Artificial Creativity: A Call for Regulation

Once upon a time, in a bustling city where skyscrapers reached for the clouds and the hum of technology was ever-present, there lived a girl named Emily. From a young age, Emily found solace in writing stories on her daily bus rides to school and crafting heartfelt poems in her spare time. Her words danced on paper, painting vivid images and evoking powerful emotions.

As Emily grew older, her love for writing only deepened. However, a shadow of fear began to loom over her passion. With the rapid advancements in artificial intelligence (AI) and machine learning, there were whispers in the literary world that AI could one day replace human creativity, including writing.

Despite her worries, Emily's curiosity led her to explore the world of data science. She discovered the beauty of analyzing data, finding patterns, and uncovering insights. Yet, her heart still yearned for the magic of storytelling and the art of weaving words together.

You might be thinking: "that story sounds lame." Who cares about some random girl named Emily in a random bustling city writing some random little poems? Anyway, the writing to that introduction was clunky and strange and oh my gosh— we need to sound the AI writing alarms!



You

tell a story about a girl who grew up writing stories on the bus and poems in her spare time. talk about the fear she experiences as AI will replace creative writing, despite the fact that she began to study data science.





ChatGPT

Memory updated

Once upon a time, in a bustling city where skyscrapers reached for the clouds and the hum of technology was ever-present, there lived a girl named Emily. From a young age, Emily found solace in writing stories on her daily bus rides to school and crafting heartfelt poems in her spare time. Her words danced on paper, painting vivid images and evoking powerful emotions.

Now, let me tell you my story. Hold on. Given the context of my introduction: let **me** tell you my story. I grew up as a lonely only child in a metropolitan city. As a kid, both of my parents worked full time. My father strongly encouraged me to read, and filled my shelves with books — the only thing I had waiting for me at home after school. After spending hours upon hours reading, I naturally started writing stories. I wrote my first book on a scrappy notebook I bought from the school store for 30 cents when I was

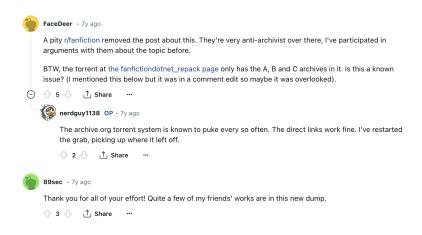
eleven. When I was in eleventh grade, I won the regional short story Gold Key for the Scholastic Art and Writing awards. Before my college admissions, I put together a creative writing portfolio to apply to the prestigious Iowa Writers' Workshop. I never submitted it. Because I was worried about not being able to support myself and my family as a writer, I chose to study a more traditional, professional path.

In April of 2023, when the Writer's Guild of America officially went on strike. One of the core tenants of the strike was to draw up formal regulations against the use of Al in major studio writer's rooms as a cheaper alternative to human writers. The training of any Al tool to write falls under the development of a Large Language Model (LLM), which requires huge amounts of training data as close to the work it is meant to imitate. This means that Al creative writing models can only be created by feeding a LLM existing creative writing as training data. Writers should be fairly compensated for their labor in producing works that have gone on to train these LLM — however, right now there are no such laws that require fair compensation or even consent (Rogin et al., 2023).

As someone who has grown up deeply invested in the literary arts, I know exactly where to find treasure troves of creative writing published by independent hobby writers. I grew up writing on Storybird, a creativity tool that allows anyone to publish web-novels inspired from custom and credited art submitted to the site. Many of my friends grew extremely attached to fan communities such as Harry Potter, Percy Jackson, and other famous books and cartoons. They started writing fan works about the characters and lore, expanding on the existing works; these words have been affectionately dubbed as "fanfiction". At an impressive twelve million words, esteemed novelist Marcel Proust currently holds the Guiness World Record for the longest novel in the world. More impressively, the fan work *The Loud House: Revamped* (based off the Nickelodeon cartoon Loud House) charts 16 million words and counting. The internet is a vast resource of endless amateur writing. But it's also a vast resource to learn about machine learning.



A short internet search also led me to find this, an open-source Github repository with a custom package for web scraping fanfiction. It even has custom attributes such as "canon type", which indicates the extent in which the fanfiction follows the "canonical" events of the source material. This package is clearly stated to be used for non-commercial, research purposes. And on my brief perusal of online community sites like Reddit, Quora, and Fanfiction.net there is a good amount of people who scrape these fanfiction sites regularly to preserve works in case of a site shutdown. But this access to web-scraping is completely unregulated.



If I wanted to flex my fingers and have a practice try at building an LLM, say for a harmless school project or resume portfolio, I would have easy and free access to an endless black hole of fan writing without even anyone asking me to pause and think about whether it is ethical for me to be scraping this data. Did the writers consent? Do I need to give them compensation? What's clear to me is that our regulations are struggling to keep up with the growth of artificial intelligence. But more than that, the internet and social media are also severely under-regulated; in context of the 2016 US Elections manipulation scandals and other controversies, we have barely been able to control the tool that revolutionized our world in 1983.

"Al stresses me out." That's a quote from Elon Musk, one of the primary investors and co-founders of OpenAi, which rocked the world when ChatGPT burst onto the scene in 2022 with the most revolutionary "Hello world" in a long time. Elon Musk called for a six-month pause in Al research (Naranayan et al., 2023). But is that enough? We need a consistent and dedicated effort to regulate this new technology. The democratization of online resources to both train and build Al models as powered by the internet means that Al will grow faster and in more unexpected ways than ever before. Writing is already

seen as a difficult job with little financial stability. They contribute so much to the entertainment and joy that we experience, but by nature of their profession, they are already financially vulnerable. We need to be ready to protect the labor, livelihoods, and sheer passion that our writers have, or we will lose literary culture. Are we ready?

Sources:

- Guiness World Records. (n.d.). *Longest novel*. Guinness World Records. https://www.guinnessworldrecords.com/world-records/longest-novel
- Jamesdean5842. (n.d.). *The loud house: Revamped Chapter 1: Loud House revamped prologue, ...* Fanfiction.Net. https://www.fanfiction.net/s/12715870/1/The-Loud-House-Revamped
- Naranayan, J., Mukherjee, S., Coulter, M., & Hu, K. (2023, April 5). *Elon Musk and others urge ai pause, citing "risks to society"* | ... Reuters. https://www.reuters.com/technology/musk-experts-urge-pause-training-ai-systems-that-can-outperform-gpt-4-2023-03-29/
- Rogin, A., Corkery, A., & Roscoe, J. (2023, September 2). Why Artificial Intelligence is a central dispute in the Hollywood strikes. PBS.

 https://www.pbs.org/newshour/show/why-artificial-intelligence-is-a-central-dispute-in-the-hollywood-strikes

Multimedia:

Fanfiction Web Scraping Package: https://github.com/smilli/fanfiction

ChatGPT Introduction Query:

https://chat.openai.com/share/f6862221-b481-4555-a608-e41e8223a584