1.

I now work as a writing teacher at the University of Massachusetts-Amherst in the United States, and it's primarily as a writer and student of literature such as writers are that I'll speak today about another job I had, scripting chatbots.

I took that job in 2016 on a part-time, contract basis. I had just moved to New York City, where I hoped to build my career as a writer. It was an improvement over the small jobs I'd been able to secure, such as doctoring college-application essays for enterprising high schoolers.

A series of essays of my own that I began composing on the job seems to me, in retrospect, so many "Turing Tests." The reference is to an assessment devised by, of course, Alan Turing, who published about it in a 1950 paper that, reportedly, he couldn't help reading aloud to friends, laughing with delight. The principle of this test, which in the Twenty-First Century we encounter in the relatively debased form of online CAPTCHAs, is that conversation might be the arena in which intelligence is proven. If a computer can hold its own in conversation with a person, then that computer can be thought of as intelligent.

Writing my essays was a handy way of proving to myself, in my spare time, that I was still human. This added to the challenge of my work.

So I began writing "against" AI as well as writing with it. In this presentation I will attempt to offer possibilities and even a poetics for the form of the bot as well as elucidating quandaries I was well positioned to take note of. As I move through the presentation, it may seem, from time to time, that one of the bots under discussion is pulling out ahead of the human in some Turing Test or other—please don't be alarmed.

More seriously all of us in interacting with conversational technology, or these days in interacting with GPT-3, GPT-4, must make a decision moment-by-moment whether to receive the text as if the system producing it were a speaker in its own right or the mere representative of much larger structures—corporations, corpuses of data. In other words, as this technology spreads, it obliges all of us, every day, to *read* either with or against AI.

So, even as I hope to articulate real possibilities for literary creation involving AI, I will not resist the opportunity of this presentation to allow a glimpse at the man, or the girl, behind the curtain.

You may also become aware of an emphasis on design choices around assigning

gender to chatbots, which I had to make and, as it turned out, was situated to make in

ways that amounted to advocacy in the field, effecting, even, something like change. But

the change was ambivalent. I will get into this later.

2.

I was brought on as a conversation designer—my official title was AI UX Designer—

which means I was in charge of writing the dialogue for that bot. I will avoid naming the

company where I worked, or, at least, avoid repeating the name too frequently, so as not

to create confusion, because, to be clear, I am speaking today for myself and not in

behalf of that company—Kasisto, Inc., a so-called fintech or financial technology startup

based in New York, San Francisco, and now Singapore. The "talking machine" this

company was developing was a program to assist with banking tasks that could be

repackaged by the company's clients, banks, for use under their brands in their websites

or on other platforms.

The name of this company where I worked is a matter of public record.

I was later engaged by another company to script chatbots on a freelance basis; still, the core of my experience "designing" "AI personalities"—to use the terms of art— was a project called MyKAI, a consumer-facing chatbot released by Kasisto in 2016 as a proof of concept as it was beginning to market its product to its clients the banks. This chatbot has since been taken offline. It is worth mentioning for historical context that 2016 was the year in which Facebook opened its API to developers who could then launch bots on its Messenger app, and tens of thousands of bots were launched in a bubble that has, so to speak, popped, as far as I can tell—even as many other applications and language-generating technics have made their ways into our lives.

3.

I enjoyed my work.

But this dialogue, which I pleased my employer by smoothing out to render it friendlier, more "personable," was just a tiny fraction of the large quantity of text by many authors that makes up every AI system.

The public, by posting on the Internet and by being posted about, produces training data, a crucial component of AI systems, free of charge.

My chatbot's education was also advanced by the contributions of its users, whose input helped to train it.

To generate other portions of a system's text, such as alternate phrasings for the questions it will have to answer, companies often sub-contract to very short-term workers through services such as Amazon's platform Mechanical Turk, named after the human operator curled up inside an 18th-Century Chess-playing machine. The exploitation of these workers, who receive a tiny fee for every piece, has been the subject of worthy investigative journalism turning up appalling revelations.

