**Why A.I. Isn't Going to Make Art** (2024)
by Ted Chiang

In 1953, Roald Dahl published "The Great Automatic Grammatizator," a short story about an electrical engineer who secretly desires to be a writer. One day, after completing construction of the world's fastest calculating machine, the engineer realizes that "English grammar is governed by rules that are almost mathematical in their strictness." He constructs a fiction-writing machine that can produce a five-thousand-word short story in thirty seconds; a novel takes fifteen minutes and requires the operator to manipulate handles and foot pedals, as if he were driving a car or playing an organ, to regulate the levels of humor and pathos. The resulting novels are so popular that, within a year, half the fiction published in English is a product of the engineer's invention.

Is there anything about art that makes us think it can't be created by pushing a button, as in Dahl's imagination? Right now, the fiction generated by large language models like ChatGPT is terrible, but one can imagine that such programs might improve in the future. How good could they get? Could they get better than humans at writing fiction—or making paintings or movies—in the same way that calculators are better at addition and subtraction?

Art is notoriously hard to define, and so are the differences between good art and bad art. But let me offer a generalization: art is something that results from making a lot of choices. This might be easiest to explain if we use fiction writing as an example. When you are writing fiction, you are—consciously or unconsciously—making a choice about almost every word you type; to oversimplify, we can imagine that a ten-thousand-word short story requires something on the order of ten thousand choices. When you give a generative-A.I. program a prompt, you are making very few choices; if you supply a hundred-word prompt, you have made on the order of a hundred choices.

If an A.I. generates a ten-thousand-word story based on your prompt, it has to fill in for all of the choices that you are not making. There are various ways it can do this. One is to take an average of the choices that other writers have made, as represented by text found on the Internet; that average is equivalent to the least interesting choices possible, which is why A.I.-generated text is often really bland. Another is to instruct the program to engage in style mimicry, emulating the choices made by a specific writer, which produces a highly derivative story. In neither case is it creating interesting art.

I think the same underlying principle applies to visual art, although it's harder to quantify the choices that a painter might make. Real paintings bear the mark of an enormous number of decisions. By comparison, a person using a text-to-image program like DALL-E enters a prompt such as "A knight in a suit of armor fights a fire-breathing dragon," and lets the program do the rest. (The newest version of DALL-E accepts prompts of up to four thousand characters—hundreds of words, but not enough to describe every detail of a scene.) Most of the choices in the

resulting image have to be borrowed from similar paintings found online; the image might be exquisitely rendered, but the person entering the prompt can't claim credit for that.

Some commentators imagine that image generators will affect visual culture as much as the advent of photography once did. Although this might seem superficially plausible, the idea that photography is similar to generative A.I. deserves closer examination. When photography was first developed, I suspect it didn't seem like an artistic medium because it wasn't apparent that there were a lot of choices to be made; you just set up the camera and start the exposure. But over time people realized that there were a vast number of things you could do with cameras, and the artistry lies in the many choices that a photographer makes. It might not always be easy to articulate what the choices are, but when you compare an amateur's photos to a professional's, you can see the difference. So then the question becomes: Is there a similar opportunity to make a vast number of choices using a text-to-image generator? I think the answer is no. An artist—whether working digitally or with paint—implicitly makes far more decisions during the process of making a painting than would fit into a text prompt of a few hundred words.

We can imagine a text-to-image generator that, over the course of many sessions, lets you enter tens of thousands of words into its text box to enable extremely fine-grained control over the image you're producing; this would be something analogous to Photoshop with a purely textual interface. I'd say that a person could use such a program and still deserve to be called an artist. The film director Bennett Miller has used DALL-E 2 to generate some very striking images that have been exhibited at the Gagosian gallery; to create them, he crafted detailed text prompts and then instructed DALL-E to revise and manipulate the generated images again and again. He generated more than a hundred thousand images to arrive at the twenty images in the exhibit. But he has said that he hasn't been able to obtain comparable results on later releases of DALL-E. I suspect this might be because Miller was using DALL-E for something it's not intended to do; it's as if he hacked Microsoft Paint to make it behave like Photoshop, but as soon as a new version of Paint was released, his hacks stopped working. OpenAI probably isn't trying to build a product to serve users like Miller, because a product that requires a user to work for months to create an image isn't appealing to a wide audience. The company wants to offer a product that generates images with little effort.

It's harder to imagine a program that, over many sessions, helps you write a good novel. This hypothetical writing program might require you to enter a hundred thousand words of prompts in order for it to generate an entirely different hundred thousand words that make up the novel you're envisioning. It's not clear to me what such a program would look like. Theoretically, if such a program existed, the user could perhaps deserve to be called the author. But, again, I don't think companies like OpenAI want to create versions of ChatGPT that require just as much effort from users as writing a novel from scratch. The selling point of generative A.I. is that these programs generate vastly more than you put into them, and that is precisely what prevents them from being effective tools for artists.

