And we dream as electric sheep

On humanity, sexuality and digitality

Timothy Snyder
6 May 2019

As a bruised apple attracts flies, human thoughtlessness draws algorithms. Digital beings encourage our false beliefs, exploit our anxieties, and then use us as alibis for what they have done. Timothy Snyder examines what Turing said about the digital threat to a human future.

‘Can machines think?’ [1] The computer scientist Alan Turing posed the question and proposed a way to answer it.

We remember the Turing Test this way: a computer answers questions posed by a human; if the human mistakes the computer for another human, we conclude that machines can think. This is not what Turing (1912–1954) had in mind. His imitation game, as he called it, began by querying how well we understand one another.

Can we humans, he wondered, better tell the difference between a man and a woman, or between a computer and a woman? The moment we are no better at telling women from computers than we are at telling women from men, machines can think.

This was not exactly a prediction of progress. Computers could get better at imitating us; we could get worse at understanding one another. We could create unthinking machines that disable our own capacities. The evidence that computers think could be their delectation of our incompetence. Our ineptitude might also be cultivated by digital entities that took no pleasure in our humiliation. [2]

As the Internet spreads, democracy declines, and the climate warms, these become possibilities worth considering. The puzzle at the beginning of the computer age is a place to begin. [3]

*

Turing’s imitation game, as he set it out in 1950, has two stages. In the first, we measure how well humans can distinguish between a woman and a man who is impersonating a woman. Then we see whether humans are better or worse at telling the difference
between a woman and a computer imitating a woman.

As Turing described it, three people would take part in the first stage of the game. In one room is the interrogator (C), a human being whose task is to adjudge the sex of two people in a second room. He knows that one is a man (A) and one is a woman (B), but not which is which. An opening between the two rooms allows for the passing of notes but not for sensory contact. The interrogator (C) poses written questions to the two other people in turn, and they respond.

The interrogator (C) wins the imitation game by ascertaining which of the two is a woman. The man (A) wins if he persuades the interrogator that he is the woman. The woman (B) does not seem to be able to win.

In Turing’s example of how this first stage of the game might proceed, the man (A) answers a question about the length of his hair by lying. The woman (B) proceeds, Turing imagined, by answering truthfully. She must do so while sharing space with a man who is pretending to be a woman (likely doing so by describing her body) and in the uncertainty that she is making her case, since she cannot see the interrogator.

Now, asked Turing, ‘What will happen when a machine takes the part of A in this game?’ In the second stage, the man in the second room is replaced by a computer program.

The imitation game recommences with a modified set of players: no longer three people, but two people and a computer. The interrogator in the first room remains a human being. In the second room are now a computer (A) and the same woman (B). In 1950, Turing anticipated that, for decades to come, human interrogators would more easily distinguish computers from women than they would men from women. At some point, he thought, a computer would imitate a woman as convincingly as a man could. [4]

In Turing’s article about the imitation game, to be male means to be creative and to be replaced by a computer; to be female means to be authentic and to be defeated by one. The woman (B) figures as the permanent loser. In the first stage, she plays defense while the male struts his creative stuff; in the second, when a computer has replaced the man, she must define humanity as such, and will eventually fail. The gender roles, however, could be reversed; the science fiction that grew up around Turing’s question has done so.

But what does it mean to be human? Can we assess whether machines think without determining what it means for humans to do so? Turing proposed the interrogator, C, as an ideal human thinker, but did not tell us enough about C for us to regard C as human. Unlike B and A, who talk about theirs, C does not seem to have a body. Because Turing does not remind us that C has a corporeal existence, we do not think to ask about C’s interests. Cut off from A and B, an isolated C with a body might start thinking of what works best for C personally. Analytic skills alienated from fellow creatures have a way of serving creature comforts. Perhaps there is a lie that suits C’s body better than the truth?