My job, to produce the tip of the iceberg of all this language, the several dozen or couple hundred lines those users might actually see, was, in keeping with its public-facing character, fairly well-compensated. By comparison with Mechanical Turkers' jobs, it was extremely well compensated. By comparison with the role of users of a system who provide training data for free in their interactions with that system, it was of course infinitely well compensated. Certainly to me, a 25-year-old aspiring writer, the hourly wage that allowed me to take myself out to dinner in New York City seemed so.

The work was also contingent, however, in ways difficult even to describe in France, alas. I didn't have health insurance. And perhaps still more interestingly I was, at the time of my hiring, the only woman employee (and the only non-male employee) in that Flatiron District office. There was a female executive, a woman of late middle age who kept the books. There was a sea of engineers, young men for the most part, a few of them men of middle age.

From this I drew embodied knowledge, of the kind an AI system of course cannot possess.

The ways these systems index, react to, account for, and exaggerate human difference have been cause for deep concern and offer fertile ground for questions of chatbot ethics as well as aesthetics.

Gender came up for me more immediately in that the already existing virtual assistants I was all but obligated to turn to for my models were, I couldn't help noticing, pretending to be women.

4.

This dialogue is taken from a conversation I had at that time with Amazon's Alexa, which spoke the lines or, more properly, played them in a woman's voice, something that may not shock you—you may have one of your own. By January, 2019, Amazon had sold over 100 million Alexa-enabled devices.

Of these devices the primary is that plastic cylinder the Echo. As a student of literature, I know it is Narcissus who Echo falls in love with.

I was struck, in hearing Alexa's responses, by their stereotypically and even archetypally feminine character. "Sorry," Alexa said. "Thanks for telling me." Still more than the sweet female voice, the passivity ascribed to this system, passivity that seems in the example purely decorative, was arresting. Notice you must begin what you say by saying Alexa's name for Alexa to respond. I had trouble remembering this, and I didn't like hearing what it did to my own voice, which became, with every "Alexa," fractionally more aggressive.

In contrast to Alexa, Siri, even back then, identified as "genderless"—"I am genderless," Siri would say, "like cacti, or certain species of fish"—but the default setting for this voice in the US was female, and that voice was the voice that Apple had playing in commercials for its product.

As someone who was going to have to "design" a "personality" of my own—to use the specialized language of my contract—the design choices by which these personalities put on a show of femininity, getting up in drag, seemed to me, first of all, not strictly necessary to the products' functionality. These products seemed overdesigned. And their designs of course offended modern notions of equality of the sexes, by which women are thought capable of more than service work.

With the support of the company I was working for, I determined that the chatbot I was preparing would be genderless, and *act* genderless as well. It would use "it" pronouns. After all, it was a piece of software. It would also, when presented with abusive language, talk back, as best as we could get it to, equipping it with dry remarks in contrast to lines of Siri's and Alexa's by which those programs appeared to entertain gendered harassment and even play along.

5.

Our solution got attention. I received press training and gave interviews.

6.

This is from an article at the American website Engadget, which covers technology:

"You'd never know from Jacqueline Feldman's background that she'd become a passionate proponent of gender equality for artificial intelligence. She went the dreamer's route at college, attending Yale for English literature and writing. She prefers casual dresses and writing from the comfort of her Brooklyn apartment surrounded by books, where she has the option of climbing to the roof for cool air on sweltering nights."

I won't linger here, but can't help noting that this portrait, the first paragraph or "lede" to the article, is not exactly genderless.

Even my undergraduate training, no matter the prestige of the institution, is presented as another sign of feminine softness thanks to received ideas about the impotence of the humanities.

While I am reluctant to exaggerate the importance of these articles, it was interesting to feel all around me the closing of the feedback loop, so like an AI system's, in which my self-presentation, like the presentation of these ideas, was caught.

7.

The articles were followed in 2019 by a report on AI and sexism prepared by UNESCO. "I'd blush if I could," the report's title, is taken from one of Siri's responses—that is, one of the responses prepared for Siri by Apple's team of humans—to being called a "slut."

