

Poetry & digital personhood

by Carmine Starnino

When a book of brazenly surrealistic poetry and prose was published in 1984, attributed to a mysterious figure named “Racter,” it was hard to know what to make of it. *The Policeman’s Beard Is Half-Constructed* was a fever vision of weirdness. “I need electricity,” declared the poet in a signature moment. “I need it more than I need lamb or pork or lettuce or cucumber./ I need it for my dreams.” That same tone, at once charming and confounding, charged Racter’s aphorisms, limericks, fictional riffs, bits of dialogue, and odd attempts at nursery rhyme (“there once was a ghoulish sad snail”).

Reviews were mixed. Most conceded that nothing like *The Policeman’s Beard Is Half-Constructed* had ever been seen before. But Racter’s patter didn’t always impress. While the strange skips in logic gave off an idiosyncratic energy, the verse also made readers feel like they were eavesdropping on the rantings of a somniloquist. One critic called the 120-page collection “Metaphysical poetry as interpreted by William Burroughs and William Blake, with a dyspeptic dash of Rod McKuen and Kahlil Gibran thrown in.” Another critic insisted Racter’s inscrutable ingenuity revealed not a literary maverick but a “coffeehouse philosopher who knew a great deal once, but whose mind is somewhere else now.” With its bright-red cover, the volume attracted a cult following. Copies soon became scarce, which only added to Racter’s mystique.

That mystique wasn’t at all harmed by the fact that Racter didn’t exist. Not as an independent scribe, anyway. The entity responsible for

insights like “When my electrons and neutrons war, that is my thinking” or “A tree or shrub can grow and bloom. I am always the same. But I am clever” was actually a piece of code. Racter (short for raconteur) had been hatched on an early desktop computer programmed with the rules of English grammar. The algorithm could conjugate verbs, assign gender to pronouns, match adjectives with nouns, and discern singular from plural. With a vocabulary of several thousand words, Racter knew just enough to string together sentences randomly but coherently, at least from a grammatical standpoint. It had no awareness of the “syntax directives” steering those sentences and took no pleasure in their twists and turns. In fact, it was Racter’s developers who sorted through the copious amount of text their instrument churned out, compiling the most striking results for publication. Racter’s daring, improvisatory style was a ruse, a party trick, hocus pocus. It was also a coup. Computer scientists had been trying to coax machines to write verse since at least the 1960s, and Racter was a singular example of how something mindless could create something meaningful. Indeed, it led the avant-garde poet Christian Bök to wonder if humans were needed to produce literature at all. *The Policeman’s Beard Is Half-Constructed*, he argued, was an “obit for classic poets.” Awaiting us was an era of “robopoetics.”

And true enough, we are overrun with Racter’s kin. Dozens of websites, with names like Poetry Ninja or Bored Human, can now generate poems with a click of a key. One

tool is able to free-associate images and ideas from any word “donated” to it. Another uses GPS to learn your whereabouts and returns with a haiku incorporating local details and weather conditions (Montreal on December 8, 2021, at 9:32 a.m.: “Thinking of you/ Cold remains/ On Rue Cardinal”). Twitter teems with robot verse: a bot that mines the platform for tweets in iambic pentameter it then turns into rhyming couplets; a bot that blurts out Ashberyesque questions (“Why are coins kept in changes?”); a bot that constructs tiny odes to trending topics. Many of these poetry generators are DIY projects that operate on rented servers and follow pre-set instructions not unlike the fill-in-the-blanks algorithm that powered Racter. But in recent years, artificial-intelligence labs have unveiled automated bards that emulate, with sometimes eerie results, the more conscious, reflective aspects of the creative process. Microsoft’s “empathetic” AI system, Xiaoice, designed to explore emotion in language, has composed millions of impassioned poems in response to images submitted by users. Deep-speare, the brainchild of Australian and Canadian researchers, caused a stir when it taught itself to write Shakespearean sonnets.

