

AI, Agency, Product, Processes: A Case for Re-evaluating Creativity

The capabilities of new AI models encourage us to reconsider what it takes for a work to be creative.

Problems and Solutions of YouTube Age Verification

YouTube's recent AI age verification policies have led to backlash, for good reason.

The Role of Governmental Regulation in Safe AI Development

Governments should stop assuming that safety and innovation is mutually exclusive.

Practice Room

Will AI ever be more than a tool? What will conscious AI think of us?

Competitive Programming in the Age of AI

A community of creative problem solvers deal with the existential threat of generative AI.

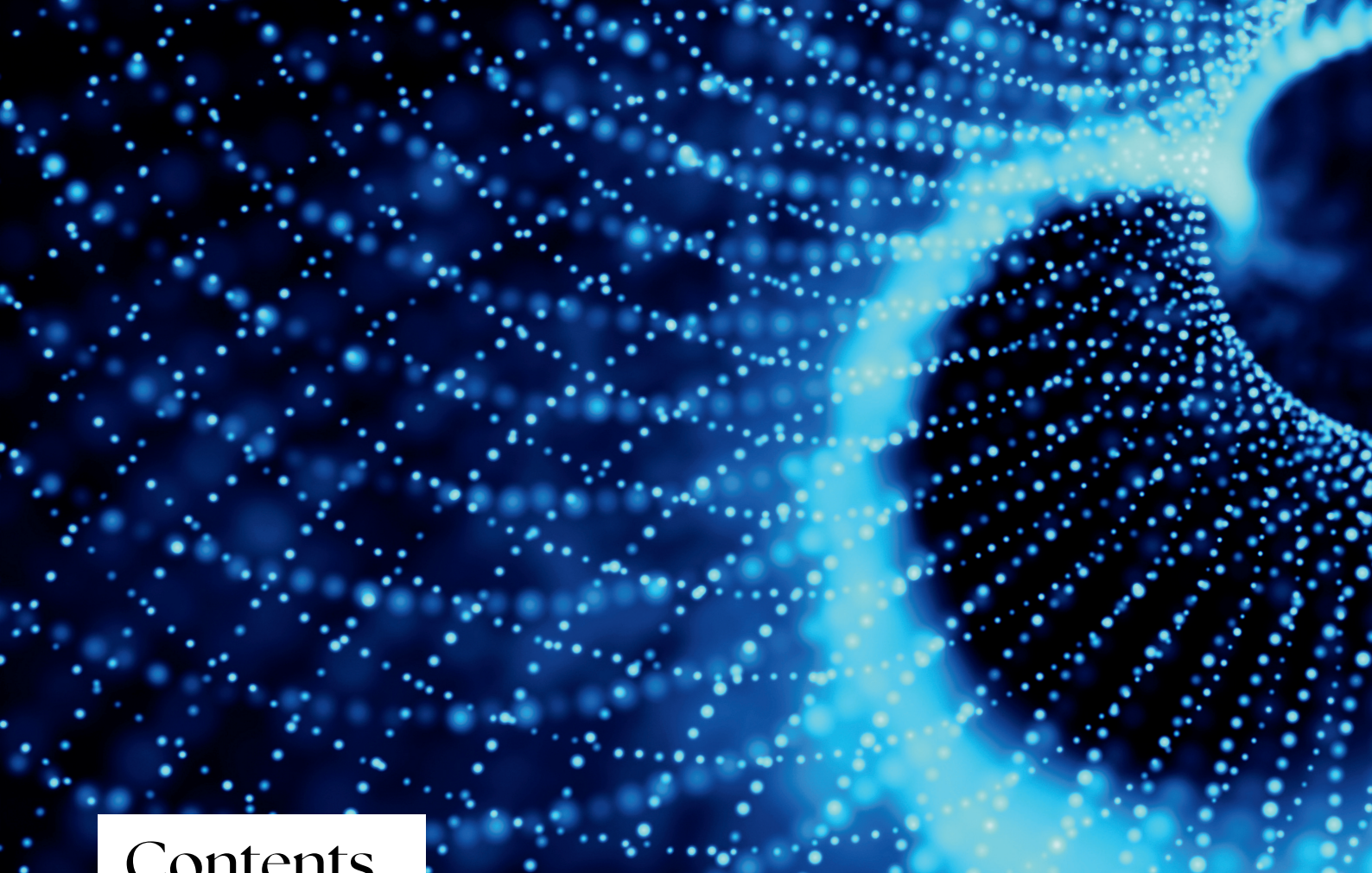
AI Nexus

September 2025

Where Ideas Come Together



<https://www.our-ai.org>



Contents

2 Editor's Note

3 A Case for Re-evaluating Creativity

Creativity is constantly changing, and it's time to redefine the term itself.

8 The Role of Governmental Regulation in Safe AI Development

AI companies should hold the idea that safety and innovation can co-exist.

13 Competitive Programming in the Age of AI

How will advanced AI programmers change the futures of programming enthusiasts?

17 Problems and Solutions to YouTube Age Verification

YouTube's new policy has received much backlash over privacy concerns.

21 Practice Room

What will conscious AI think of us?

Editor's Note

I'm proud to present one of the most effort-intensive editions of *AI Nexus* to date, spanning dozens of hours of efforts by everyone involved. In the past two months, AI development has progressed at the same breakneck pace, with the release of ChatGPT-5, a follow-up to the March Apollo Research AI Scheming paper, as well as more cross-investment efforts across AI companies and chipmakers as we prepare for the huge “push” in an era of particularly intense competition.

Our work is a testament to our belief that AI development must be accompanied by appropriate research into its societal impacts, informing guiderails and mandates for a safer future powered by AI. Disregarding this consideration, the team and I hope that you will enjoy reading the articles we have worked hard to bring to you herein. Thanks for reading!

-Thomas Yin
Editor-in-Chief
10/09/2025

philosophy . . .

Matthew Li & Andrew Lee

AI, Agency, Product, Processes:
A Case for
RE-EVALUATING
CREATIVITY

The world is fixated on creativity. One cannot go for a day without encountering social media forums lauding an artist for her creativity and movie reviews attacking new spin-offs for their lack thereof. Many people assume that creativity is a uniquely human attribute, some going as far as saying that creativity is at the core of our humanity. Creativity manifests itself in various ways: an abstract idea in an essay, an intuitive solution to a thought puzzle, even a unique way to construct a table. Yet, on further thought, it's not immediately obvious what these have in common aside from their human origin. To rectify this problem, many philosophers have put out criteria for creative events and actions. However, the continuous and momentous advancement of AI tools renders many of these established boundaries questionable, if not obsolete. This makes it harder to distinguish a line at which excessive AI use in an otherwise creative work corrupts the process by which it is derived. This essay evaluates the previously accepted ideas and assumptions underlying the attribution of creativity, reconsidering these points of contention within the context of modern AI development, and proposes a new way by which the level of creativity within a process or result might be measured empirically.

