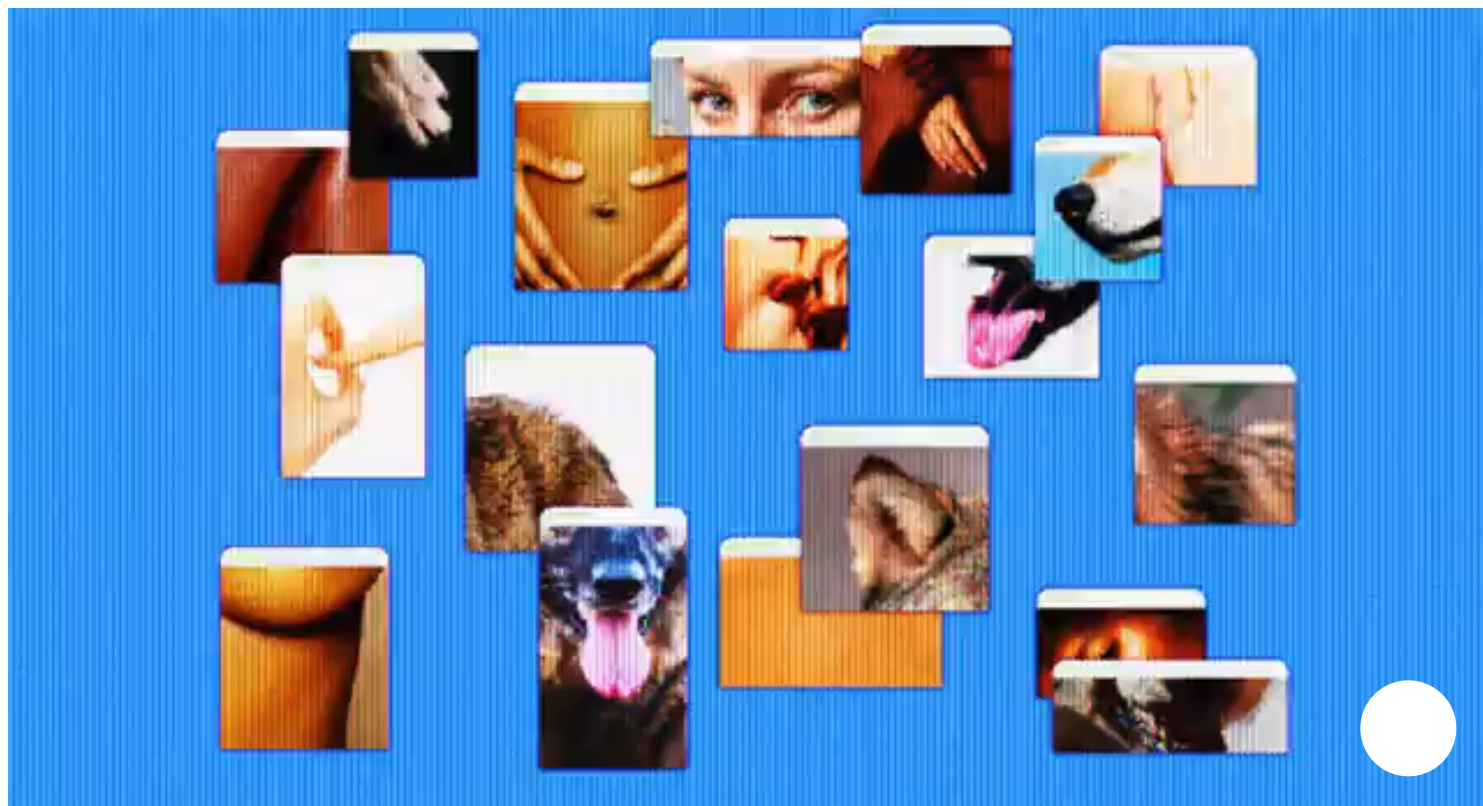


ROSE EVELETH CULTURE MAY 15, 2023 6:00 AM

The Fanfic Sex Trope That Caught a Plundering AI Red-Handed

Sudowrite, a tool that uses OpenAI's GPT-3, was found to have understood a sexual act known only to a specific online community of Omegaverse writers.



VIDEO: JAMES MARSHALL; GETTY IMAGES



THESE DAYS, SO-CALLED generative AI can (allegedly) make art, write books, and compose poetry. Systems like Stable Diffusion, Midjourney, and ChatGPT are seemingly quite good at it. But for some artists, this creates problems. Namely, determining what legal rights they have when their work is scraped by these tools.