These initiatives have now been dwarfed by Racter’s newest descendant. Released in 2020 by OpenAI, a San Francisco start-up, GPT-3 is an AI tool that was force-fed a vast portion of the internet (the entirety of English-language Wikipedia adds up to only a fraction of the billions of words ingested). Endowed with algorithms that help it make sense of all that data—“neural” algorithms modeled after the circuitry of the human brain—GPT-3 can produce, from a simple prompt, astoundingly human-like writing of any kind: recipes, actuarial reports, film scripts, real-estate descriptions, technical manuals. Of course, there is buzz for GPT-3’s poetic chops too. In one example, the American poet Andrew Brown asked the software to take the perspective of a cloud gazing down on two warring cities. GPT-3 delivered a rhyming poem that began, not uncharmingly, with “I think I’ll start to rain.” Stephen Marche, writing an article for *The New Yorker*,

assigned GPT-3 maybe the trippiest poem in the English canon: Samuel Taylor Coleridge’s fifty-four-line “Kubla Khan”—an opium dream interrupted in the middle of its composition in 1797 and never completed. GPT-3’s mission? Finish the fragment. What the program fantasized was so sophisticated (“The tumult ceased, the clouds were torn,/The moon resumed her solemn course”) that readers unfamiliar with the original poem might have had a hard time discerning where Coleridge ended and the computer began. With an estimated billion dollars in backing, GPT-3 isn’t a better Racter. It’s a godlike Racter. *Forbes* named it the AI “Person” of the Year. Anyone who believed AI to be “nothing like intelligence,” said one expert, “has to have had their faith shaken to see how far it has come.”

The extent to which journalists, academics, and computer developers are genuinely troubled by GPT-3’s believability suggests that AI’s holy-grail goal of passing the Turing Test—that is, building a machine that can persuade us it is thinking—might be getting too close for comfort. The test draws its name from Alan Turing, the British mathematician and wartime codebreaker who was interested in the links between computation and cognition. In 1950, he presented his colleagues with a challenge: if a device exhibits behaviors indistinguishable from what a human would do or say, what reason would we have for denying it the capacity for intelligence? The thought experiment, which Turing called “the imitation game,” has haunted AI research and inspired scores of competitions. The Neukom Institute for Computational Science, for example, ran a well-known contest in which AI teams tried to fool judges with the artistry of their machine-made submissions, from limericks to choral music, and, in 2014, an online game called *Bot or Not*—which asked users to decide whether a specific poem was replicant or real—caused a sensation. The Turing Test is now, in a sense, philosophical clickbait, and, like any clickbait, it’s better at raising uncomfortable questions about human consciousness than answering them. Are our aptitudes and abilities really unique, it asks, or are we just robots of a different kind? The test is also,

strange to say, a means of self-protection: as long as our creations can't imitate us, we feel safe. But GPT-3 represents a new threat level. In a series of controlled trials using poems devised by the AI system, University of Amsterdam researchers found that, half the time, GPT-3's eloquent put-up job prevailed. "People," they wrote, "are not reliably able to identify human versus algorithmic creative content."

It's tempting to think these incidents demystify human virtuosity: how replicable it is, how unspecial. That's why Bök loved Racter. He believed that, by automating the free play of language, it had rendered inspiration—and any trait once deemed beyond computation, such as poetic genius and originality—irrelevant. Many have found GPT-3 no less cruel in unmasking creativity as fundamentally algorithmic, a product of mere procedures. But more interesting than any Turing-esque reality check AI can inflict is what poetry can teach us about the limits of algorithms. At the moment, those limits are hard to see. AI composes symphonies and hit songs. It sculpts and choreographs. It dabbles in haute couture. It sings. In 2018, Christie's sold its first piece of AI-produced art for \$432,500. Such feats seemed impossible even a decade ago. They are also part of a larger automation revolution that has seen AI produce more accurate medical diagnoses, faster drug development, cheaper goods, and safer cars. There is little, it seems, robots won't take over. As they strengthen their foothold in the humanities, it's not clear if any higher ground is left for humans to retreat to. Much in the way certain skills and industries are in danger of being replaced, artists are being asked to contemplate their own obsolescence. "Clearly, AI is going to win," one developer has said. "It's not even close."

But what would it mean for AI to "win" at poetry? And what kind of poem would finally convince us? The answer depends less on what we believe a computer can do and more on what we believe suffices as poetry. "Most people have so little of an idea of what poetry is," wrote Paul Valéry, "that this vague idea is their definition of poetry." Indeed, for all the confidence with which AI researchers

throw around the term, few seem to have ever stopped to examine their assumptions about poetry. Those unexamined assumptions have a role in this debate. They are, in part, the reason so many believe we will close the rift between poet and machine. They also likely account for any misconceptions about creativity now being channeled into the bid to build an artificial writer. Maybe developers see poetry's self-evident artificiality—how it's marked by statistically measurable formal elements such as meter, rhyme, and assonance—and think equivalences are possible. If humans can master the mechanics of verse forms, why can't a machine? Or maybe developers understand "poetic" as the Racter-like by-product of placing certain words in a certain order, with suggestive properties arising from their grouping. The Turing Test, after all, has shown that readers have a weakness for rhetoric, grand gestures, and feelingful murk—all of which algorithms easily mimic. If this is what we mean when we say AI will one day rival human poets, then it will surely win, and indeed may already have.