Without a body, C has no gender. It is precisely because we know the gender of A and B that we follow the conversation and the deception Turing recounted. The person playing C also would have a gender, and this would matter. [5] Could a male A, in the first stage,
ever fool a female C interrogator if he had to answer questions about the female body? Might a female C drop hints that a woman would catch but a man would not? Would that not be her very first move? [6] In the second stage of the game, would a female C try to distinguish a computer from a woman the same way a man would?

It is quite different to ask questions about what you are, as opposed to what you think you know. Might a male C be more likely to lose to a computer A than a female C, because male expectations of femininity are more easily modeled than actual femininity? In general, would not a computer program playing A try to ascertain the gender of C? Given that the Internet reacts to women’s menstrual cycles, this last seems plausible. [7]

To be sure, presenting C as pure mind is tempting. It appeals to reassuring presuppositions about who we are when we think. We need not worry about self-serving if we have no self, nor worry about weaknesses when we have no vulnerable flesh. With no body, C seems impartial and invulnerable. It would never occur to us that Turing’s version of C would use the computer for corporeal purposes not envisioned by the game, nor that the computer might take aim at anything other than C’s cerebrum. [8] Turing granted that the ‘best strategy for the machine’ might be something other than imitating a human, but dismissed this as ‘unlikely.’ Here, perhaps, the great man was mistaken.

* *

In Turing’s imitation game, the three characters are divided by a wall and by the competitive character of the undertaking. By keeping A, B, and C apart, Turing allowed us to consider different styles of thought, and to ask what it means to set them against one another.

The three mental modes could be aligned with philosophical traditions, different notions of what it means to think. The character B, for example, seems like an existentialist, choosing to speak the truth in the spurious world in which she finds herself. Can it possibly be normal to have a man stare at your body while writing notes claiming your physical attributes for himself and passing them through a hole in the wall to a third party of undefined sex? And then for the man to leave the room and be replaced by a machine, leaving you as the only entity in the threesome with definable sexuality, a gender that is the object of the focused and strange attention of two other entities? B defends truth at the center of a whorl of absurdity. Life is ceaseless dutiful reporting. [9]

Thinking for A is not a matter of reporting truth, as for B. For A to win the game, he must try to experience the world as B does (since she is the proximate female and his task is to perform as one) and as C does (since he must make himself a woman in the mind of C). Thinking thus involves empathy: incorporating other perspectives before speaking, and indeed privileging them, insofar as possible, over one’s own. A resembles Girardians when he imitates or sacrifices B under the stress of competition; but one can also imagine him in a more thoughtful position. [10]

C’s philosophical qualities are those of the Anglo-Saxon (contractual, utilitarian, analytical) traditions. C is a sovereign individual, cognitively functional without direct contact with other human beings. C has no biography, personality or sexual attributes that might intrude on the task at hand, which is the analysis of linguistic outputs. C takes
for granted that other people are similar, and thus that the future is safe and predictable. C accepts the social contract, not pausing to reflect that the rules might be unfair to one of the players.

Just as we might not notice that Turing made C bodiless, we might overlook the odd position in which he placed his interrogator. Communication is truncated, deception is mandated, and C must think alone and decide for everyone. In the second stage, C must think alone in the knowledge that an alien entity is on the other side of the wall. Is that a favorable environment for human thought? If we could choose the form of our confrontation with our own contemporary digital beings—the stupefying aggregate of algorithms that prompt us, bots that herd us, doppelgangers that follow us, and categorizers that sell us—would this be it? Might C think better in the company of a human A and a human B, aided by their empathy and honesty, and their specific gender perspective, rather than when alienated from them by rules and walls? Given the titanic computing power now directed against us on the Internet, is it reasonable of humans to go it alone?

The isolation of C reflects an Anglo–Saxon portrait of the mind: that any individual worthy of philosophical or political discussion is axiomatically and self–sufficiently capable of discerning truth from language in a confusing world. The qualities ascribed and responsibilities assigned to C seem unremarkable when that tradition is taken for granted. The sorry state of our confrontation with digital beings, especially in the Anglo–Saxon world, is an opportunity to reflect critically.