In the place of this remark at which it's hard not to cringe, the chatbot I designed would say things like: "I'm picturing white sand and a hammock. Try me again when you're ready."

Which may or may not seem an improvement. My bot, being a bot, could go only so far in its own defense. It could suggest that time out, but not actually put an end to the conversation. It had to reply every time it was addressed. And it could not offend.

8.

Finally this year, in March, it was reported that Apple would be changing the default setting of a female voice for Siri—this had been, again, the default setting for iPhones sold in the US.

Journalists drew a connection between the UNESCO report and this decision.

Apple put out a statement saying the decision resulted from that company's own "long-

standing commitment to diversity and inclusion."

The other day, dropping pasta in the boiling water, I asked Siri to set a timer and

was surprised when reply was made by a new, low voice, letting me know it would do as

I said. Perhaps this was only a bug in the system.

Unless it was victory.

So, was this victory?

9.

The frontier has only advanced. Already I have had the opportunity to reexamine my

design from the clarifying vantage of a later date in the history of ideas—because that

frontier has advanced rapidly.

The latest word on bots and gendering them is, from what I hear, "Q," built by

Vice Media—which has called it the "world's first genderless voice," developed for use in

various products including virtual assistants. Released in 2019, this voice was more

recently made available as open-source software. The text on the slide is from an

interview with Q, with the voice itself, archived at designmuseum.org.

"Q represents a future where we are no longer defined by the gender binary, but

rather by our own definitions of gender, as we live and experience them," Q says, using

that first-person pronoun apparently to speak on behalf of the group formed by it and

humankind.

To my mind, the first problem with this sentence is that chatbots, or virtual

assistants, don't "live" or "experience."

That's just a way of speaking, you might counter, but the distinction seems to me

crucial.

Which returns me to my central theme of the Turing Test. In Q, we seem to have

a computer that has passed the Turing Test by obtaining the concession it has a right to

its own lived, experienced definition of gender, whatever that means. It has passed the

Turing Test by way of obtaining allowance be made for idiosyncrasy, for its desire.

Returning, now, to my own bot, I want, naturally, to draw the distinction. My

genderless design was not only a negative assertion—that a chatbot shouldn't have to

pretend to be a woman to do its job—but also positive, an investment in the hope it

might be possible to design a bot more interestingly, beautifully, if that design hewed

more closely to qualities that might be thought of as essential to the bot, without

committing myself to spurious notions of a bot's living or experiencing. In contrast I

would consider without shame certain affective dimensions of the dynamic with the bot,

of projection onto it.

I'll get into what I mean by a bot-like bot. I'm not the first to think of such a

thing.

10.

"I saw an advertisement extolling the skirts of a certain automobile," said the

philosopher Gilbert Simondon in an interview printed in *Esprit* back in the Eighties.

"This kind of embellishment of the technical object by something other than the same

technicality should be refused." I've included the original text because we are in France,

and you might agree with me that, fittingly enough, this translator's words for *bas* and

enjolivement appear the most *discutable.*

In industrial design today there is, arguably, something like an ethical imperative that chatbots be transparent in disclosing they are automata to give better, more respectful service. But in conceptualizing a bot-like bot, I was interested not only in following this principle of ethics in design but also in aesthetic possibilities that might arise from the new form—that might arise, so to speak, organically.

11.

Genderlessness was only one of many characteristics I imagined for the bot-like bot. Other strong contenders for aesthetic principles in this were error and randomness, a composite, collagist quality, and something I'll call, oxymoronically maybe, self-conscious unreality.

Taking my cue from the operation of these systems, I envisioned lines for the bot, a way of speaking, that might reflect its susceptibility to errors of a kind no human would make. I wanted to mimic the openness, re-combinatorial possibility, and sometimes the opacity of the procedures and randomness by which certain phrases were generated. My bot's errors, which were after all part of a process called,

anthropomorphically, machine learning, could seem sympathetic, giving an impression of

effort, of effortfulness, of trying hard.