But there's another kind of poetry AI will have to beat—poetry as an art of brilliant accuracies, of reality re-described in ways that bind sound to perception. And here AI's deficiencies are brutally exposed. Because to compete at this imitation game, a machine has to show that, by micro-adjustments of effect, it can draw our senses to the highest pitch of expression. It will need to be able to match Les Murray's depiction of beans as "minute green dolphins at suck," or Peter van Toorn's realization that flying dragonflies have "a great rattle of rice in their wings," or Elizabeth Bishop's noting a fish's "coarse white flesh/ packed in like feathers," or how, for Seamus Heaney, love was "like a tinsmith's scoop/ sunk past its gleam/ in the meal-bin." To play at this level, a machine has to imbue words with the most intimate associations and, turning inward, confess hardships, regret irrevocable choices, ponder its ultimate demise. It has to hit the same mark Robert Frost does at the end of his sonnet "Design," when, watching a spider readying itself to eat a moth, he asks what the grim scene reveals about nature—and if there is any moral code to such predation. "What

but design of darkness to appall? —/ If design govern in a thing so small.” The word “appall” here logs the shock perfectly. In its French root, the word means to make white. It also contains “pall”: a sheet, laid over a coffin, usually of white linen. Thus Frost’s diction hones our cognition, schooling us to see the world in a fresh way.

None of that is possible with GPT-3. Short for “Generative Pre-trained Transformer,” the model is unique not simply because of what it does, but also how it does it. It learns about language from watching grammar and syntax in action. The algorithms effectively train themselves. They pick up patterns in the data and, through a relentless process of trial and error, approximate them. That’s how, under the right conditions, GPT-3 can parrot impressively realistic paragraphs of text. Its credibility, however, drops to zero the longer you spend with it. Eventually you realize it is vacantly yoking bits of colloquialized detritus, bobs and tags of speech. Of course, the system makes a nice show of making sense, so we forgive its failures. But the failures are no less real: malfunctioning tones, misfires of inflection. GPT-3’s output is a shell of hyperactivity with nothing inside—a mesmerizing mix of materials without a center, language on autopilot. This is what led the *MIT Technology Review* to call GPT-3 “a fluent spouter of bullshit” and the researcher Timnit Gebru to warn against giving too high a mark to the program, citing the human tendency to “impute meaning where there is none.”

Gebru was referring to the alarming ease with which people fall under AI’s spell, a phenomenon memorably embodied in one of the earliest forays into text-based algorithms. Released in 1966 by the MIT professor Joseph Weizenbaum, ELIZA was the world’s first chatbot. Designed to impersonate a therapist, it would reflect back a user’s statements with open-ended questions and prepared responses (“My mother never loved me” would trigger “please go on” or “tell me more.”) Weizenbaum’s goal was to explore a computer’s capacity for conversation. Instead, he was alarmed by how completely users were taken in by ELIZA’s shallow repartee; his own

secretary once insisted he leave the room so she could talk to the program in private. Credulity even extended to graduate students who had watched him build ELIZA from scratch. Sherry Turkle, a social scientist and Weizenbaum’s colleague, called it “the Eliza effect,” which she defined as “human complicity in a digital fantasy.” We can see this effect in the love-struck language Racter’s programmers used to describe the moment their creation came to life. “We stared at each other (as Keats might have) ‘with a wild surmise . . .’ A machine performing arithmetic operations had just addressed us in our own language.” The programmers knew that “address” was an illusion—they had created it. Yet Racter’s scripted, anthropomorphic performance was enough to enchant and disarm. When it comes to AI, it seems, we can always be counted on to be easy marks, to fall for the trap we ourselves set. We *want* AI to win.

Why would GPT-3 be any different? So prepared are we to believe in its powers—to praise, as Meghan O’Gieblyn did in *n+1*, its “syntax of profundity”—we look past any evidence to the contrary. When it comes to its poetry, we will accept the claim for the deed—even as GPT-3’s poems ring dead. They have no sense of conviction, build no internal pressure, achieve no meaningful closure. GPT-3’s supercharged subroutines can, on demand, crank out endless, near instantaneous, strangely compelling turns of phrase (made to channel Wallace Stevens, it once yielded: “I must have/ Grey thoughts and blue thoughts walk with me/ If I am to go away at all.”) But drawing readers into a shared experience begins with mastering the trick of joining up the best of your phrasemaking into a standalone work, a memorable unity. And GPT-3 can’t.