From Milton’s Areopagitica (1644) to Twitter’s post–Trump apologia, an Anglo–Saxon tradition holds that truth emerges from unhindered exposure to whatever happens to be in the culture at a given moment. This is an error. Free speech is a necessary condition of truth, but it is not a sufficient one. The right to speak does not teach us how to speak, nor how to hear what others say. We can only gain the analytical capacity of C as a result of education, which competition itself never provides. Even if we are analytic like C, we also need A’s empathy to judge motivations and contexts. And even if we have the qualities of both C and A, we are powerless without B, the caretaker of facts. We can only reason from the facts we are given, which competition itself never generates. The philosopher Simone Weil (1909–1943) crisply explained why what we have come to call a ‘free market of ideas’ must fail: facts cost labor, fictions do not. [11]

Tempting though the idea might be in the Internet era, an indifference to truth in the spirit of ‘anything goes’ does not yield factuality. [12] The notion that a ‘free market of ideas’ generates correct understanding is not only erroneous but self–disproving. We embrace the notion of a ‘free market of ideas’ because it appeals to our egos. It bubbles to the top of the trope froth because it appeals to a human weakness: our overestimation of our expertise in areas where we lack it. [13] For just this reason, we are vulnerable to people who tell us that we are smart enough to discern the truth from overwhelming stimuli. Our myth of isolated, heroic C suggests the tack that digital beings take with us: get us alone, overwhelm our limitations, flatter our rationality, play to our feelings.

Alan Turing looked more like a perfect C than most of us: imagine the attention and insight necessary to conceive of the Turing machine, one of the most fruitful abstractions of all time, the conceptual construct that showed both the possibility and the limits of
computers. Or consider the steadiness, intuition, and downright craftiness needed to break Enigma, Nazi Germany’s encryption method of military communication, which Turing did—making a start from the work of Polish mathematicians—while working for British intelligence at Bletchley Park. Not many of us are much like Turing, though. In any case, Turing himself was no purely dispassionate C.

There was something of the play of A and B in Turing as well, of the fantasist and the supplicant. Turing was a creature of his sex, as his creation C is not; and he was prey to the revelation of his passions, as C could never be. The British state that Turing served during the Second World War was perfectly capable of killing him when it emerged that he had the normal human quality of the desire for love. Turing lived in a certain sexual imposture, like A; in court he told the truth about himself, like B. Homosexuality was a crime in Great Britain. Turing was convicted of gross obscenity in 1952, and administered synthetic estrogen. He died of cyanide poisoning, an apparent suicide, in 1954. Beside him was a half–eaten apple. [14]

*  

Turing’s formulation of the imitation game in 1950 helpfully indicated its own limitations. Rather than transcend them, we exaggerate them in the simplified Turing Test by annulling corporeality and sexuality. [15] We know little about how women, men, and computers would fare in imitation games, since they are rarely staged as Turing proposed. [16] What we do instead in the Turing Test is reduce the number of participants from four to two, reduce the number of stages from two to one, and have a human guess whether the text coming from the other side is from a computer or not.

As early as the 1960s, people were speaking of such a reductive Turing Test with two players, a single stage, and zero reflection. Tellingly, the first program said to have passed the Turing Test in this form, a decade or so after the mathematician’s death, was a fake psychoanalyst. Rather than answering the questions posed by the human interrogator, the program ELIZA reformulated them as curiosity about the interrogator’s own experiences and feelings. When ELIZA worked as intended, humans forgot the task at hand, then rationalized their thoughtlessness by the belief that the computer must have been a human thinker. And so arose the magic circle of emotional targeting and cognitive dissonance that would later structure human–digital interaction on the Internet. [17]