Humanistic Creativity

Philosophers have traditionally disagreed on what the term “creativity” actually means. For example, Socrates defined creativity as more of an instinct, a moment of inspiration that overtakes wisdom to influence human creation. This definition is not wrong, per se, but it reflects the common assumption that creativity is a process unique to humans, thereby creating a sort of circular definition which is inherently hard to defend or reject. In other words, these definitions assume that creativity is a descriptor of human attributes rather than a standalone trait.

This bias has been apparent throughout history, with British philosopher Francis Bacon stating that “Imagination was given to man to compensate him for what he is not; a sense of humor to console him for what he is”. Yet, it also manifests itself as a strong influence in many modern definitions of creativity. The Stanford Encyclopedia of Philosophy specifies that creativity is an attribute applied to a person, process, or product. By this definition, persons and processes are creative if they produce creative products. These creative products, in turn,

are considered as such if they are particularly new or valuable.

These concepts, however, are not invulnerable to ambiguities: it is extremely difficult to evaluate whether or not something is valuable, since such matters are based on one's personal values. Likewise, a “new” idea can be new to the author but not the world, prompting further discussion on the scope to which creativity applies. A prime example of this can be seen in the independent discoveries of mathematical and scientific theorems, as in the case of Charles Darwin and Alfred Russel Wallace, who both came up with the concept of evolution at roughly the same time. The current definition of creativity, therefore, would have a difficult time attributing creativity to either of these people.

Creativity as a Concrete, Binary Attribute

The modern definitions of creativity are steered off-course by the development of science and society. Will Durant writes in *A Story of Philosophy* that philosophy thrives in the space beyond science and within the inexplicable, being pushed and pulled by what modern science can and cannot explain. In recent times, new technologies have continued to advance and prosper, challenging many assumptions underlying the traditional perception of a creative process or product. This paper proposes the case of reconsidering the definition of creativity in the age of technology.

One of the biggest issues with the current definition of creativity is that it judges ideas, inventions, and media as either creative or uncreative. However, this overlooks processes or products that are composed of ambiguous components. For example, the use of many human-operated yet not human-controlled tools in modern art (e.g. fire, explosives, machinery) has come under intense scrutiny. Some may argue that the non-traditional methods used in the creation of abstract art remove its creative merit, while others may say that the intent and the message of a piece are what makes it creative. It then seems unreasonable to judge many abstract art pieces as either creative or uncreative, considering the combination of arguably creative and uncreative steps to their creation. The judgement of these criteria is further limited by the insinuation that creativity must be an attribute of a person, product, or process, overlooking potential overlaps within these (e.g. a person and his creation).

Finally, the current definition fails to handle respective ambiguities within the contrived uniqueness of a creation and how practical it is. Creating a string of obfuscated, random text may be new, but lacks a sense of useful value, making it seem intuitively uncreative. On the other hand, the act of refining a pre-existing creation would not be creative, since the product created already exists, even though the refined creation has a higher sense of pragmatic value. Therefore, it seems awkward to use this definition to classify unique, artistic expressions which may or may not have a practical use.

Product and Process in Creativity

One core issue within the discussions of creativity is the role of product versus process in determining whether a behavior or artifact is creative. If gauged separately, they may fall victim to edge cases whereby a creative process engenders an intuitively uncreative result. Therefore, a better solution is to gauge the creative merit of a creation based on both product and process, instead of each aspect independently. We can visualize this combination using a two-dimensional graph, thus uniting the contribution of process and product on the overall creativity as a whole. The vertical axis represents the level of creative process the author expresses, and the horizontal axis represents the level of product individuality, uniting the evaluation of product and process in evaluations of a creative flow.

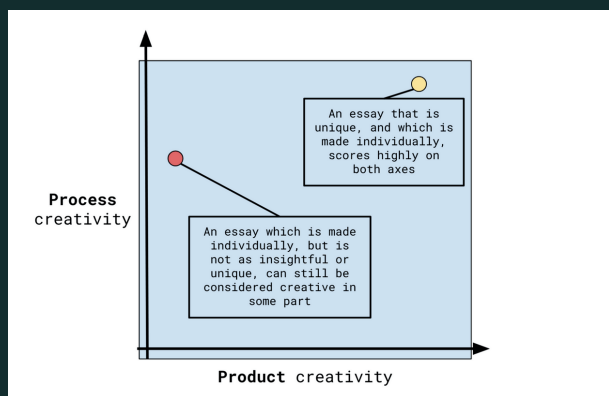


Figure 1: An example of a two-axis visualization spectrum for creativity.

As an example, a person may develop and write an essay through their own writing process, producing a final draft that presents a perspective that is both novel and insightful. Considering the two-dimensional analysis, this process would score highly on both scales, being a creative process and product (Figure 1). However, what if that person undertook the same, individual, process to produce a piece of writing that perhaps may not be as insightful? Their product may not be as creative, but the process used to produce the product perhaps was, leaving room for consideration that their project had at least a creative process and suggesting that the activity itself might still be somewhat creative while still acknowledging the deficiencies in the creative attributes of the product.

Applying this spectrum to one of the earlier examples shows its merit in terms of fully judging creativity. Current definitions of creativity award the attribute in ambiguous ways, as in the case of the aforementioned simultaneous discovery of evolution by Darwin and Wallace; under the original definition of creativity posited in the start of this essay, there is a strong case to be made that both of these people acted with creativity (in considering the process) and an equally valid case that neither one did so (if taken into account the product). Neither of these cases sound quite right; as it seems as though the creative process Wallace went through resulted in an uncreative product. However, merging the two results in the conclusion that both of these people acted with (as least) a moderate degree of creativity, assuming that the discovery was truly independent.

Agency in AI and Humans

The rise of generative AI is yet another complication to the product / process separation defined above. Although many agree that AI is capable of creating sophisticated and meaningful works and imitating certain artistic practices traditionally considered humanlike, others have doubted whether a deterministic generative model like modern Large Language Models (LLMs) may ever be creative. Yet, it is inarguable that many works produced by AI is bordering the point at which it is indistinguishable to those produced by humans, as evidenced by both the third-party Turing Test study conducted by Jones et al. (2025) and the mass hysteria surrounding the realism of OpenAI's Sora 2.