That failure matters. As NLP, or Natural Language Processing—the field GPT-3 belongs to and has dramatically extended—takes center-stage in AI, machines that can, as Turing put it, “compete on equal terms” with poets have become prestige projects. It’s no coincidence that each time a new threshold is smashed, poetry is soon offered up as evidence of the breakthrough. The most profound exercise

of full human consciousness, poetry has long been coveted as a benchmark for silicon-based minds, the ultimate proof of concept. Its principles were not only present at the founding of artificial intelligence as a field—the 1956 conference that set out to design machines able to “use language, form abstractions and concepts”—but every step in eroding the line between robots and people has been marked by a poetry generator. When the famed futurist Ray Kurzweil wanted to sell the public on the idea of a thinking machine in the late 1980s, he began by inventing a “Cybernetic Poet.” In fact, you can even argue that the pursuit of machine poetry has driven entire sectors of AI, helping push the limits of what language models can now do. A 2013 symposium on artificial intelligence held at Exeter University declared the art form a “particularly valuable domain” because solving it would lead to advances on many fronts—including fixing the problems that still plague AI’s *sensation du jour*, GPT-3.

Is it only a question of time? If AI optimists claim bragging rights, it’s because they have demonstrated, again and again, that there is no task a machine won’t eventually crack. Chess was once thought beyond the reach of algorithms because of the cunning and foresight required to play it well. We also doubted AI could ever cope with Go, the ancient Chinese board game notorious for the astronomical number of positions its rules allow. On both counts, we were wrong: grandmasters were routed. Why would poetry do a better job of holding a mechanical mind at bay?

Poetry’s edge arises from the “terrible error” that the AI researcher Kate Crawford believes wrong-footed AI at its conception—namely, the belief that minds are like computers and vice versa. “Nothing,” she insists, “could be further from the truth.” To understand what she means, recall that when Garry Kasparov squared off against Deep Blue in 1997, the IBM supercomputer appeared capable of counterintuitive thought with a baffling move that left Kasparov profoundly unnerved. Eight days later, the greatest chess player of all time lost the deciding match. Something similar happened in 2016 when Lee Sedol, then the world’s best Go

player, faced Google’s AlphaGo. The AI program, using self-learning algorithms that outstripped even Deep Blue’s incomprehensible calculations, landed a move that so stunned Sedol with its strangeness, he needed fifteen minutes to recover. AlphaGo went on to win. “It’s not a human move,” remarked another Go champion at the time; “I’ve never seen a human play this move.” Why did AI beat us at these games? Because it learned to think like human beings about chess and Go, only smarter and faster? Hardly. It beat us because it learned to think in an entirely inhuman way. The scale of AI’s processing power—able to mull millions of strategies and pit itself against those strategies millions of times—found bizarre but superior solutions that centuries of flesh-and-blood play never considered, solutions so removed from normal reasoning as to be alien.

Poetry, however, is inexorably linked to *how* humans think—a kind of undeluded self-questioning that, as T. S. Eliot wrote, helps us become “a little more aware of the deeper, unnamed feelings which form the substratum of our being.” It’s also tied to the *need* to think this way. A poem’s mental force derives from the set of intentions driving it, intentions that push poets into action. But when GPT-3 gets the call to write a poem, it doesn’t know it’s writing “poetry,” or what “writing” even is. That last part is anything but trivial. Style is a sentient act: you strive for it. My point is that a computer will never replicate what poets do unless it can also replicate why they do it. Machines that write do so like machines: their poems are a statistical by-product of having absorbed vast strata of ready-made data no human mind will ever contain—the same brute-force method AI uses to decode whale language or track new particles in physics. For GPT-3 to pull off the real thing will require that algorithms not only move data, but are moved by it; that they not only consume our experiences but feel the fleetingness of our lives. How do you confer knowledge of mortality? There is no computational shortcut for that. Compressed into Frost’s choice of “appall” was a lifetime’s insight on loss. This is why poetry, unlike so much else our species has mastered, cannot be copied. It’s an artifact of

introspection that can only be mastered by our species. There is no superhuman way to write poems because we write them by virtue of being what a computer isn't: human.