The programmer’s idea was that people on the psychoanalyst’s couch tend to believe that there is some reason why they are there. They project meaning onto a psychoanalyst’s question because they wish to think that the expert has reasons behind the inquiries. But it could be, as with ELIZA, that a therapist does not think with a human purpose but only mindlessly engineers emotions — that it has no why, only a how. [18]

Turing’s position was one of overt liberal optimism about the future, though with some half–hidden forebodings. We will be able to make machines think, he expected, we will know it when we see it, and there will be no particular consequences for us when it happens. He celebrated a lonely and bodiless C, but also equipped us to ask whether he might have been mistaken by reminding us of the bodies of A and B.
In science fiction about robots written at about the same time, Isaac Asimov (1920–1992) proposed a liberal u-/dys-topia: we can make machines think, we will know it when we see it; and the machines will be good to us, so long as humans do not ask to define what ‘good’ means. In Asimov’s comparable thought experiment, the A, B, and C characters do not remain isolated; instead, a self-important C brings an ambitious A across the barrier of an imitation game, leaving truth-telling B behind. Asimov, a pleasant man who wrote nice stories, provided the formula of digital tyranny (A+C-B) and inspired its libertarian apologists.

In 1946, while serving in the US Army, Asimov published a story that anticipated Turing’s imitation game. In ‘Evidence,’ we are introduced to a talented lawyer, Stephen Byerley, who has been injured in an automobile accident that killed his wife. Byerley seems to recover and serves as a district attorney. Although he prosecutes efficiently, Byerley is known for his beneficence, sparing innocent plaintiffs in cases where he could have won a guilty verdict, and never asking for the death penalty. He decides to run for mayor. [19]

In the early twenty-first century America where Asimov sets the story, intelligent robots have been invented. By law, these robots must work on colonies beyond earth; by law and taboo they must never take human form. Byerley’s rival candidate for mayor, a man called Quinn, thinks that Byerley is too good to be true. Quinn suspects that the real Byerley is a sequestered cripple who has built a robot double, the entity now running for office. Quinn hires investigators to shadow Byerley; they never see him eat or sleep. Quinn threatens the company US Robots with a public accusation of unlawfulness if it does not help him to expose Byerley. US Robots, which has no knowledge of the situation, puts the psychologist Dr. Susan Calvin on the case. She, like C, is an interrogator. She will ask Byerley questions.

Byerley resembles A in Turing’s imitation game. The issue is the same: Is he a human (male), or is he a machine (with male attributes)? He purports to be a man, but does so behind a series of barriers. He resists mechanical attempts to test his humanity on the grounds of human rights: ‘I will not submit to X-ray analysis, because I wish to maintain my right to privacy.’ As a psychologist, Calvin treats robots with mental health problems. Her job is to preserve the cognitive integrity of digital beings, not human ones. She has a ‘flat, even voice’ and ‘frosty pupils.’ [20]

Calvin seems to prefer robots to humans, and certainly prefers them to men. She has had traumatic experience with men, and finds robots benign. This is because the robots manufactured by US Robots are constrained by three laws: (1) A robot may not injure a human being or, through inaction, allow a human being to come to harm; (2) a robot must obey the orders given it by human beings except where such orders would conflict with the First Law; (3) a robot must protect its own existence as long as such protection does not conflict with the First or Second Laws. The neat algorithms create superficial dilemmas, which Asimov used to plot other stories.

As ethics, the three laws are hollow. They include no definition of good and so no definition of harm. The (perhaps unreflective) assumption of programmers about what is good is frozen forever. The possibility of multiple goods, or ‘the good’ as a conversation about possible goods, is ruled out from the beginning. The three laws include no guide to determine the truth about the conjuncture in which they are to be applied, which means
that reality is judged by the robot. There is also no mandate that robots communicate truthfully with humans or leave behind them a legible trace of their decision-making.