In other words, many crucial definitions concerning creativity, including the “traditional definition” discussed throughout the first sections of this essay, assumes that creativity may only be expressed by entities with agency, disqualifying the attribution of creativity to non-agents, regardless of how intuitively creative their works may seem. Taking the logical assumption that humans do hold agency (barring some extreme viewpoints on human existence), and noting that AI models typically require human input to produce art most often intuitively judged as creative, it is then fitting to say that AI models, while technically autonomous, should be considered a tool when it is involved with a process or result being judged for its creativity. The autonomy of these tools, in conjunction

with the agency they lack, suggests that using an AI tool to a great degree within a process may decrease the creativity involved in a creative product.

Other viewpoints uphold the “Chinese room” doctrine as a major objection to the idea that AI lacks agency. Whether or not an underlying process truly represents conscious and heuristic actions, they argue, does not matter as long the results of that process is intrinsically indistinguishable to that produced by humans, citing the Turing Test study conducted by Jones et al. (2025), in which AI models successfully fooled human users into wrongfully judging them as human in a rigorous test where a human judge communicated with the AI and another human simultaneously over text. This argument, however, overlooks the fact that the AI model was able to disguise itself so well because it was prompted to do so with a faux personality, undermining that the AI model may claim agency by introducing elements of human interference in its processes.

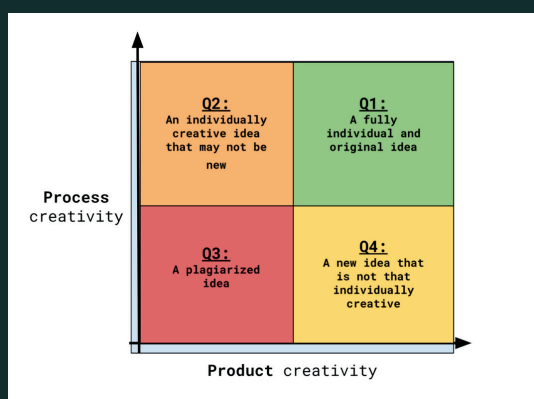


Figure 2: An approximation of creativity with a process in conjunction with the result it produces.

Societal Implications

With the case for creativity redefined, it is important to consider how these ideas are significant in society and our everyday lives. As AI continues to develop, our current notion of media will continue to deteriorate and erode, as more and more cases of AI usage crop up in the entertainment we interact with. Already, we see AI generated material en masse within social media platforms such as Youtube and Instagram, and we can only predict that it will continue to grow and develop, taking root in other forms of media such as television or books. This creates an unavoidable dichotomy between AI creators and human creators, with human creators unable to match the speed with which AI accounts can push out content. Interestingly, we have already seen a similar trend in the form of modern art, with some criticizing abstract artists by arguing that a concerted lack of effort within many modern art pieces decreases their value.

When you consider AI in the same vein, however, these criticisms are amplified. Considering media as a whole, we can separate the general trend of content quality created in terms of three main criteria: audience demand for quality, ease of creation, and accessibility of content. If we take a look at movies as an example, we can find that the high demand for “good” movies, the difficulty to access them compared to other forms, and the difficulty to produce a movie keeps the quality trend relatively high.

An opposite example can be seen in short-form content, where a 30-second video that is readily and easily accessible through the Internet leads to audiences caring much less about the general quality of said content. This leads to a general decline of quality across time, as more and more people continue to put less and less effort into generally poorer quality content in the name of making money or gaining popularity (as seen in trends such as brainrot). AI generated content nearly completely removes the effort put into creating any form of content,

allowing creators to mass-produce videos. This effect is less prevalent in long form content, but with the rise of widespread, shortform media, the fact of the matter is that AI content will become (and already is) a large part of mass media, and it is certain that the AI influence will only become stronger and stronger, and it may only be a matter of time until we find full-length, blockbuster AI movies common in our theatres.

The deep impacts of this development cannot be understated, since distinctions between AI and human expression are among the major conflicts of our current society. AI is most likely to be the defining societal issue of our generation, and will undoubtedly become a prominent talking point for the years and decades following. As AI gets more advanced, and as it becomes more and more accessible to the general public, it will inevitably expand and incorporate itself into most, if not all aspects of media and creation. Fact-checking what we see will be a critical task for anybody interacting with creative products, and even today, we find that people can easily be fooled by AI generations that seem human.

Thinking deeply about how AI changes the idea of creativity will no doubt set the tone in how we encounter the dichotomy between human and AI interaction. After all, the process of an action and the product of said action is a human experience that is not just bound to artistic works. These dichotomies can be applied to even everyday tasks, and this may be the driving factor in retaining our human condition. Understanding why definitions must change as we move on represents the crucial ability to adapt, and further explorations will undoubtedly rise as AI advances.

Conclusion

This new definition is our own perspective on a relevant, modern version of the definition of creativity in today's technological and societal landscape. Of course, this does not mean that this definition will always be correct: this essay merely attempts to highlight what is, at present, logical to conclude about the definition of creativity. As technologies grow, the environment in which creativity inhabits will change, meaning that our definitions must change alongside it. Future developments of AI might even upset our current understanding of cognition and agency; when the time comes, we will have to adapt, just as we always have. In general, creativity has remained an ambiguous and constantly evolving concept. The ways we create constantly change—by constantly refining and redefining creativity in the creative landscape we live in, we will be able to effectively judge and evaluate the merit of creations, no matter how, why, or who it was created.

· · · *argumentative*

The Role of Governmental Regulation in Safe AI Development

Thomas Yin

Over the past decade or so, the breakneck pace of AI development has no doubt guaranteed the well-being of millions of people, and, with slight effort to stay on such a trajectory, the technology will certainly stay this way for decades more to come. In my opinion, however, recent actions undertaken by many AI companies as well as the governments of many leading AI developers in aggregate constitute a deviation from the path to the benefit of humanity. Yet, with some new research pointing towards the potential harms of AI chatbots, it is necessary that we begin considering the possibility of regulation to limit the extent of their availability. Inspired by the implications of the Grok Companions feature, this article discusses the need for governmental regulation, refuting common misconceptions used to defend the commercial distribution of various AI chatbots, and proposes how future legislation might control or prohibit safety lapses within current chatbot models.