In the end, it may not matter. What have algorithms taught us about creativity? That we don't actually think about it all that much. It's a lifestyle amenity, a commodity: something you switch on, download, stream, pirate, meme, mint as a digital token. There is no question poetry will be subsumed, and soon, into the ideology of data collection, existing on the same spectrum as footstep counters, high-frequency stock trading, and Netflix recommendations. Maybe this is how the so-called singularity—the moment machines exceed humans and, in turn, refashion us—comes about. The choice to off-load the drudgery of writing to our solid-state brethren will happen in ways we won't always track, the paradigm shift receding into the background, becoming omnipresent, normalized. The self as a source of memory and observation will be replaced by the self as a passive conduit for the almost limitless data surrounding us, a vast weather system of zeros and ones subtly, invisibly, and irrevocably breaking us down into parts that can be targeted and tapped. Could poems be created on the fly to optimize screen time, with our attention sold to advertisers? The book deals, legions of followers, and prodigious sales that followed the success of "instapoetry"—a phenomenon impossible without Instagram—has already shown that algorithms can legitimize, and merchandise, literary taste.

In a 1967 essay, the poet Howard Nemerov anticipated that if people grew to love poems written by computers, it wouldn't be because "the machine had imitated the subtlety of the mind, but that the mind had simplified (and brutalized) itself in obeisance to its idol the machine." That we struggle to recognize machine verse means our expectations for the human stuff are lower now. In his 2010 book *You Are Not a Gadget*, Jaron Lanier reminds us that the Turing Test "cuts both ways." For AI to pass, humans have to fail. We are judged as much as the machine. And what the test exposes is

how far our sense of poetry has strayed, how ready we are to be persuaded, to credit anything as genius. As machine poetry spreads, it will create a tolerance for things bots can do. AI will heighten, and push us to honor, poetry as a "construct," a system of vocabularies, a remote-controlled theatricality. We may end up cherishing the superficial and arbitrary effects most feasible for algorithms, becoming bored with interiority. Writing will appear less risky, less troublesome. We will be free of the expectation actually to understand it. We will also be free of its judgment on us—the demand that, as Rilke put it, "you must change your life." Maybe we will come to prize poetry that doesn't have any human reality in it. We will value deepfaked emotions, seeing them as better. Hand-woven stanzas will become vintage *objets d'art*: an artisanal good peddled on Etsy-like storefronts in the metaverse.

This isn't a debate about whether AI can write poetry. It's a debate about how much longer it matters that humans can. In swooning over Racter, Bök intuited a collapse decades in the making: the movement toward what the novelist Tao Lin called "the de-consciousnessed thing." Nearly a century of sense-scrambling experimentation—from the Dadaists and Surrealists to the L=A=N=G=U=A=G=E school of poetry—has not only primed us to admire what GPT-3 is capable of but also opened the door to humans writing machine poetry. What began in the early 2000s with Flarf poets turning random Google searches into tawdry collages has become a scene obsessed with repurposing, recycling, and remixing. Poetry, for many, is no longer something you compose but something you cannibalize—what the avant-garde poet Kenny Goldsmith called "uncreative writing." Why add to the unprecedented amount of information already surrounding us? Instead, he argues, steal. Powerful data-mining tools that comb public-domain content have made cut-and-paste verse easier (there is now an American press devoted solely to such books). But the mood doesn't require software. Case in point: the popularity of erasure poetry, in which documents are redacted into saying something new. Or consider the revival of patchwork forms like the cento and

glosa, which lift lines from other poems, thus making a muse of plagiarism. Words are now bits of information deflected, or misdirected, from one place to another, infinitely revisable, severed from any history or context. We keep making the Turing Test easier and easier.

With more start-ups getting funding in NLP than in almost any other category of AI, analysts say we are on a brink of a Cambrian explosion in language software. GPT-3 already fuels hundreds of bots, apps, corporate blogs, social-media feeds, and content farms. It is responsible for billions of words of functional and grammatically accurate copy per day. OpenAI recently promised a new version—GPT-4—that

will be five hundred times more powerful, able to sort through even vaster amounts of data. As systems get bigger, they get better: better at pruning mistakes that lead to bad sentences and better at repeating strategies that lead to good ones. GPT-4 will shock us with an even more uncanny ability to find just the right words, to arrange them artfully—with not a syllable wasted—all of it adding up to something that sounds human. And we will fall over ourselves to praise it. But as long as the ability to write poems remains a barrier for admission into the category of personhood, robots will stay Racters. Against the onslaught of thinking machines, poetry is humanity's last, and best, stand.