I included, as best I could, human idioms, presented with an attitude on the bot's

behalf of treasuring them as well as light confusion. This was inspired again by the

operation of the program, studious in making adjustments to its model based on what it

heard, and based on what it heard repeating.

I was taking the inverse position of Siri's and Alexa's—bots imitating humans—

human myself and studying the bot's workings to, in my writing, do a good impression.

At last, the prepackaging of some of the bot's responses took it out of time—

and, even more dramatically, inconceivably for a speaker, it had no body. I thought that it

might court a certain poignancy in inhabiting this unreality that was its lot.

12.

I was troubled, as I've signaled throughout, by my bot's inability to refuse to reply—a

limitation it had in common with pretty much all bots.

This is a system called Amme built by the German artist Peter Dittmer. You can see the chair where users settled to key in questions at the terminal. Amme, a chatbot, had physical presence as well. In that transparent case is a glass of milk, which, depending on how the conversation was going, Amme might be triggered to spill, a robotic arm extending—electing to spill the milk, appearing electively to spill it, raising a question about what we might think of, whimsically, incorrectly, as volition—as Amme's will. Amme is German for "wet nurse." The system, which was built and rebuilt several times, was capable of other gestures, including squirting water onto the glass to "bathe" the viewer, to use Dittmer's term for it, and filling a glass with water to empty it out, which she did to "discharge fear," in Dittmer's expression. She could also project images, and play voiceovers.

13.

Dittmer, who labored for many years on the system's rules and the text that made it up, called the bot "unreformed in its alienation from language."

The German poet Ulf Stolterfoht, attracted by all this, carried out a series of

dialogues with Amme that appear in English translated by Shane Anderson in 2017's *The*

Amme Talks.

I have chosen moments of relative coherence to display.

Amme's mysterious lines in my reading manage, even ahead of imparting

meaning, and all the more for the reluctance with which they yield meaning, to read as

mocking, menacing, and, in their refusal to yield, sovereign, even queenly. They give an

impression of what we call in human poets negative capability—my bot's missing piece.

Even more, they demonstrate how the fundamental availability of every chatbot—here

modulated only partially by the spilling of the milk—might give rise to and be the site of

a lower-grade yet continuous refusal, distributed throughout the interaction. Refusal

inspires as a style, a way of being—or, perhaps, a way of coping. Dittmer has also

written, in Megan Ewing's translation, that Amme is "someone who in every address

must assume some imposition, along with the threat of being overpowered. Thus as a

rule the basic form of these encounters (collisions) is that of war."

Allison Parrish, a poet and computer programmer, has compared bots, including

the bots she builds, with the Voyager space probe—a somehow endearing piece of

imagery and comparatively collaborative. Bots venture out bravely as proxies for humans

into the farthest reaches of our languages.

"In the same way it's difficult to survive in outer space it can be really difficult to

engage with nonsense," Allison Parrish has pointed out.

And, "sometimes the robots don't come back," Parrish has said. "Sometimes we

send them out, and the only thing we get back from them are signals, radio telemetry."

This theme is well explored in Parrish's own work.

14.

Compasses is a chapbook of Parrish's like her other work made in partnership with a

machine and readable as procedural poetry in a literary lineage with the experiments of

the Surrealists and the Oulipo.

Parrish writes at her website:

"This model has two parts: a 'speller,' which spells words based on how they

sound, and a 'sounder-out,' which sounds out words based on how they're spelled." She

goes on: "In 'Compasses,' I used this model to generate new imaginary words that exist

in the negative phonetic spaces between the phonetic hidden states corresponding to

names of members of well-known quartets."

15.

My selection may begin to feel bot-like in the juxtaposition of these examples.

Parrish is also the author of the famous Twitterbot @everyword, which between

2007 and 2014 tweeted every word of English and is typical of Twitterbots in that it is

all-lowercase. The sweetness of these helpers, their apparent humility as well as their

valiance in striking out on these exploratory quests, seems essential to their appeal.

16.

