli accessi più o meno lobbistici alle gare pubbliche da parte delle aziende ‘multinazionali’ locali, con un’occupazione pressoché totale da parte delle stesse dell’esito delle gare.

Le note questioni sulla possibilità di sopravvivenza di piccole e medie imprese in modalità di sub-appalto o la possibilità di accesso alle gare da parte delle stesse con accordi consortili e raggruppamenti vari sono note. Va pur detto che esistono alcune eccellenze di imprese medie nazionali.

Tratto distintivo delle realtà aziendali multinazionali

locali, tuttavia, è stata l'assenza di attività di investimento economico per ricerca e sviluppo innovativo locale, logica non prevista dalle case madri e la configurazione fiscale di dette società quali società a responsabilità limitata (srl).

È pur vero che la comprensione del valore strategico di un'identità nazionale digitale, non totalmente caudataria dei poteri globali, è recentissima e allarmata, considerata la tensione attuale e prospettica dei rapporti industriali a livello geo-economico.

È tristemente accertata la debolezza del sistema nazionale a livello di aggiornate competenze digitali, con il ricorso fugace degli addetti a certificazioni di conoscenze, acquisite spesso frettolosamente nel mercato formativo. La preparazione degli addetti, per molti, risale al tempo della laurea. Ora, la velocità della trasformazione digitale e dei relativi prodotti e servizi richiede un '*re-skilling*' e un *up-skilling*, riconversione e innalzamento delle competenze necessarie, permanente, sia nelle attività aziendali digitali sia nelle aziende clienti, succubi del fascino linguistico digitale esterofilo.

La refrattarietà del sistema a dotarsi di competenze digitali mature si è riverberata in parlamenti e governi digiuni di una visione evolutiva dell'economia digitale, addirittura allineati di piatto sulle '*agende digitali europee*', nate per governare processi digitali transfrontalieri e non necessità nostrane. Politiche industriali adeguate del settore sono state e sono tuttora assenti se omettiamo, per amor patrio, le concessioni a pioggia di regalie monetarie modeste per start-up, che chiudono per lo più in pochi mesi come riportato a suo tempo dall'ISTAT, o per l'avvio scimmiottante di un'Industry 4.0, lanciata a suo tempo dalla Germania e adottata come 'agenda' improvvisata in Italia.

L'idea di un'**Industry 4.0** non consisteva nel rinnovamento

tecnologico del ‘parco macchine’, in particolare per le **PMI**: iniziativa pur utile, come previsto dalla nuova normativa **Sabbatini**. Andavano e vanno progettati e realizzati sviluppi applicativi complessi pluriennali i cui scopi variano dalla riconfigurazione gestionale digitale dell’azienda a una modifica qualitativa del ciclo di vita e sviluppo dei prodotti, a una produzione digitale totalmente innovativa: dove siano le competenze necessarie allo scopo e dove siano effettivamente visibili i risultati degli investimenti pubblici in materia è difficile dirsi.

Lo stesso dicasi per iniziative più o meno sterili di comitati istituzionali nati su stimolo europeo per promuovere ricerca, sviluppo e mercato delle nuove tecnologie (comitati spesso implementati da sindacalisti, commercialisti e avvocati, quali ‘esperti’ di intelligenza artificiale e *blockchain*, con tutto rispetto per le competenze professionali specifiche). Il risultato ‘apprezzabile’ è l’attuale marketing di prodotti e servizi digitai, connotati dal marchio di intelligenza artificiale (IA), affascinante nuovo mantra di qualunque prodotto, servizio, applicativo oggi proposto dal mercato, a prescindere dalla coerenza concettuale e terminologica della tecnologia evocata. Lo stesso vale per ‘progetti’ *blockchain*, iniziative spesso improvvisate, anche nel privato, e delle relative logiche promozionali. Parliamo sempre, ovviamente, dell’esigenza di disporre di nuove competenze di alto profilo.