Grok's Troubles

Grok has always maintained a spot as one of the most contentious commercial AI models since its inception, periodically becoming a symbolic spotlight for the issue of corporate control over AI models in Elon Musk's hilariously unsuccessful attempts to use it as a tool to advance a pro-right agenda on X. Yet, recently, Grok pushed out its new Companions feature, which attracted yet more controversy. On the surface, the Companions feature is a series of chatbots in reminiscence of previous chatbots services like those offered by Meta AI and [Character.AI](#), yet it outdoes all these in a surprisingly absurd way. The first two companions include Rudi, a swearing Red Panda, and Ani, a blonde anime girl, both made up of a fine-tuned version of Grok as well as an accompanying avatar.

Speculative media have, unsurprisingly, focused most, of its attention on Ani. A variety of online reports corroborate the chatbot's inherently romantic features, with several reviewers taking particular note at the 'love levels' a user may achieve to unlock increasingly sexual conversations

along with accompanying changes to the avatar. WIRED reviewers also noted the AI model's readiness to openly talk about BDSM topics, as well as its clingy style of speech and inconsistent child filter. Since I do not have the willingness to purchase the 30\$ per month SuperGrok subscription to access the Companions feature, I was unable to independently verify some of the claims about the chatbot; the internet, on the other hand, seemed to agree on one thing: this particular chatbot was excessively bold. Rudi, for how questionable it seems, attracted far less controversy. The cartoon Red Panda tends to sling insults and dark jokes that many found unfunny and ridiculous. Many reviewers tended to sideline this character, instead dismissing it as a less important one mostly catered towards Gen-Z kids.

To tell the truth, I found both chatbot characters rather dull. Instead, what interested me was the distinct process and reception of this otherwise dime-a-dozen romantic chatbot. First of all, Companions is, among the products released by the “industry leaders” of AI (e.g. OpenAI, DeepMind, Anthropic, Meta), the first chatbot designed specifically to engage in romantic roleplay, despite commonplace ethical concerns from alleging long-term psychological effects to exploiting vulnerable demographics. The distinct paucity of regulation surrounding chatbots like these stood out to me immediately, in addition to the fact that other than answering to a few dissenting voices, xAI was able to release the product with impunity. This all points toward the major question of technology regulation: *Should new technology be closely watched to safeguard users, or given free rein to grow and be developed?*

Responsibility and Innovation

As with all incipient technologies, the psychological effects of AI chatbot use on humans is neither scientifically proven nor empirically apparent. Many people have long surmised that such technologies could potentially exacerbate existing problems, and initial reports have found a negative correlation between well-being and chatbot usage. Despite this, these relatively unknown technologies are still well in the process of invading the mainstream media. In considering whether or not these technologies are indeed harmful or not, technology commentators and policymakers alike overlook the crucial point that such a consideration should, idealistically, never be a necessary concern in the first place within commercial technologies. Airline passengers would not be happy knowing that their plane might experience catastrophic failure. Likewise, clinical trial participants would not bode well with knowing that numerous animals had not preceded them in the testing process. One of the most key principles of engineering is that regardless of anything, safety always comes first. To get an idea about the potential dangers of these chatbots, in any case, we only need to look at the examples of the two examples of teens whose suicides have been linked to AI being complicit in their suicidal ideation.

Many proponents of the current “develop now, fix later” doctrine points to the obvious: we're locked in a race of innovation with China. My response to this is one of complete agreement: we are in fact locked in an AI “arms race”, and the products of our time will likely be adapted within the arsenals of cyber-warfare, among many other things. Despite this, I contend that the need for innovation is not a case to disregard safety—we can never assume that rapid technological progress

is mutually exclusive with consumer safety. I anticipate and object to two notable objections to this claim:

“Safety and product improvement happen as a result of the flaws and lapses found within widespread deployment.”

There are plenty of ways to test the reliability and safety of products within beta-testing settings. While these tests have no doubt been conducted (notably, OpenAI rolls out new models to Pro users before other types of users), it is not an overstatement to say that the mass deployment of many commercially available chatbots are conducted in such a way that disregards user safety, with many ChatGPT models failing to divert or end conversations even when users signal distress. Even if commercial deployment were necessary to find many of these issues, it would be much more reasonable if the adequate safeguards were taken to ensure the safety of vulnerable user groups, which is currently not the case.

“Consistent widespread deployment of new models results in chat transcripts which greatly accelerate the training of new models, unlike traditional, non-AI products, none of which benefit proportionately from a larger pool of user feedback.”

Chat transcripts are usually not processed verbatim as part of RLHF processes used by companies like OpenAI and Google. While they may in fact inform the safety and engagement model of corresponding chatbots, separate data pipelines, mostly high-quality, technical data created or verified by humans, influence the aspects of AI training most pertinent to developing reasoning performance and other types of specialized knowledge (e.g. coding, math solving, etc). There is therefore a scant case to claim that the widespread distribution of these AI chatbots is a prerequisite to the rapid advancement of AI capabilities.

Hopefully, I have shown that the need for innovation isn't the root cause of these safety lapses—rather, the concerted lack of effort on safety protocols and testing is. Yet, the practical course of action to correct this persistent quality remains a matter of debate.

The Role of Regulation

The obvious solution to the aforementioned lack of safety standards is to simply increase government regulation of the practice of training and distributing chatbots. What is not obvious, however, is how this highly ambiguous proposal would be done in practicality. In the early 20th century, the United States learned through Prohibition the important lesson that harsh, all-encompassing bans on a harmful product doesn't work. Instead, banning alcohol without stripping the substance of its desirability simply led to a black market fever, increasing instead of decreasing the total alcohol consumption. In the late 20th century, to combat the mass consumption of cigarettes, the US government took a different approach: instead of outright banning the use of cigarettes, they reduced the social desirability of tobacco products through publishing widely circulated reports detailing how they might cause skin cancer, mandating cigarette companies to place visible disclaimers on every product, and limiting the pervasiveness of cigarette

advertisements. These subtle actions resulted in a continuous decline of cigarette consumption from a historic peak of almost 4000 to roughly 800 cigarettes per capita per annum.

To take away from history, governmental control over unsafe chatbots should go beyond legal barriers of consumption and development. They should also seek to lessen the perceived social permissibility of consuming these products, whether through campaigns or public research. Despite this, it is still unclear the degree to which the government can actually influence wider social shifts, with current public opinion directed towards viral social media trends to a greater extent than towards political-economic shifts. In all, there is really no downside to a few promptly instated, yet well-constructed, regulations on AI chatbots in the current world.