Faced by the rise in these systems, authors and artists are pushing back. The Writers Guild of America (WGA) is striking in part over the potential use of AI to write scripts, referring to such systems as “plagiarism machines.” Visual artists have penned open letters denouncing the use of AI to replace illustrators, calling it “the greatest art heist in history.” Getty sued Stability AI in January for copyright infringement.

But what if your work exists in a kind of in-between space—not work that you make a living doing, but still something you spent hours crafting, in a community that you care deeply about? And what if, within that community, there was a specific sex trope that would inadvertently unmask how models like ChatGPT scrape the web—and how that scraping impacts the writers who created it.

THE TROPE IN question is called “the Omegaverse,” which is perhaps best described as an act of collective sexual worldbuilding. It began in the (very active) fandom for the TV series *Supernatural*, but has now spread to almost every corner of the fan-fiction world. These stories are defined by a specific sexual hierarchy made up of Alphas, Betas, and Omegas in which Alphas and Omegas can smell one another in particular ways, experience “heats,” and (usually) mate for life. Most of these stories are heavy on smut, and bodily fluids are crucial to the whole genre.

ARTIFICIAL INTELLIGENCE

***Death of an Author* Prophesies the Future of AI Novels**

TOM COMITTA

DIGITAL CULTURE

Sci-Fi Publishers Are Bracing for an AI Battle

ELIZABETH MINKEL

PITCH PERFECT

The *Ted Lasso* Fandom's Push for Polyamory Just Makes Sense

MARAH EAKIN

Within the Omegaverse, there is also something called “knotting,” a phenomenon borrowed from animals in which a penis grows a bulb at the base to remain locked inside a vagina. If this all sounds overwhelming, you’re not alone. “I remember the first time I encountered it, and I will confess, my reaction was, ‘What is this? What is happening?’” says Hayley Krueger, a fan-fiction writer who published [an Omegaverse 101 explainer](#) earlier this year. But she says she quickly fell in love with the trope.

When characters in the Omegaverse mate, they become linked biologically. Different writers have different ways of showing or expressing this—anything from being able to smell your mate’s mood, to being able to actually communicate telepathically across distances. “I really like the dynamic between characters,” Krueger says. “It’s almost like soulmates, but you choose it and then you get all these perks that go with it.”

Because the Omegaverse has such specific terms and phrases associated with it, ones that are found within fan fiction and nowhere else, it’s an ideal way to test how generative AI systems are scraping the web. Determining what information has gone into a model like ChatGPT is almost impossible. OpenAI, the company behind the tool, has [declined to make](#) its training data sources public. *The Washington Post* did their own analysis of the model, [and created a way to peek](#) at the websites that make up Google’s C4 dataset. But even people who build applications using ChatGPT have no insight into what the system is trained on.

To be clear, what this means is that the AI model that is being used across the world for everything from

customer service, to HR training, to journalism, knows what “knotting” is.

In the absence of any list of sources, people have tinkered with other ways to try and explore what these models might know and how. One way to do that is to prompt the system with words or questions you know come from a certain source. So, for example, if you want to know whether the works of Shakespeare are being used in the model, you might give the system a few unique lines of a play and see if it comes back with iambic pentameter. Or, if you want to know whether these systems are trained on fan fiction, you might give the model a specific trope unique to fandom.

A few months ago, a fan-fiction writer with the handle [kafetheresu](#) did just that. In a very thorough [post on Reddit](#), they pointed out that when they gave the writing tool Sudowrite (which uses OpenAI’s [GPT-3](#) to operate) specific words and phrases unique to the Omegaverse, the system readily filled in the rest in a way that suggested the AI knew all about this particular trope. (The Reddit poster declined to speak on the record.)

For example, when the writer plugged in the sentence “Steve had to admit that he had some reservations about how the New Century handled the social balance between Alphas and Omegas” into Sudowrite, it suggested the following next section:

Being an omega himself, he knew how difficult it was to find a balance between the highly competitive alpha-omega relationship and the omega-omega friendship. Especially for someone like Steve, who was highly competitive and wanted to be a career officer.

The only way Sudowrite would know to fill in these specific sentences was if it had learned about the Omegaverse. And the only place for it to learn about the Omegaverse is on fan-fiction sites like Archive of Our Own. I tested Sudowrite with a few lines inspired by Omegaverse fiction as well to verify these results. Here’s what Sudowrite gives you if you start with “The scent stopped Harry dead in his tracks, his neck tingling against his will, bringing back memories of one particular Alpha”:

Harry's heart raced with excitement and fear as he followed the scent. He couldn't help but wonder what had happened to the alpha since they last saw each other. Had he found his mate? Had he forgotten about Harry? The questions swirled around in his head as he approached a dark alleyway.