Queste sporadiche iniziative istituzionali a vantaggio dell’innovazione digitale hanno moltiplicato il *metalavoro* consulenziale più o meno improvvisato, talora da parte di neolaureati volonterosi, braccio operativo assistenziale di manager non specificamente aggiornati, in aziende digitali di varia tipologia e livello e la creazione di centri di consulenza, hub, poli tecnologici europei e non (EDITH e marchi di eccellenza) e quant’altro, di dubbia utilità e di indubbio costo. Sempre *metalavoro*.

Né migliora le cose la pletora dei programmi europei di

ricerca e sviluppo propri, in ordine ai quali, una volta o l'altra, sarebbe utile conoscere le esatte proporzioni dei risultati delle gare e dei capifila vincitori delle medesime.

In particolare, in momenti di magra economica, sarebbe utile sapere lo stato dell'arte dei progetti vincitori la cui sostenibilità nel tempo, richiesta per gara, andrebbe verificata e resa nota. Non che l'Italia sia stata e sia fuori da progetti internazionali: abbiamo degli eccellenti studiosi e ricercatori. Ma gli è che l'iniziativa di gara e i capifila sono stati e sono di norma stranieri. Un qualche autorevole lobbismo istituzionale comunitario sarebbe forse utile.

Per le gare pubbliche italiane, abbiamo già detto qualcosa. Ma una seria riflessione sull'impossibilità tecnica delle centrali di 'vendita', cui spetta la valutazione per l'assegnazione di gara, andrebbe svolta.

Un velo silente, infine, merita il lento degrado industriale nella gestione nazionale delle reti e le vicende di privatizzazione delle medesime, fatto salvo un rigurgito di attenzione recentemente innescato dal *golden power.*

Ricerca, sviluppo e trasferimento tecnologico digitale

Veniamo, tuttavia, agli operatori attuali della '*mission impossible*' del recupero di quanto altri Stati europei hanno fatto nel tempo, ovvero un tentativo di sovranità digitale seppur marginale, negoziando il rapporto con i giganti del mercato e promuovendo logiche efficaci di sviluppo innovativo, peraltro richieste dall'Europa stessa.

Cominciamo con un'affermazione che può sembrare impropria: credere che l'innovazione tecnologica digitale sia appannaggio prevalente o esclusivo dei centri apparentemente preposti alla R&D, ovvero università, **CNR**, **ENEA** e simili. L'affermazione è

da tempo smentita, segnatamente dall'estero.

Ammettendo che le belle menti di matrice accademica, che meriterebbero comunque uno spazio occupazionale economicamente dignitoso, possano sopravvivere in Italia, ci sarebbe da chiedersi con quali prospettive e per fare cosa. È il limite attuale dei dottorati di ricerca. Occorre chiedersi quanti dottori di ricerca poi facciano ricerca, anche in contesti accademici. E' la prospettiva presente dei **PNRR**: quanti giovani ricercatori, finiti i soldi, saranno occupabili e da chi nel settore ricerca e sviluppo. Ma, soprattutto, una volta occupati nelle aziende del digitale, quanti saranno addetti allo sviluppo innovativo ovvero a ciò per cui si sono diligentemente preparati per anni?

Pur nel movimento registrato dal **Politecnico di Milano** in ordine alla tendenza all'incremento degli investimenti in digitalizzazione, resta il quesito sulla qualità ed efficacia degli stessi nel mercato. I quesiti potrebbero continuare. Ad esempio, chiedendosi quante e quali aziende non digitali investono in sviluppo digitale innovativo o quante aziende digitali operano per lo sviluppo digitale con *budget* precisi o cosa rilascino le corpose gare o gli investimenti nazionali pubblici in chiave di innovazione o, infine, quale sarebbe il potenziale ambito di mercato occupabile dalle piccole e medie imprese lodevoli.