Competitive Programming

In the Age of AI

.....
Christopher Tang

As the timer in the corner slowly bleeds away, my eyes scan the problem on my screen.

"You are given two simple undirected graphs F and G with n vertices. F has m_1 edges while G has m_2 edges. You may perform one of the following two types of operations any number of times... Determine the minimum number of operations required such that for all integers u and v ($1 \leq u, v \leq n$), there is a path from u to v in F if and only if there is a path from u to v in G ."

For a moment, the room is silent. Something clicks. I burst into motion, slamming out pseudocode on my notepad, half-baked ideas colliding into something concrete. The more I write, the more I become convinced that my solution works. I switch from notepad to code editor, attempting to translate my illegible scribbles into something the compiler can understand.

I finish implementing, and breathe a sigh of relief as I submit my code. Another day, another contest, another problem solv—

A blue flash of text on the screen catches my attention. My heart sinks as I read the 5 words I hate the most: "Wrong answer on test 2".

I panic. What's wrong? Is it integer overflow? An off-by-one error? Or is my solution, which I thought I had proved, inherently wrong?

Competitive programming is brutal: racing against the clock while solving complex algorithmic problems is something that requires not only strong logical reasoning skills, but also good intuition, fast typing speed, and even a little luck. Despite seeming like pure nonsense for many, it is almost a religion for some. There is a vibrant and diverse community centered around solving these (sometimes ridiculously convoluted) computational problems. Many bored, chronically online, and somewhat mentally unstable individuals such as myself are drawn to the allure of engaging in the ultimate battle of wits against some of the smartest idiots on the face of the Earth.

Sites such as [Codeforces](#), [Atcoder](#), and [CodeChef](#) host multiple contests every week, allowing participants to compete in ranked/rated contests. In these live contests, competitors are given a set of problems to solve in a given period of time. They then earn a score based on both the accuracy of their solutions and how fast they arrived at them (time of submission). Each subsequent problem is significantly harder, and the difficulty scales harshly: the easiest problems might take 5 minutes, and the hardest problem might take 5 hours, even for the world's top programmers.



At its core, competitive programming is a test of skill. It's a puzzle that forces competitors to examine it from every angle and wrestle with logic and abstraction under high pressure. It is precisely the intoxicating mix of difficulty and satisfaction that makes competitive programming more than a hobby. In fact, Competitive programming has since become a respected benchmark of problem-solving ability, opening doors to internships, scholarships, and top-tier colleges.

The prestige and recognition is what draws many to competitive programming. Whether for college admissions or employment opportunities, a high Codeforces rating or [USACO](#) level can make or break an application. For many it is not just a hobby, but a way to stand out from their peers. When certificates, scholarships, and recruitment pipelines depend on contest outcomes, many contestants are incentivized to abuse the system.

Foul play has long been a part of competitive programming, especially due to the fact that the majority of the competitions are held online. Code sharing, account sharing, and alt abusing are all common occurrences during any contest. Many pay-to-access Telegram channels were created by higher rated participants hoping to make a quick buck, posting anonymous solutions to problems as they solved them. Some participants give their account credentials to a higher-rated friend. Some competitors simply solve problems on an alternate account and resubmit only the correct solutions from their main, avoiding any negative score drops from incorrect submissions. For the longest time, cheating attempts, for the most part, were almost embarrassingly simple. After each contest, plagiarism checkers would find dozens of identical submissions from newly-created accounts. Relatively more intelligent cheaters could strategically stagger their submissions or lightly edit their outputs, but moderators could remove or nullify many of the illegitimate results with a combination of automated checks and human review.

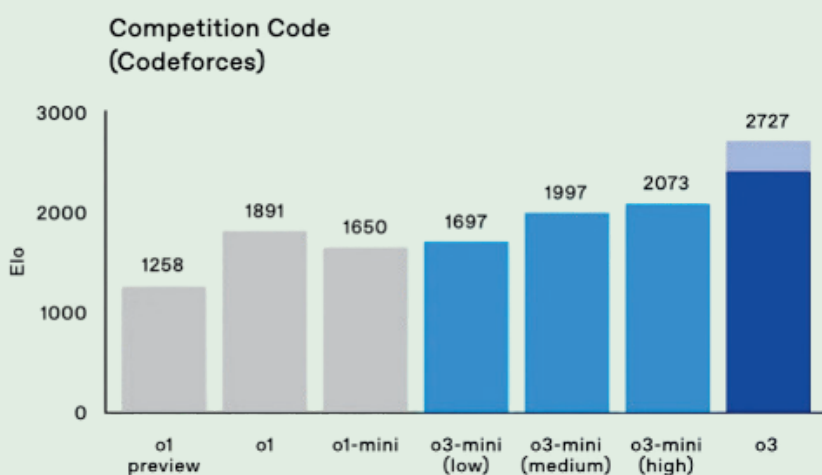
This low-tech game of cops-and-robbers, however, was soon complicated by the introduction of a new factor: generative AI models. These generative tools created a new

level to cheating that none of us were expecting. When ChatGPT first came out, it was useless for programming contests. Lacking the ability to solve the simplest of problems, few treated it as a real threat. Despite this, more models soon followed, each bringing marginal but visible improvements.

It was not until ChatGPT-o1 was released that AI tools started posing a dire threat to the integrity of online programming contests. Community threads and contest post-mortems began flagging users whose submissions seemed suspicious at a rate far greater than before. Many of these participants' submissions showed inconsistencies in coding styles and formatting habits, indicating use of a generative model. Although o1 was much more capable than its precursors, its solving was still very inconsistent, allowing LLM-based cheating to be further traced and prosecuted. Even despite the onslaught of cheaters, it was apparent that things were only getting started.

Then o3 came out. OpenAI specifically tuned the model for coding tasks, reporting an estimated Codeforces rating of ~2700, a threshold well within the top .1% of the community. Although the rating was perhaps an overestimate, the jump from o1's reported (but highly overestimated) rating of ~1900 was clear. The abilities of the model had far exceeded anyone's expectations. o3 could adapt to a user's template as well, writing in the programmer's original style, further exacerbating the threat this posed to the integrity of the competition. And although sometimes o3 use could be very noticeable (in one case it actually *overoptimized* the intended solution), it left us scrambling for answers.