The companies promoting generative-A.I. programs claim that they will unleash creativity. In essence, they are saying that art can be all inspiration and no perspiration—but these things cannot be easily separated. I'm not saying that art has to involve tedium. What I'm saying is that art requires making choices at every scale; the countless small-scale choices made during implementation are just as important to the final product as the few large-scale choices made during the conception. It is a mistake to equate "large-scale" with "important" when it comes to the choices made when creating art; the interrelationship between the large scale and the small scale is where the artistry lies.

Believing that inspiration outweighs everything else is, I suspect, a sign that someone is unfamiliar with the medium. I contend that this is true even if one's goal is to create entertainment rather than high art. People often underestimate the effort required to entertain; a thriller novel may not live up to Kafka's ideal of a book—an "axe for the frozen sea within us"—but it can still be as finely crafted as a Swiss watch. And an effective thriller is more than its premise or its plot. I doubt you could replace every sentence in a thriller with one that is semantically equivalent and have the resulting novel be as entertaining. This means that its sentences—and the small-scale choices they represent—help to determine the thriller's effectiveness.

Many novelists have had the experience of being approached by someone convinced that they have a great idea for a novel, which they are willing to share in exchange for a fifty-fifty split of the proceeds. Such a person inadvertently reveals that they think formulating sentences is a nuisance rather than a fundamental part of storytelling in prose. Generative A.I. appeals to people who think they can express themselves in a medium without actually working in that medium. But the creators of traditional novels, paintings, and films are drawn to those art forms because they see the unique expressive potential that each medium affords. It is their eagerness to take full advantage of those potentialities that makes their work satisfying, whether as entertainment or as art.

Of course, most pieces of writing, whether articles or reports or e-mails, do not come with the expectation that they embody thousands of choices. In such cases, is there any harm in automating the task? Let me offer another generalization: any writing that deserves your attention as a reader is the result of effort expended by the person who wrote it. Effort during the writing process doesn't guarantee the end product is worth reading, but worthwhile work cannot be made without it. The type of attention you pay when reading a personal e-mail is different from the type you pay when reading a business report, but in both cases it is only warranted when the writer put some thought into it.

Recently, Google aired a commercial during the Paris Olympics for Gemini, its competitor to OpenAI's GPT-4. The ad shows a father using Gemini to compose a fan letter, which his daughter will send to an Olympic athlete who inspires her. Google pulled the commercial after widespread backlash from viewers; a media professor called it "one of the most disturbing commercials I've ever seen." It's notable that people reacted this way, even though artistic

creativity wasn't the attribute being supplanted. No one expects a child's fan letter to an athlete to be extraordinary; if the young girl had written the letter herself, it would likely have been indistinguishable from countless others. The significance of a child's fan letter—both to the child who writes it and to the athlete who receives it—comes from its being heartfelt rather than from its being eloquent.

Many of us have sent store-bought greeting cards, knowing that it will be clear to the recipient that we didn't compose the words ourselves. We don't copy the words from a Hallmark card in our own handwriting, because that would feel dishonest. The programmer Simon Willison has described the training for large language models as "money laundering for copyrighted data," which I find a useful way to think about the appeal of generative-A.I. programs: they let you engage in something like plagiarism, but there's no guilt associated with it because it's not clear even to you that you're copying.

Some have claimed that large language models are not laundering the texts they're trained on but, rather, learning from them, in the same way that human writers learn from the books they've read. But a large language model is not a writer; it's not even a user of language. Language is, by definition, a system of communication, and it requires an intention to communicate. Your phone's auto-complete may offer good suggestions or bad ones, but in neither case is it trying to say anything to you or the person you're texting. The fact that ChatGPT can generate coherent sentences invites us to imagine that it understands language in a way that your phone's auto-complete does not, but it has no more intention to communicate.

It is very easy to get ChatGPT to emit a series of words such as "I am happy to see you." There are many things we don't understand about how large language models work, but one thing we can be sure of is that ChatGPT is not happy to see you. A dog can communicate that it is happy to see you, and so can a prelinguistic child, even though both lack the capability to use words. ChatGPT feels nothing and desires nothing, and this lack of intention is why ChatGPT is not actually using language. What makes the words "I'm happy to see you" a linguistic utterance is not that the sequence of text tokens that it is made up of are well formed; what makes it a linguistic utterance is the intention to communicate something.

Because language comes so easily to us, it's easy to forget that it lies on top of these other experiences of subjective feeling and of wanting to communicate that feeling. We're tempted to project those experiences onto a large language model when it emits coherent sentences, but to do so is to fall prey to mimicry; it's the same phenomenon as when butterflies evolve large dark spots on their wings that can fool birds into thinking they're predators with big eyes. There is a context in which the dark spots are sufficient; birds are less likely to eat a butterfly that has them, and the butterfly doesn't really care why it's not being eaten, as long as it gets to live. But there is a big difference between a butterfly and a predator that poses a threat to a bird.

A person using generative A.I. to help them write might claim that they are drawing inspiration from the texts the model was trained on, but I would again argue that this differs from what we

usually mean when we say one writer draws inspiration from another. Consider a college student who turns in a paper that consists solely of a five-page quotation from a book, stating that this quotation conveys exactly what she wanted to say, better than she could say it herself. Even if the student is completely candid with the instructor about what she's done, it's not accurate to say that she is drawing inspiration from the book she's citing. The fact that a large language model can reword the quotation enough that the source is unidentifiable doesn't change the fundamental nature of what's going on.