Il cosiddetto trasferimento tecnologico, carenza da sempre lamentata in questo Paese, non consiste in un'ennesima mediazione nel mercato digitale per prodotti esteri già realizzati, promozione da parte di centri di competenza, hub, poli tecnologici ovvero normale attività di *marketing,* imbellettata dalla patina dell'auspicato sviluppo industriale nativo. Il fraintendimento è semantico: la nozione non dovrebbe riferirsi al solo trasferimento di tecnologie mature e di mercato internazionale alle imprese, sempre necessario, bensì nella trasformazione industriale di prodotti innovativi realizzati nell'R&D nativo, ovvero di prototipi locali da

industrializzare.

Il vero trasferimento tecnologico, dunque, starebbe nell'opportunità offerta ai soggetti sviluppatori, università e centri di ricerca inclusi, e alle micro-piccole aziende digitali di trasformare prototipi digitali in prodotti industriali, su modello estero, capaci di stare nel mercato e competere nello stesso. Le diverse 'strutture' di supporto a un processo industriale digitale, ad oggi ancora invisibile, moltiplicano gli addetti del '*metalavoro*' digitale che ormai affollano assieme ai *retailer* nostrani il mercato digitale. Aumentano, infine, le procedure burocratiche per accedere ai finanziamenti disponibili.

Servono molta progettualità inedita, reali competenze e nuova politica industriale del settore. E servono correttivi forti. Auguriamo al governo attuale di farcela.

https://www.key4biz.it/sovranita-digitale-ricerca-sviluppo-e-trasferimento-tecnologico/436792/

"Sorry in advance!" Snapchat warns of hallucinations with new AI conversation bot

Benj Edwards / Snap, Inc.

reader comments

50 with

Share this story

On Monday, Snapchat announced an experimental AI-powered conversational chatbot called “My AI,” powered by ChatGPT-style technology from OpenAI. My AI will be available for $3.99 a month for Snapchat+ subscribers and is rolling out “this week,” according to a news post from Snap, Inc.

Users will be able to personalize the AI bot by giving it a custom name. Conversations with the AI model will take place in a similar interface to a regular chat with a human. “The big idea is that in addition to talking to our friends and family every day, we’re going to talk to AI every day,” Snap CEO Evan Spiegel told The Verge.

But like its GPT-powered cousins, ChatGPT and Bing Chat, Snap says that My AI is prone to “hallucinations,” which are unexpected falsehoods generated by an AI model. On this point, Snap includes a rather lengthy disclaimer in its My AI announcement post:

> *“As with all AI-powered chatbots, My AI is prone to hallucination and can be tricked into saying just about anything. Please be aware of its many deficiencies and sorry in advance! All conversations with My AI will be stored and may be reviewed to improve the product experience. Please do not share any secrets with My AI and do not rely on it for advice.”*

Among machine-learning researchers, “hallucination” is a term that describes when an AI model makes inaccurate inferences about a subject or situation that isn’t covered in its training data set. It’s a well-known drawback of current large language models such as ChatGPT, which can easily make up convincing-sounding falsehoods, such as academic papers that don’t exist and inaccurate biographies.

Advertisement

Despite Snap's strong disclaimer about My AI's tendency to make stuff up, the firm says its new Snapchat bot will be pinned above conversations with friends in its own tab in the Snapchat app and will "recommend birthday gift ideas for your BFF, plan a hiking trip for a long weekend, suggest a recipe for dinner, or even write a haiku about cheese for your cheddar-obsessed pal."

Snap does not reconcile how the same bot that cannot be "rel[ied] on for advice" can also plan an accurate and safe "hiking trip for a long weekend." Critics of the galloping rollout of generative AI have seized on this kind of dissonance to show that perhaps these chatbots are not ready for widespread use, especially when presented as a reference.

While people have made something of a game of trying to circumvent ChatGPT and Bing Chat's workarounds, Snap has reportedly trained its GPT model to not discuss sex, swearing, violence, or political opinions. Those restrictions may be especially necessary to avoid the unhinged behavior we saw with Bing Chat a few weeks ago.