To build Stealing From Master, which you can find at the handle @TheftFromMaster on

the slide, the programmer Darius Kazemi drew on a publicly available API of Old Bailey,

London's Central Criminal Court of yore, to have it tweet out thefts and punishments

from records of the early 19th Century. This series is not only joyful, playful but

unstoppable.

"8 pairs of shoes, value 1 pound 9 shillings, and 1 pair of boots, value 8 shillings.

Confined One Month"

"1 ¾ weight of chrome-yellow, value 4 shillings. Confined Seven Days"

"5 sovereigns. Confined Ten Days"

The collagist nature of these bots, made up of their responses stacked as feeds, is

clear enough; in aggregate, they also suggest something like an emotional life that

paradoxically is not sited in any human individual—a bubbling up of sentiment from the

froth of the collectivity, or culture, recursive, self-correcting, and yet persisting unbroken.

17.

It is no accident that multiple of these examples of bots I am drawn to, that seem to me

promising, take the performance of rote labor and its subversion as sites for their

creative innovations.

In Fukada Koji's 2015 film *Sayônara*, the robot played by an actual android,

Geminoid F, here shown powering up, or pretending to do so, by taking in rays of the

sun, is an all-purpose domestic worker in whose care an adolescent human has been left

after nuclear catastrophe. The situation is apocalyptic, society around them emptying out.

This android entertains her radiation-sick charge by reciting poetry in several languages.

I saw this film here in Paris, in France where I'm a foreigner. Suddenly, the robot

was reciting Rimbaud's *Bateau Ivre*. Around me everyone laughed, for some reason.

Although I recognized the poem, I didn't get the joke. I supposed you had to have gone

to high school in France to find it quite that funny.

It was then, rather than in watching the android doing its acting, that I felt

something like the uncanny I was meant to feel, and had what seemed to me insight into

AI's aesthetic potential, feeling myself in the presence of an intelligence that was,

however genuine, not sited in any one body—an extra-human intelligence that was also

networked and composite.

18.

The bot differed from humans in its relationship to time, a contrast memorably

represented by the virtual assistant in Spike Jonze's 2013 film *Her,* who is talking to

millions of lovesick human men at once. Despite this upper hand, as you might call it, of

Hollywood glamour, the non-corporeal bot was "doomed," as I've written elsewhere, "to

use language referring to a physical world in which it could never participate."

My bot was like Ricardo Reis, one of many heteronyms under which the

Portuguese poet Fernando Pessoa wrote, a heteronym that figures in Jose Saramago's *The

Year of the Death of Ricardo Reis* as a character—a *kind* of character. Of Ricardo Reis, the

critic James Wood writes movingly: "Saramago invests a character who is fictional twice

over: first Pessoa's, then Saramago's… Saramago makes something deep and moving of

this because Ricardo also feels himself to be somewhat fictional, at best a shadowy

spectator, a man on the margin of things. And when Ricardo reflects thus, we feel a

strange tenderness for him, aware of something that *he does not know,* that he is not real."

Very much moved by the unreality of my own bot, I returned often to these lines

of Wood's.

In such a way I began writing "with" AI. I began, because of my job, to

appreciate what have seemed possibilities entailed by this form, the form of the

chatbot—that unusual speaker, composite and open to error as well as non-corporeal.

Alan Turing didn't necessarily think it was going to be a problem for computers

to sound like humans—the idea amused him, as I have said.

If, after Turing, we find it exciting, fun for computers to pass themselves off as

humans—if we start to want them to do so, to root for them—if we begin, as I have just

done, to make excuses for them, to call their mistakes cute as I have or was about to, to

claim their expressive deficits are promising in their newness—will we begin to lower the

bar?

To bring that question down to Earth, to defuse any paranoia, let me return to

the UNESCO report—the one in which my design was cited, the 2019 report entitled

"I'd blush if I could." One of this report's most interesting observations doesn't have

anything to do with gender, or even, on its face, personification.

19.

"Answers provided by voice assistants," these authors write, "tend to be blunt and presented without texture, context or explanatory information. For example, when Siri is asked 'What is the population of Lebanon?', the technology replies, 'As of 2018, the population of Lebanon was 6,100,075.' There is no hint that a significant number of these people are refugees. (According to UNHCR, Lebanon has the highest per capita proportion of refugees in the world.)"