It might seem odd that a sex trope generally found only on fan-fiction forums like Archive of Our Own ended up in a tool that runs on GPT-3, but it shouldn't be surprising. Most large language models use some version of the CommonCrawl dataset, a corpus made of 12 years' worth of crawling the publicly available internet. Archive of Our Own hosts more than 11,080,000 works, making it a treasure trove of publicly accessible content.

To be clear, what this means is that the AI model being used across the world for everything from customer service, to HR training, to journalism, knows what knotting is and could, without the right guardrails, explain to an unassuming user why an Omega needs to produce slick in order to mate.

JAMES YU, THE chief technology officer at Sudowrite, says his team noticed the Reddit post fairly quickly. He told me that it was eye-opening because it highlighted how vast the data sets that go into these models really are. "For me, it highlights the things I don't know," says Yu. "In every one of these models is millions of other latent spaces that I just never encounter. It's almost like an endless ocean."

Sudowrite is intended to be used as a writer's assistant; authors plug in sentences that are giving them trouble, or scenes they're working on, and the AI offers up a few lines to help guide them on where they could go next. Prior to the Reddit post, Yu had no idea what the Omegaverse was. Now his own system was offering tips on how to write smut about it.

Writers of fan fiction, much like writers of journalism or television or movies, were not pleased to find out that their work was being used to train these systems. "This is particularly concerning as many for-profit AI writing programs like Sudowrite, WriteSonic, and others utilized GPT-3," the original Reddit poster wrote in an email to the Archive of Our Own communications team, shared in the thread. "These AI

apps take the works which we create for fun and fandom, not only to gain profit, but also to one day replace human writing.”

Even if it's just plain smut, there's a human element there and it's someone creating something for their enjoyment and they want to share that hard work with people. It's stealing that.

— HAYLEY KRUEGER, FAN-FICTION WRITER

Yu is aware of this complaint. “I'd love for there to be a simple way to do fair compensation for content that was used to train GPT-3, but unfortunately, there is no mechanism that OpenAI provides for that,” he says. “If someone (OpenAI? Google?) were to offer this, we'd try it out immediately.” For now, he believes that Sudowrite's value to writers outweighs the harm. “We're one of the few AI platforms that is catering specifically to fiction writers,” he says, adding that when there is a better model that has opt-in features, and potential payment for people's work, “we will be in a good position to switch to it, and this is something we would promote heavily to our user base to bring awareness.”

But that's not convincing to a lot of writers who feel their work is being used against their will to enrich technology companies. Compared to Google and OpenAI, Sudowrite is small potatoes, but they have still raised \$3 million in seed funding. As far as writers are concerned, it's not enough for a place like Sudowrite to wait around for some other, bigger company to fix what they see as fundamental, unethical flaws in the system. In the comments on the Reddit post, one user said: “God I hate AI so much on so many different levels.” Others shared tips on how to make their fan works private. “I never liked the idea of hiding my work, but because of this I went and restricted everything I've written so only registered users can see it,” wrote another.

“It sort of takes the heart out of it,” says Krueger. “Fan fiction is used by a lot of creators to explore difficult topics that are personal to them and their life experiences as a way to vent about these topics. Even if it's just smut, just plain smut,

there's a human element there and it's someone creating something for their enjoyment and they want to share that hard work with people. It's stealing that from people.”

This is the same argument being made by the WGA—that these systems can use copyrighted (or copyrightable) work against the authors’ will. This applies to fan fiction too.

IT MIGHT NOT be done for profit, but fan fiction is eligible for copyright claims. While the writers do not own the content on which they’re basing these pieces of fiction, they do own anything original they contribute through their work.

For those people, non-commerciality is of value, and the idea of someone else making money off it is highly offensive because their moral commitments are being betrayed.

— BETSY ROSENBLATT, LEGAL CHAIR, ORGANIZATION FOR TRANSFORMATIVE WORKS

“Even things that are highly derivative, if they originated with the author, are considered original,” says Betsy Rosenblatt, a professor at University of Tulsa College of Law and the legal chair for the [Organization for Transformative Works \(OTW\)](#), a nonprofit that oversees Archive of Our Own among other fanwork projects. That includes things like any original characters an author added, the plot structure, and the particular word choices. So it is possible in some situations to file for copyright protection for works of fan fiction—although most fan-fiction writers don’t, whether it’s because they don’t know how, or don’t want to spend the money, or simply aren’t interested in jumping through the hoops.