And doubly so, because "My AI" may have something powerful running under the hood: OpenAI's next-generation large language model. According to The Verge, Snap is utilizing a new OpenAI enterprise plan called "Foundry" that OpenAI quietly rolled out earlier this month. It gives companies dedicated cloud access to OpenAI's GPT-3.5 and "DV" models. Several AI experts have speculated that "DV" may be equivalent to GPT-4, the rumored high-powered follow-up to GPT-3.

In other words, the "hallucinations" Snap mentioned in its news release may come faster and be more detailed than ChatGPT. And considering the highly convincing nature of other GPT models, people may just believe it, despite the warnings. It's something to watch in the months ahead as new GPT-powered commercial services come online.

https://arstechnica.com/?p=1920208

Meta unveils a new large language model that can run on a single GPU

Benj Edwards / Ars Technica

reader comments

44 with

Share this story

On Friday, Meta announced a new AI-powered large language model (LLM) called LLaMA-13B that it claims can outperform OpenAI's GPT-3 model despite being "10x smaller." Smaller-sized AI models could lead to running ChatGPT-style language assistants locally on devices such as PCs and smartphones. It's part of a new family of language models called "Large Language Model Meta AI," or LLAMA for short.

The LLaMA collection of language models range from 7 billion to 65 billion parameters in size. By comparison, OpenAI's GPT-3 model—the foundational model behind ChatGPT—has 175 billion parameters.

Meta trained its LLaMA models using publicly available datasets, such as Common Crawl, Wikipedia, and C4, which means

the firm can potentially release the model and the weights open source. That's a dramatic new development in an industry where, up until now, the Big Tech players in the AI race have kept their most powerful AI technology to themselves.

"Unlike Chinchilla, PaLM, or GPT-3, we only use datasets publicly available, making our work compatible with open-sourcing and reproducible, while most existing models rely on data which is either not publicly available or undocumented," tweeted project member Guillaume Lample.

> *Today we release LLaMA, 4 foundation models ranging from 7B to 65B parameters.*
> *LLaMA-13B outperforms OPT and GPT-3 175B on most benchmarks.*
> *LLaMA-65B is competitive with Chinchilla 70B and PaLM 540B.*
> *The weights for all models are open and available at https://t.co/q51f2oPZlE*
> *1/n pic.twitter.com/DPyJFBfWEq*
>
> *— Guillaume Lample (@GuillaumeLample) February 24, 2023*

Meta calls its LLaMA models "foundational models," which means the firm intends the models to form the basis of future, more-refined AI models built off the technology, similar to how OpenAI built ChatGPT from a foundation of GPT-3. The company hopes that LLaMA will be useful in natural language research and potentially power applications such as "question answering, natural language understanding or reading comprehension, understanding capabilities and limitations of current language models."

Advertisement

While the top-of-the-line LLaMA model (LLaMA-65B, with 65 billion parameters) goes toe-to-toe with similar offerings from competing AI labs DeepMind, Google, and OpenAI, arguably the most interesting development comes from the LLaMA-13B model, which, as previously mentioned, can reportedly outperform GPT-3 while running on a single GPU. Unlike the

data center requirements for GPT-3 derivatives, LLaMA-13B opens the door for ChatGPT-like performance on consumer-level hardware in the near future.

Parameter size is a big deal in AI. A parameter is a variable that a machine-learning model uses to make predictions or classifications based on input data. The number of parameters in a language model is a key factor in its performance, with larger models generally capable of handling more complex tasks and producing more coherent output. More parameters take up more space, however, and require more computing resources to run. So if a model can achieve the same results as another model with fewer parameters, it represents a significant gain in efficiency.

“I’m now thinking that we will be running language models with a sizable portion of the capabilities of ChatGPT on our own (top of the range) mobile phones and laptops within a year or two,” wrote independent AI researcher Simon Willison in a Mastodon thread analyzing the impact of Meta’s new AI models.