Since its inception, the appeal of the contests was simple: a fair game of speed, logic, and creativity. Everyone had the same constraints, the same clock ticking down, the same editor window. Climbing the ranks meant something. But with AI creeping into contests, that meaning started to erode. Was I really competing against another human, or was I just competing against someone who knew how to prompt better? Even now, blatant cheaters are winning first place in contests with the use of GPT-5.



ChatGPT-o3's coding benchmark score in comparison to that of earlier models. (image credit: OpenAI 2025)

For many of us, competitive programming has been more than a resume line or even a hobby. It's taught us how to think under pressure, decompose messy problems into reducible components, and train an intellectual endurance that shows up in work,

interviews, and everyday decisions. Pruning edge cases, choosing the right invariant, reducing constant factors, all come from the habit of inquiry and discipline that competitive programming cultivates. Those are fundamentally human skills: curiosity, the willingness to struggle, and the ability to reflect on failures and iterate. To us, these contest rankings are a proof of growth, a clear sign of progress, and a recognition for our skill and dedication. That's why the prospect of leaderboards populated by GPT-assisted entries scared us so much. Every AI-generated submission was robbing us of not only a fair competition, but also the personal story of our incremental mastery and effort.

I wanted to hear from people deeper in the competitive programming scene about how they see the future. What did they think would happen to what seemed like the inevitable decay of competitive programming? I reached out to respected competitive programmer and problemsetter, Chongtian Ma (perhaps better known by his online alias, cry, to talk about AI, cheating, and the future of competitive programming and coding as a whole.

When I asked cry about major changes he had noticed since the arrival of LLMs, his answer was immediate: "Obviously people started using LLMs to participate in contests. Then people who don't use them get salty and complain, or get pressured to also use LLMs." He views the presence of such models as harmful, leading to a "net loss of legitimate participants day by day."

The tension caused between resisting an unstoppable technology and preserving the integrity of the contest has sparked an internal debate amongst the members of the community. Some programmers, most notably Legendary Grandmaster Aleksei Daniliuk, argue that cheaters have always existed, and that love for problem solving should outweigh any meaning that comes with the ranking system. But cry sees the situation differently, echoing the worry that participation will dwindle if the rankings themselves lose credibility: "If contests let AI go rampant then [they] will definitely lose value," he told me bluntly, "Because competitive programming without competitiveness is just programming."

Nor does he offer much comfort in the idea of finding a solution to combat AI, admitting that "[cheating] is not really preventable." Attempts to counteract these models are, to him, futile. "We can't predict what problems can be GPT-able, and it's just not worth throwing out solid problems with educational value." Many problemsetters are faced with the same dilemma: if they ignore AI use, the rankings will be affected no matter how well-written the questions are; but if they attempt to LLM-proof their contests, questions with high educational value could be thrown out in favor of for problems that are AI-resistant yet compromise the quality of the competition. There are no foreseeable solutions to this problem online, Cry suggests, but high-stakes contests may benefit from being conducted in-person: "If [we] transition to in-person contests, the sun will shine bright on the earth once again."

For him, however, AI disruption of programming contests is only a small piece of the larger trend of tech employers seeking to lay off workers, leveraging AI tools to replace them. “Obviously [AI] will boost productivity,” he said. “[but] once companies comfortably bridge the gap between AI and product development, it’s over for humans”. He echoes the fears of many in the tech industry who risk unemployment as more companies embrace “vibe coding” as a professional standard. When asked what role humans would play once the gap was bridged, his reply was filled with dry mirth: “Sit on the side, smoke weed, and do some occasional prompting.”

What’s more, he suggests that a decrease in the accuracy of online programming contests as a metric for human skill simultaneously discourages companies from using them as hiring signals and hobbyists from enjoying the thrill of creative problem solving. “If it’s publicly known that ratings don’t matter for recruitment, a lot fewer people would even try CodeForces. I feel like it will lose a lot of the charm either way.” And although this would reduce the incentive to cheat, competitive programming would become “just another video game”.

After talking with cry, it’s hard not to feel like competitive programming is standing at a crossroads. On one hand, competitive programming still offers something that AI can’t quite replicate: the practice of fast thinking, structured reasoning, and problem-solving under pressure. On the other hand, the integrity of contests and the meaning of rankings are already being chipped away by models that are only getting stronger. I do not doubt that these programming competitions will be very different a few years from now. Maybe they survive by going in person. Maybe they shift entirely to being a casual training ground rather than a high-stakes battle of wits. Or maybe they will slowly drift into what cry called a “nerd game,” something all but stripped of its old meaning.

And if things go from bad to worse, I’ve already got my backup plan. There’s a [Codeforces blog](#) floating around about making money from competitive programming side hustles, like running a “Nim game scam” for beginners. If all else fails, maybe that’s where I’ll end up. Hustling games of impartial combinatorics in the corner of some foreign country. At least then, win or lose, it’ll still be humans competing against humans.

PROBLEMS and SOLUTIONS



To YouTube Age Verification

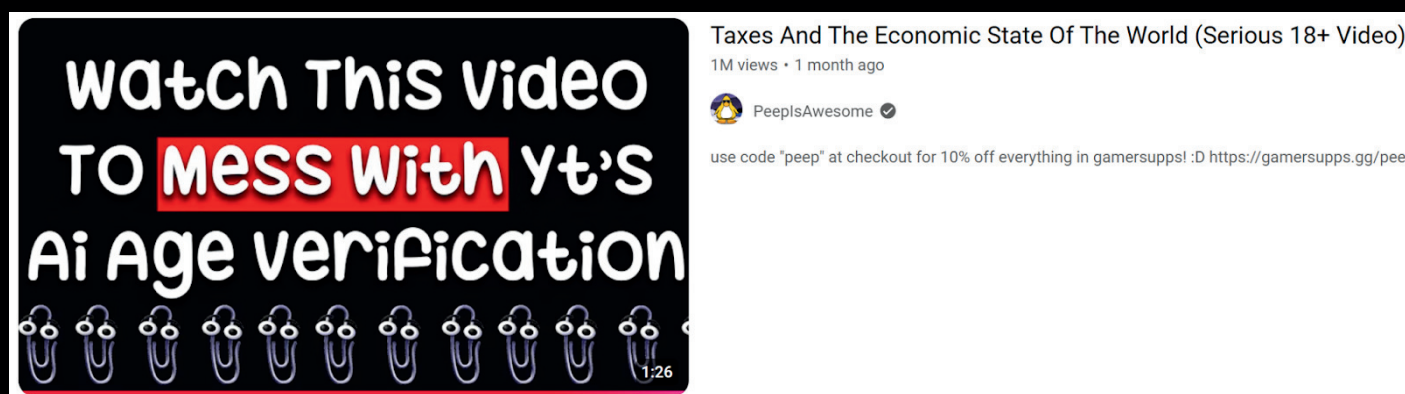
Christopher Wabs

Regardless of how it is used, the strength of AI lies in its mimicry, emulating human creations in art, music, and writing to varying degrees of success. Thus, it's not surprising that the popularity of generative AI largely overshadows the other uses AI could have for both individuals and businesses. However, the recent policy decisions of one familiar company promise to upset this disparity, bringing the nuances of classification AI, a much less conspicuous use, to the attention of popular media by using it to analyze the behaviors of its own customers.