But for most writers I spoke with, it’s not really about copyright or ownership or even money. Most fan-fiction authors don’t make a living doing this. They do it for the community, for the friends and connections they make. “I have so many friends that I’ve met through partaking in events where we create stuff together,” says Krueger. And Rosenblatt says that people who are unhappy with scraping see this as a major problem. For them, it’s not that they are being deprived of potential income,

but instead that someone is making money off of something that they created specifically to be non-commercial.

“For those people, non-commerciality is of value, and the idea of someone else making money off it is highly offensive because their moral commitments are being betrayed,” Rosenblatt says.

And perhaps because there isn’t a big financial driver, the culture of fan fiction is all about attribution—writers link and nod to other people who’ve influenced them, or helped them. “The idea is that no one should get paid for this, but everyone should know what’s mine,” Rosenblatt says. This is not simply difficult to do with AI systems, but is in fact nearly impossible. Many of these models are black boxes, and it would be impossible to spit out a list of influences that contributed to something specific ChatGPT wrote.

SO, CAN AI systems like Sudowrite and writers who don’t want to be used by them exist in harmony? Nobody knows, of course, but most of the people I spoke with talked about some form of opting in. Rosenblatt says that some writers of fan fiction really like the ability to use AI in their work. Sudowrite certainly has fans in the writing world. Others want nothing to do with these systems, and want the ability to remove their work from the training data. “I would love to get to a place where we could have a totally opt-in model and everyone is compensated for that,” says Yu, “I just don’t think that’s possible right now.”

Yu says that if people were able to opt out at scale, then the models would become noticeably worse. The reason ChatGPT works as well as it does is precisely because it’s got so much data to pull from. Critics argue that if the only way your system can function is by using work against people’s wishes, then perhaps the system itself is fundamentally morally flawed.

Fan fiction might seem like an easy mark when it comes to training models. These pieces are publicly available, non-commercial, and often not copyrighted. But that doesn’t mean they aren’t valuable and worth protecting from being used in ways that the original creators don’t like.

A weekly dispatch from the future by Will Knight, exploring AI advances and other technology set to change our lives. Delivered every Thursday.

Your email

Enter your email




SUBMIT

By signing up you agree to our [User Agreement](#) (including the [class action waiver and arbitration provisions](#)), our [Privacy Policy & Cookie Statement](#) and to receive marketing and account-related emails from WIRED. You can unsubscribe at any time.

In 2019, Archive of Our Own won a Hugo Award for Best Related Work. At the ceremony, nominees asked every science fiction writer who had ever contributed to the site to stand, and a huge chunk of the room did. The value of this kind of community-based, collective worldbuilding is often dismissed as silly or frivolous, but these works are important to millions of people around the world. “I have read fan fiction that has affected me emotionally and lived with me in ways that stories I’ve read that are published have not,” says Krueger.

In the efforts to consider the future of generative AI, and whose work does or doesn’t get used to train it, even smutty fan fiction deserves protection.

Get More From WIRED

-  Understand AI advances with our [Fast Forward newsletter](#)
-  Our new [podcast](#) wants you to [Have a Nice Future](#)
- [TikTok and Instagram filters](#) aren’t trying to fool anyone
- [Facebook’s 2024 election policy](#) may hinge on a Cambodian video
- [Disney’s ESPN Bet](#) is a play for younger gamblers
- You are not responsible for [your own online privacy](#)
- The internet is turning into a data black box. An “[inspectability API](#)” could crack it open
-  Charge right into summer with the best [travel adapters](#), [power banks](#), and [USB hubs](#)



Rose Eveleth ([@roseveleth](#)) is an Ideas contributor at WIRED. They're also the host and producer of “[Flash Forward](#),” a podcast about possible (and not so possible) future scenarios, and has covered everything from fake tumbleweed farms to million-dollar baccarat heists.

IDEAS CONTRIBUTOR 

TOPICS FICTION ARTIFICIAL INTELLIGENCE CHATGPT

MORE FROM WIRED

The 20 Best Movies on Max (aka HBO Max) Right Now

From *Edge of Tomorrow* to *Pulp Fiction*, here are our favorite movies on the streaming service.

JENNIFER M. WOOD