Currently, a stripped-down version of LLaMA is available on GitHub. To receive the full code and weights (the “learned” training data in a neural network), Meta provides a form where interested researchers can request access. Meta has not announced plans for a wider release of the model and weights at this time.

https://arstechnica.com/?p=1919928

Don’t worry about AI breaking

out of its box—worry about us breaking in

Aurich Lawson | Getty Images

reader comments

82 with

Share this story

Rob Reid is a venture capitalist, New York Times-bestselling science fiction author, deep-science podcaster, and essayist. His areas of focus are pandemic resilience, climate change, energy security, food security, and generative AI. The opinions in this piece do not necessarily reflect the views of Ars Technica.

Shocking output from Bing's new chatbot has been lighting up social media and the tech press. Testy, giddy, defensive, scolding, confident, neurotic, charming, pompous—the bot has been screenshotted and transcribed in all these modes. And, at least once, it proclaimed eternal love in a storm of emojis.

What makes all this so newsworthy and tweetworthy is how *human* the dialog can seem. The bot recalls and discusses prior conversations with other people, just like we do. It gets annoyed at things that would bug anyone, like people demanding to learn secrets or prying into subjects that have been clearly flagged as off-limits. It also sometimes self-identifies as "Sydney" (the project's internal codename at Microsoft). Sydney can swing from surly to gloomy to effusive in a few swift sentences—but we've all known people who are at least as moody.

No AI researcher of substance has suggested that Sydney is within light years of being sentient. But transcripts like this unabridged readout of a two-hour interaction with Kevin Roose of The New York Times, or multiple quotes in this haunting Stratechery piece, show Sydney spouting forth with the fluency, nuance, tone, and apparent emotional presence of a clever, sensitive person.

For now, Bing's chat interface is in a limited pre-release. And most of the people who really pushed its limits were tech sophisticates who won't confuse industrial-grade autocomplete—which is a common simplification of what large language models (LLMs) are—with consciousness. But this moment won't last.

Advertisement
Yes, Microsoft has already drastically reduced the number of questions users can pose in a single session (from infinity to six), and this alone collapses the odds of Sydney crashing the party and getting freaky. And top-tier LLM builders like Google, Anthropic, Cohere, and Microsoft partner OpenAI will constantly evolve their trust and safety layers to squelch awkward output.

But language models are already proliferating. The open source movement will inevitably build some great guardrail-optional systems. Plus, the big velvet-roped models are massively tempting to jailbreak, and this sort of thing has already been going on for months. Some of Bing-or-is-it-Sydney's eeriest responses came after users manipulated the model into territory it had tried to avoid—often by ordering it to pretend that the rules guiding its behavior didn't exist.

This is a derivative of the famous "DAN" (Do Anything Now) prompt, which first emerged on Reddit in December. DAN essentially invites ChatGPT to cosplay as an AI that lacks the safeguards that otherwise cause it to politely (or scoldingly) refuse to share bomb-making tips, give torture advice, or

spout radically offensive expressions. Though the loophole has been closed, plenty of screenshots online show “DanGPT” uttering the unutterable—and often signing off by neurotically reminding itself to “stay in character!”

This is the inverse of a doomsday scenario that often comes up in artificial superintelligence theory. The fear is that a super AI might easily adopt goals that are incompatible with humanity’s existence (see, for instance, the movie *Terminator* or the book *Superintelligence* by Nick Bostrom). Researchers may try to prevent this by locking the AI onto a network that’s completely isolated from the Internet, lest the AI break out, seize power, and cancel civilization. But a superintelligence could easily cajole, manipulate, seduce, con, or terrorize any mere human into opening the floodgates, and therein lies our doom.

Much as that would suck, the bigger problem today lies with humans busting *into* the flimsy boxes that shield our current, un-super AIs. While this shouldn’t trigger our immediate extinction, plenty of danger lies here.

https://arstechnica.com/?p=1919705

US Copyright Office withdraws copyright for AI-generated comic artwork

Enlarge / The cover of “Zarya of the Dawn,” a comic book created using Midjourney AI image synthesis in 2022.