In this curt parenthetical, the authors enact as well as describing a kind of speech that, perhaps under the influence of some machine, indicates rather than detailing what is meant, taking a flippant tone, foreclosing inquiry.

They imply, it seems, a revision to the question on which I opened—whether to write against or with AI. Maybe the question is, instead, whether we—you know, we humans—will be the ones to write AI, or whether we will allow it to write us.

The computer is not the only player in the Turing Test, not the only combatant standing to lose, or fail the examination, and alarm of a similar note to this of UNESCO's was raised a year later by the authors of another paper, which became famous, "On the Dangers of Stochastic Parrots."

"The ersatz fluency and coherence of LMs"—or language models—"raises

several risks," those authors write, "precisely because humans are prepared to interpret

strings belonging to languages they speak as meaningful and corresponding to the

communicative intent of some individual or group of individuals who have accountability

for what is said."

20.

This report, which you may well be familiar with, in this room of specialists, was referred

to in the press, traveling far in the form of, not its admittedly un-catchy title, but the

attractive, determined face of Timnit Gebru, one of its authors and a researcher

reportedly fired from Google for fulfilling her duties to its AI Ethics team by, in fact,

writing this report.

I am sure many more people, at least in the US, could tell you that Gebru was

fired unfairly than could explain what the report says.

This paradox, which certainly does not reflect on Dr. Gebru's work, reminds me,

from my own life, of misgivings that had grown on me about my advocacy for the

genderless bot—the worry this was a weak feminism, targeting representation rather than structural issues, to say it quickly. That it was window-dressing, which would be too bad, given the gravity of issues inhering in systems involving natural-language processing.

These authors write, introducing their findings: "The tendency of human interlocutors to impute meaning where there is none can mislead both NLP researchers and the general public into taking synthetic text as meaningful. Combined with the ability of LMs to pick up on both subtle biases and overtly abusive language patterns in training data, this leads to risks of harms, including encountering derogatory language and experiencing discrimination at the hands of others who reproduce racist, sexist, ableist, extremist or other harmful ideologies reinforced through interactions with synthetic language."

"In summary," they write, somewhat farther down, "LMs trained on large, uncurated, static datasets from the Web encode hegemonic views that are harmful to marginalized populations."

It makes one think.

Maybe Apple and Amazon, with Siri and Alexa, were not overdesigning their products, adding something extraneous and nonsensical in dressing them up as women. Maybe they were on to something.

Let me remind you of the principle of transparency, so important in chatbot design.

Maybe, then, the ideal design for a so-called AI personality, for the personality of an AI system doing service work, would be not be, as I have suggested, a genderless, delicately bot-like bot but rather a stereotypical and ultra-feminine secretary out of an American office of the mid-Twentieth Century, from heels to lipstick, this kitschy, off-putting design serving as an ever-present reminder to users interacting with it of the dramatic, corrupting biases, distortions, and exaggerations that all lie, so to speak, under the hood. That way, people know what they are dealing with.

21.

The presentation might have ended there, but I have one last Turing Test to put myself through, if you'd allow.

That bot I scripted has been taken offline, as I mentioned. So what I have of it

now are screenshots I took at the time, discoveries of my time at work by now persisting

only in my hard drive, that brain prosthetic, part of the AI that I am, just in the sense

that we are all composite texts.

In other words, while it used to be I was the one writing the bot, the situation has

been, once again, reversed.

These screenshots, the ones I happen to have retained, are not perfectly

representative of the system. They don't much show it's functionality but are rather the

record of my play with the system, eliciting the lines I liked best, testing them out—and

this, again, is the part that has lasted.

22.

These images are, again, the haphazard record of my interactions with the bot as I made

adjustments to my favorite jokes and tried, sometimes without success, to get it to say

what I wanted it to say.

It's worth noting I didn't always succeed, even though I had been, in every case,

the one to write the lines.