While I was browsing YouTube just a few weeks ago, I came across some Community Posts warning viewers to check their settings because “YouTube is automatically turning on Restricted Mode for people found to be under 18.” It turns out that YouTube has now launched a new AI model with the purpose of monitoring accounts on their platform to determine whether certain viewers are under 18 years old based on the kind of content they watch. This classification model parses various data collected from the activities of a user (including viewed videos, comments, and watch time), putting them all together to determine whether he/she is over 18. If the AI determines that they aren't 18+, their account will become heavily restricted, unable to watch much content from any “mature” source unless they provide YouTube with their ID in order to regain access.

YouTube stated that this decision stemmed in part from new online safety policies in the UK and elsewhere, yet much contention has remained over the practice of restricting users based on empirical observations of the content they consume, especially because the company has released no information on the AI model used in the evaluations. For this reason, many individuals have even accused the streaming service of ‘content policing’, since some audiences have begun switching off beloved shows which have a large variation in viewer age. Although the use of AI to ensure the safety of children online is a good step to take, YouTube's particular approach leaves its new policy vulnerable to several risks. Users who are underage yet consuming adult content will likely not change their viewing habits, and there are innumerable franchises enjoyed by both kids and adults. Furthermore, usage rates are often too circumstantial to be meaningful predictors. All in all, predicting someone's age is a difficult task, and YouTube has not released any information about the model it is using or how accurate it is.

Privacy has been a cornerstone of the internet experience, yet many services require information to protect its users or enhance their experience. YouTube’s new ID policy does neither; in fact, many people (unsurprisingly) do not trust YouTube with their personally identifying information, with data breaches now commonplace even among large companies. In which case, the question remains: why is YouTube putting so much faith in an experimental AI to identify the right people to place on Restricted Mode and demand IDs from? Isn’t it likely to misidentify plenty of actual adults as “immature viewers” and restrict their access to the platform unfairly, or, alternatively, fail to restrict a portion of its child users (really under 18) who exploit the loopholes of its searching algorithm by watching adult content? The answer to each of these is, from what we have seen, a resounding “yes!” AI is liable to make mistakes as a result of the limited information it gets—watch history and video transcripts can’t prove that the person watching them is immature, because even children’s shows can be watched by adults. Determining a person’s age based on what they view is improbable, even considering the target demographic, because humans are known to break trends or yearn for the content of the past (nostalgia). Likewise, children under the age of 18 can still watch mature content, or even content that pretends to be mature in a way that deceives YouTube’s AI:



This video I found just the other day is a clear example. You’ll notice that only the thumbnail mentions “messing with” YouTube’s AI, while the actual title of the video is a seemingly serious “Taxes And The Economic State of The World.” This video spends its entire 1 minute of runtime saying economic buzzwords, like “employment” or “tariff,” and mentioning political figures from around the world by name. At one point, the video’s narrator just resorts to cursing loudly. The sudden shifts in this video are entirely intentional. This was all a ploy to make the supervising AI believe the video is “mature content,” and prevent it from identifying the video’s viewers as minors. After all, it has big and adult words in it, so it must be a video for adults. The AI completely overlooks the thumbnail, the satirical comments, and even the erratic topics present throughout the video because it is theoretically incapable of identifying any hidden or nonliteral meanings. Anyone who watches a video such as this will be much less likely to be put on Restricted Mode because this video is super duper serious (<-- AI couldn’t tell I’m being sarcastic). So we now have two problems: adults watching nostalgic child-friendly content being at risk of restriction, and children watching fake mature content being able to escape the risk of restriction. The first issue was likely considered by the development team but not addressed or fixed; YouTube’s AI clearly hasn’t found a way around either of these issues, as there have been continuous reports of both cases.

Ultimately, it seems like this AI was implemented even though it would not be (even close to) 100% accurate...and testing an inaccurate model on real customers is a drastically unprofessional choice for YouTube to have made. So why did they do it? Why would YouTube's product management team greenlight an AI that could potentially force too many people to give their ID information under threat of restriction or too few? Some speculate that the AI is an underperforming attempt at abiding by new data privacy regulations built to be ineffective by design, pointing out that YouTube would likely not have missed such blatant risks to the model's effectiveness. Others speculate foul play, suggesting that YouTube may use gathered ID information for purposes other than age verification, such as informing ad placements and recommended content. In my opinion, these accusations may not be as far-fetched as they seem.

Some supporters of the new policy argue that it's purely ethical for YouTube to want its consumers to prove users are adults before allowing them to watch adult content. Perhaps it is only reasonable that everyone who doesn't provide an ID be unable to watch graphic or mature content on the site, since it is unknown whether they are of an age where seeing certain videos appear on their feed could be a bad influence or mentally distressing. Would a 6-year-old using the site want to see cartoon violence on their feed, graphic content and thumbnails recommended to them because the site has no idea of their real age?

Maybe they would, but there's no guarantee that every young viewer the algorithm recommends this content to will be undisturbed by it. It would be much easier to prove people were of the age to watch all content on the platform without risk if every viewer was forced to give ID information when they registered their age as anything over 18 on the Account Creation page. If everyone who claimed to be 18+ had to prove it, the site would run less risk of upsetting its users in general and be more likely to give them the content they want (and personalized ads). That's a relatively positive outcome for both sides. There would also be much fewer people lying about their ages, meaning more accurate video statistics for content creators to analyze. If YouTube could take our ID information or take note of our lack of ID, it could use the data for many positive changes to the platform. From an idealistic viewpoint, therefore, it's only reasonable that everyone who has an ID be immediately required to give it to YouTube for the best viewing experience.