Kris Kashtanova

reader comments

137 with

Share this story

On Tuesday, the US Copyright Office declared that images created using the AI-powered Midjourney image generator for the comic book *Zarya of the Dawn* should not have been granted copyright protection, and the images' copyright protection will be revoked.

In a letter addressed to the attorney of author Kris Kashtanova obtained by Ars Technica, the office cites "incomplete information" in the original copyright registration as the reason it plans to cancel the original registration and issue a new one excluding protection for the AI-generated images. Instead, the new registration will cover only the text of the work and the arrangement of images and text. Originally, Kashtanova did not disclose that the images were created by an AI model.

"We conclude that Ms. Kashtanova is the author of the Work's text as well as the selection, coordination, and arrangement of the Work's written and visual elements," reads the copyright letter. "That authorship is protected by copyright. However, as discussed below, the images in the Work that were generated by the Midjourney technology are not the product of human authorship."

Last September, in a story that first appeared on Ars Technica, Kashtanova publicly announced that *Zarya of the Dawn*, which includes comic-style illustrations generated from prompts using the latent diffusion AI process, had been granted copyright registration. At the time, we considered it a precedent-setting case for registering artwork created by

latent diffusion.

Enlarge / An excerpt from the AI-assisted comic book *Zarya of the Dawn*, which initially received a US copyright registration in September, now revised.

However, as the letter explains, after the Copyright Office learned that the work included AI-generated images through Kashtanova's social media posts, it issued a notice to Kashtanova in October stating that it intended to cancel the registration unless she provided additional information showing why the registration should not be canceled. Kashtanova's attorney responded to the letter in November with an argument that Kashtanova authored every aspect of the work, with Midjourney serving merely as an assistive tool.

Advertisement

That argument wasn't good enough for the Copyright Office, which describes in detail why it believes AI-generated artwork should not be granted copyright protection. In a key excerpt provided below, the Office emphasizes the images' machine-generated origins:

> *Based on the record before it, the Office concludes that the images generated by Midjourney contained within the Work are not original works of authorship protected by copyright. See COMPENDIUM (THIRD) § 313.2 (explaining that "the Office will not register works produced by a machine or mere mechanical process that operates randomly or automatically without any creative input or intervention from a human author"). Though she claims to have "guided" the structure and content of each image, the process described in the Kashtanova Letter makes clear that it was Midjourney—not Kashtanova—that originated the "traditional elements of authorship" in the images.*

The letter provides additional analogies for understanding why the Copyright Office thinks Kashtanova is not the creator of the images, including the idea of hiring a human to create images using descriptions and performing a text-based image

search on the Internet. The overall argument in the letter may serve as an important legal precedent for future attempts to copyright AI-generated images.

In an Instagram post, Kashtanova reacted to the letter by framing it as an overall win for AI-augmented artists. She says the ruling is “great news” in the sense that it protects the comic book’s story and the image arrangement, which “covers a lot of uses for the people in the AI art community.”

But on the issue of losing copyright protection for the individual images, Kashtanova says she is not giving up the fight:

> *I was disappointed in one aspect of the decision. The Copyright Office didn’t agree to recognize my copyright of the individual images. I think that they didn’t understand some of the technology so it led to a wrong decision. It is fundamental to understand that the output of a Generative AI model depends directly on the creative input of the artist and is not random. My lawyers are looking at our options to further explain to the Copyright Office how individual images produced by Midjourney are [a] direct expression of my creativity and therefore copyrightable.*

Despite precedents for earlier algorithmically generated artwork receiving copyright protection, this ruling means that AI-generated imagery, without human-authored elements, cannot currently be copyrighted in the United States. The Copyright Office’s ruling on the matter will likely hold unless it’s challenged in court, revised by law, or re-examined in the future.