As the linguist Emily M. Bender has noted, teachers don't ask students to write essays because the world needs more student essays. The point of writing essays is to strengthen students' critical-thinking skills; in the same way that lifting weights is useful no matter what sport an athlete plays, writing essays develops skills necessary for whatever job a college student will eventually get. Using ChatGPT to complete assignments is like bringing a forklift into the weight room; you will never improve your cognitive fitness that way.

Not all writing needs to be creative, or heartfelt, or even particularly good; sometimes it simply needs to exist. Such writing might support other goals, such as attracting views for advertising or satisfying bureaucratic requirements. When people are required to produce such text, we can hardly blame them for using whatever tools are available to accelerate the process. But is the world better off with more documents that have had minimal effort expended on them? It would be unrealistic to claim that if we refuse to use large language models, then the requirements to create low-quality text will disappear. However, I think it is inevitable that the more we use large language models to fulfill those requirements, the greater those requirements will eventually become. We are entering an era where someone might use a large language model to generate a document out of a bulleted list, and send it to a person who will use a large language model to condense that document into a bulleted list. Can anyone seriously argue that this is an improvement?

It's not impossible that one day we will have computer programs that can do anything a human being can do, but, contrary to the claims of the companies promoting A.I., that is not something we'll see in the next few years. Even in domains that have absolutely nothing to do with creativity, current A.I. programs have profound limitations that give us legitimate reasons to question whether they deserve to be called intelligent at all.

The computer scientist François Chollet has proposed the following distinction: skill is how well you perform at a task, while intelligence is how efficiently you gain new skills. I think this reflects our intuitions about human beings pretty well. Most people can learn a new skill given sufficient practice, but the faster the person picks up the skill, the more intelligent we think the person is. What's interesting about this definition is that—unlike I.Q. tests—it's also applicable to nonhuman entities; when a dog learns a new trick quickly, we consider that a sign of intelligence.

In 2019, researchers conducted an experiment in which they taught rats how to drive. They put the rats in little plastic containers with three copper-wire bars; when the mice put their paws on one of these bars, the container would either go forward, or turn left or turn right. The rats could see a

plate of food on the other side of the room and tried to get their vehicles to go toward it. The researchers trained the rats for five minutes at a time, and after twenty-four practice sessions, the rats had become proficient at driving. Twenty-four trials were enough to master a task that no rat had likely ever encountered before in the evolutionary history of the species. I think that's a good demonstration of intelligence.

Now consider the current A.I. programs that are widely acclaimed for their performance. AlphaZero, a program developed by Google's DeepMind, plays chess better than any human player, but during its training it played forty-four million games, far more than any human can play in a lifetime. For it to master a new game, it will have to undergo a similarly enormous amount of training. By Chollet's definition, programs like AlphaZero are highly skilled, but they aren't particularly intelligent, because they aren't efficient at gaining new skills. It is currently impossible to write a computer program capable of learning even a simple task in only twenty-four trials, if the programmer is not given information about the task beforehand.

Self-driving cars trained on millions of miles of driving can still crash into an overturned trailer truck, because such things are not commonly found in their training data, whereas humans taking their first driving class will know to stop. More than our ability to solve algebraic equations, our ability to cope with unfamiliar situations is a fundamental part of why we consider humans intelligent. Computers will not be able to replace humans until they acquire that type of competence, and that is still a long way off; for the time being, we're just looking for jobs that can be done with turbocharged auto-complete.

Despite years of hype, the ability of generative A.I. to dramatically increase economic productivity remains theoretical. (Earlier this year, Goldman Sachs released a report titled "Gen AI: Too Much Spend, Too Little Benefit?") The task that generative A.I. has been most successful at is lowering our expectations, both of the things we read and of ourselves when we write anything for others to read. It is a fundamentally dehumanizing technology because it treats us as less than what we are: creators and apprehenders of meaning. It reduces the amount of intention in the world.

Some individuals have defended large language models by saying that most of what human beings say or write isn't particularly original. That is true, but it's also irrelevant. When someone says "I'm sorry" to you, it doesn't matter that other people have said sorry in the past; it doesn't matter that "I'm sorry" is a string of text that is statistically unremarkable. If someone is being sincere, their apology is valuable and meaningful, even though apologies have previously been uttered. Likewise, when you tell someone that you're happy to see them, you are saying something meaningful, even if it lacks novelty.

Something similar holds true for art. Whether you are creating a novel or a painting or a film, you are engaged in an act of communication between you and your audience. What you create doesn't have to be utterly unlike every prior piece of art in human history to be valuable; the fact that you're the one who is saying it, the fact that it derives from your unique life experience and arrives at a particular moment in the life of whoever is seeing your work, is what makes it new. We are all

products of what has come before us, but it's by living our lives in interaction with others that we bring meaning into the world. That is something that an auto-complete algorithm can never do, and don't let anyone tell you otherwise.