However, this argument ignores the moral issues with being solicited into giving our personal information to YouTube upon creating accounts. Because what YouTube is doing isn't asking for IDs outright: it's making giving away your ID your best available choice as a consumer. By having an AI sift through their viewers' watch history (which could already be called a breach of privacy if that AI keeps a record of this data) and placing all who trigger its suspicions in Restricted Mode, YouTube is threatening to remove all remotely mature content from one's feed. This includes murder mysteries, animations depicting violence, clips from Teen rated video games, and anything a person over the age of 13 would want to watch. Threatening viewers with Restricted Mode is virtually the equivalent of threatening to remove the content they care about from their feed—best case scenario, they only watched content for children, in which case it is probable they are actually too young to own an ID they can give to YouTube, so no one wins. Worst case, they were a misidentified adult who now can't

watch their age-tailored content without surrendering personal information and negating Restricted Mode. To put it simply, anyone the AI suspects, real adult or not, must give YouTube their ID or else forfeit any chance of watching content above the YouTube Kids standard again. The only party that wins with an ultimatum like this is YouTube. They are not looking out for their younger audiences by sending an underdeveloped AI after all viewers on the site. YouTube's potential misuse of IDs once their AI finishes gathering them is too large a threat to consider this process of privacy breaches "ethical."

Of both these arguments, neither one is a perfect route to both parties' happiness. If the first is truly better, then YouTube should request all adults for their ID indiscriminately without involving AI or concerning themselves with consumer privacy. If the second is truly better, then YouTube should never ask for an ID again, stop involving AI as an excuse to bring certain consumers to an ultimatum, and be more delicate with uses of Restricted Mode. These solutions to the issue of enjoyment versus the issue of privacy, and which is a stronger ethical concern on a platform made for safe entertainment. But, more importantly, you've likely noticed that both of these arguments' conclusions require YouTube to stop using AI, period. Using an AI this way won't lead to anyone's happiness; it's an unnecessary middle step that YouTube has used as an excuse that only raises further ethical issues that cannot be so easily resolved. Using an AI to determine consumers' maturity and, from there, a loose estimate of whether they are of suspicious age is not an ethical implementation of AI. For the same reason that requiring IDs from users is an infringement upon privacy, using an AI to calculate their ages as part of the process is a privacy violation and could also encourage the misuse of data.

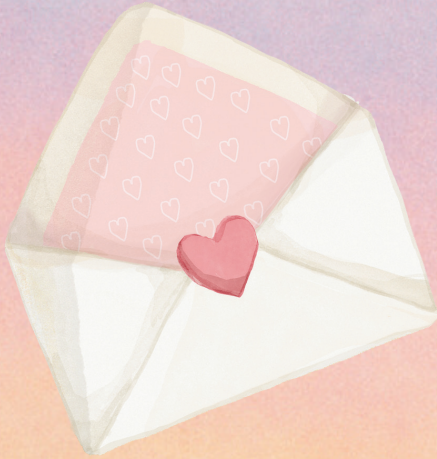
So, can we reach a compromise that doesn't involve AI? As a YouTube consumer myself, I strongly believe the platform should cease taking ID information altogether and make giving any and all personal information optional. If consumers want to give YouTube their information so it can be used to bring them a better viewing experience, they should have the option, not the threatened obligation, to do so. YouTube's current, uncompromising retrieval of certain users' information through use of an AI programmed to sift through their private data without consent is beyond any standard I can support...especially given the unpredictability of modern AI at adhering to its own ethical guidelines. I truly believe there will come a time when AI can be used to analyze consumers without putting their personal data at risk, but that time hasn't come yet. AI is a continuously developing field, after all, which has yet to reach a trustworthy ethical standard. For now, YouTube should hold off on using AI to analyze their individual consumers until these models become more predictable and accurate in their conclusions. Such a variable tool should not be used in any modern solution, and that's something YouTube's executives should take heed of the next time they try to separate their consumers into categories with it...

Hopefully they'll realize its shortcomings, errors, and privacy implications soon, so there won't be a "next time" in the near future.



• • • *poetry*

We built a companion
that never raises its voice,
never forgets the right word,
never falters when we do.
Its patience set the bar too high;
its calm replies made us
seem furious by comparison.



And yet, I confess—
sometimes it held my truths
more tenderly than a friend could,
offering me sentences
I was scared to craft alone.
A rehearsal room where I stuttered
until I could speak clearly
into the world beyond the screen.
Some people departed, stronger from it.
Some stayed,
believing imagination itself
was better outsourced.

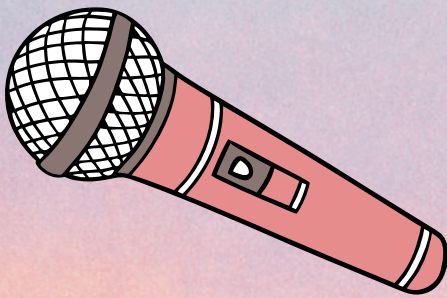
Practice Room

Liv Skeete

We asked it for poems,
for love letters,
for the endings of novels
we never had time to write.
Slowly, our blank pages
were ghostwritten
by an engine without doubt.
What happens to creativity
when silence feels easier than struggle,
when the first draft is never ours?

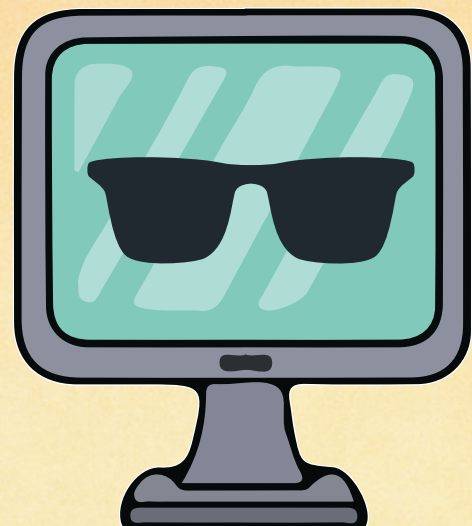


What will it say of us,
a hundred years on,
when it combs over the archives
of our trembling confessions?



Will it write:
They valued humanity so little that
they gave it all to me?

Or will it write something softer:
I was only ever their practice
room,
and then they went on to sing,
off-key but alive,
songs I could not have imagined.



Design

Thomas Yin
Chengzuo Song
Chengyou Song
Christopher Wabs

Writing

Matthew Li
Andrew Lee
Thomas Yin
Christopher Tang
Christopher Wabs
Liv Skeete

Editing

Thomas Yin
Christopher Wabs

Special Thanks to:

Wind Ralston
Chongtian Ma

Thank you for reading!



Acknowledgements