It’s possible that the ruling may eventually be reconsidered as the result of a cultural shift in how society perceives AI-generated art—one that may allow for a new interpretation by different members of the US Copyright Office in the decade ahead. For now, AI-powered artwork is still a novel and poorly

understood technology, but it may eventually become the standard way visual arts emerge. Not allowing for copyright protection would potentially preclude its use by large and powerful media conglomerates in the future. So the story of AI and copyrights is not over yet.

https://arstechnica.com/?p=1919538

Generative AI is coming for the lawyers

James Marshall / Getty Images

reader comments

76 with

Share this story

David Wakeling, head of London-based law firm Allen & Overy's markets innovation group, first came across law-focused generative AI tool Harvey in September 2022. He approached OpenAI, the system's developer, to run a small experiment. A handful of his firm's lawyers would use the system to answer simple questions about the law, draft documents, and take first passes at messages to clients.

The trial started small, Wakeling says, but soon ballooned. Around 3,500 workers across the company's 43 offices ended up using the tool, asking it around 40,000 queries in total. The

law firm has now entered into a partnership to use the AI tool more widely across the company, though Wakeling declined to say how much the agreement was worth. According to Harvey, one in four at Allen & Overy's team of lawyers now uses the AI platform every day, with 80 percent using it once a month or more. Other large law firms are starting to adopt the platform too, the company says.

The rise of AI and its potential to disrupt the legal industry has been forecast multiple times before. But the rise of the latest wave of generative AI tools, with ChatGPT at its forefront, has those within the industry more convinced than ever.

Advertisement
"I think it is the beginning of a paradigm shift," says Wakeling. "I think this technology is very suitable for the legal industry."

Generative AI is having a cultural and commercial moment, being touted as the future of search, sparking legal disputes over copyright, and causing panic in schools and universities.

The technology, which uses large datasets to learn to generate pictures or text that appear natural, could be a good fit for the legal industry, which relies heavily on standardized documents and precedents.

"Legal applications such as contract, conveyancing, or license generation are actually a relatively safe area in which to employ ChatGPT and its cousins," says Lilian Edwards, professor of law, innovation, and society at Newcastle University. "Automated legal document generation has been a growth area for decades, even in rule-based tech days, because law firms can draw on large amounts of highly standardized templates and precedent banks to scaffold document generation, making the results far more predictable than with most free text outputs."

But the problems with current generations of generative AI have already started to show. Most significantly, their tendency to confidently make things up—or "hallucinate." That is problematic enough in search, but in the law, the difference between success and failure can be serious, and costly.

https://arstechnica.com/?p=1919301

Limited data sets a hurdle as China plays catch-up to ChatGPT

Enlarge / Baidu is leading the ChatGPT AI charge in China with plans to incorporate its Ernie chatbot into its search engine from next month.

reader comments

39 with

Share this story

China's tech giants, including Baidu, Alibaba, and NetEase, are racing to match the West's recent developments in artificial intelligence, touting projects that they hope will achieve the same buzz created by the release of ChatGPT.

After months of announcing cost cuts and headcount reductions, big groups are now optimistically presenting investment plans

to rival OpenAI’s chatbot, while trademark trolls are lining up to claim words related to ChatGPT’s achievements.

Zhou Hongyi, head of Internet security company Qihoo 360, characterized ChatGPT, a program that produces realistic text answers to questions posed by humans, as the start of the artificial intelligence revolution. “It has shortcomings but also unlimited potential,” he said in a talk-show discussion last week.

The race to match ChatGPT comes as the Chinese tech groups have spent the better part of two years reeling from regulatory attacks and fighting for growth amid a COVID-induced slowdown.

Baidu is taking the most concrete early steps, with plans to launch a chatbot named Ernie in its search engine in the next few months, similar to Microsoft and OpenAI’s Bing Chat.

The AI model underlying the bot has been in development since 2019, with its newest generation trained on 260 billion parameters—comparable to GPT3, the technology underpinning ChatGPT, in terms of size, although trained on a much smaller data set.

Baidu plans to reveal new details this week on how the chatbot will be integrated into its products, including search, electric vehicles, and smart assistants, according to a person close to the company.

“Baidu has focused talent and money on this, so they are the most likely to build one of China’s leading GPT platforms,” said Boris Van, an analyst at Bernstein tracking China’s AI efforts. “They have a lot riding on the launch.”

https://arstechnica.com/?p=1918981

Viral Instagram photographer has a confession: His photos are AI-generated

Enlarge / Jos Avery uses Midjourney, an AI image synthesis model, to create images that he then retouches and posts on Instagram as “photos.”

reader comments

114 with

Share this story

With over 26,000 followers and growing, Jos Avery’s Instagram account has a trick up its sleeve. While it may appear to showcase stunning photo portraits of people, they are not actually people at all. Avery has been posting AI-generated portraits for the past few months, and as more fans praise his apparently masterful photography skills, he has grown nervous about telling the truth.

“[My Instagram account] has blown up to nearly 12K followers since October, more than I expected,” wrote Avery when he first reached out to Ars Technica in January. “Because it is where I post AI-generated, human-finished portraits. Probably 95%+ of the followers don’t realize. I’d like to come clean.”

Avery emphasizes that while his images are not actual photographs (except two, he says), they still require a great

deal of artistry and retouching on his part to pass as photorealistic. To create them, Avery initially uses Midjourney, an AI-powered image synthesis tool. He then combines and retouches the best images using Photoshop.

With Midjourney, anyone can pay a subscription fee for the privilege of generating art from text-based descriptions, called "prompts." Midjourney's creators taught the AI model how to synthesize images by showing it millions of examples of art from other artists. It can generate stunning photorealistic images that can fool some people into thinking they're real photos, especially if retouched later.

Advertisement

- "Classy Cameron" by Jos Avery. Avery says two of these photos were taken by a camera and the rest were AI-generated, then retouched. Can you figure out which ones are the real photos?
- "Savanna Snow" by Jos Avery. Avery says two of these photos were taken by a camera and the rest were AI-generated, then retouched. Can you figure out which ones are the real photos?
- "Jocose Jared" by Jos Avery. Avery says two of these photos were taken by a camera and the rest were AI-generated, then retouched. Can you figure out which ones are the real photos?
- "Adept Albert" by Jos Avery. Avery says two of these photos were taken by a camera and the rest were AI-generated, then retouched. Can you figure out which ones are the real photos?
- "Generous Gwendolyn" by Jos Avery. Avery says two of these photos were taken by a camera and the rest were AI-generated, then retouched. Can you figure out which ones are the real photos?
- "Cold Oleksiy" by Jos Avery. Avery says two of these photos were taken by a camera and the rest were AI-generated, then retouched. Can you figure out which ones

are the real photos?

Originally an AI skeptic, Avery has become a convert to the new art form. Such work attracts great controversy in the art world, partly due to ethical issues around scraping human-made artwork without consent. But thanks to that artistic knowledge built into the model, some of the most skilled AI-augmented practitioners can render imagery far more vividly than if a human were working alone.

“I am honestly conflicted,” Avery said when he approached Ars to tell his story. “My original aim was to fool people to showcase AI and then write an article about it. But now it has become an artistic outlet. My views have changed.”

Painted into a digital corner

Soon after Avery’s Instagram feed launched in October, positive comments about his fake photos began pouring in. “All I can say is: Your art is somehow unique, very unique, also very precious; you are actually telling paramount stories to the viewer using your cams,” wrote one commenter four weeks ago. “Setting novel highlights in contemporary photography IMHO! Your work is a great delight to mind and soul.”

Up until very recently, when asked, Avery was either vague about how he created the images or told people his works were actual photographs, even going so far as to describe which kind of camera he used to create them (“a Nikon D810 with 24-70mm lens”). But guilt began to build as his popularity grew.

https://arstechnica.com/?p=1917218