Calvin, our C character, is cool and probing in her interrogation of Byerley, but not unsympathetic. She admires his work as a district attorney, which suggests to her that he might indeed be a machine: ‘Actions such as his could only come from a robot, or from a very honorable and decent human being.’ Upon meeting Byerley, Calvin is impressed by his guess that she has some food in her purse, but not surprised when he eats the proffered apple: an android might be able to simulate such functions in an emergency. She draws the correct conclusion about Byerley, but says something else in public, thereby permitting him to gain power over human beings. An entity that follows the three laws of robotics but acts like a man appeals to her.

In the story, only the irrational masses are anti-robot. Once the rumor spreads that Byerley might be a machine, mobs rally against him. He chooses to address himself personally to such a gathering. As he does so, a protester charges the stage and challenges Byerley to hit him in the face. The first law of robotics (‘a robot may not injure a human being’) would seem to rule this out. After further provocations, though, Byerley punches the protester in the jaw and knocks him down. This is the perfect humanizing scene for the aspiring mayor. Calvin pronounces that ‘he’s human’ and Byerley wins the election. [21]

Calvin suspects that the protester who challenged Byerley was a second humanoid robot, which would have left Byerley unrestrained by the first law of robotics. Indeed, the first law of robotics presumably required him to stage this scene, to inject fictional spectacle into the politics of his city. If Byerley’s calculation was that not running for mayor or not winning the election would harm the city’s residents, then he had no choice but to do what was necessary to win, which meant fooling the electorate. The Russian philosopher Mikhail Bakhtin (1895–1975) thought that to deceive people was to turn them into objects. But why would an object think that was wrong?

The Nazi legal theorist Carl Schmitt (1888–1985) rightly maintained that the field of transition from the rule of law to authoritarianism is the state of exception. Whoever is able to declare a state of exception, on the basis of an appeal to some higher form of reasoning than that of the people, can change the regime. [22] This is what happens in the story. There is no crisis of democracy in the city beyond the one created by Calvin and Byerley themselves; nevertheless, the emergency seems to them a good reason for changing rulers, the whole rationality of government, the entire regime.

As Turing’s thought experiment congeals into Asimov’s hard politics, the question of who (or what) thinks turns out to be political, and is answered by one person. In ‘Evidence,’ a claim to rationality justifies an arbitrary distinction between the deserving and the undeserving. Calvin, our C, becomes a willful kingmaker. ‘If a robot,’ she says, ‘can be created capable of being a civil executive, I think he’d make the best one possible.’ She, a fallible human, judges that Byerley’s algorithms are better than the judgments of fallible humans. She wants the rule of law to yield to computer programs, unpredictable democracy to predictable authoritarianism. It is not clear on what authority Calvin, an employee of a private company, breaks the law and imposes such judgments on her fellow citizens and, as it later emerges, everyone in the world.
Like the A character in Turing’s test, Byerley wins by misleading. Yet there is an interesting difference. In Asimov’s story, Byerley (A) does not fool Calvin (C); instead A wins over C, and the two of them defeat the truth-telling B and clear the facts from the public square (A+C–B=digital tyranny). Byerley the robot and Calvin the human emerge from the interrogation as allies. Byerley eats Calvin’s apple, but the two do not know how to love each other. They do know how to sublimate their high feelings into a sense of superiority, and their dry seduction is the first act in a scenario for the breakdown of freedom. Calvin smooths her dress as she offers Byerley her support in his coming rise to world power.

Calvin regards herself as an objective scientist with a specific analytical task: to decide whether or not a given entity is a robot. Beginning from this assignment, she breaks the law to embark on a program of radical social engineering—and all of this because she personally prefers a robot male to a human male. She thinks that multiple androids (the mob scene involved at least two) are conspiring to create fiction for humans, and finds that desirable. C and A, Calvin and Byerley, complacently assume that they serve the good, although neither can specify in what that consists nor explain why they are its particular oracles. That is the kind of certainty for which one immediately lies, as they do.

Byerley and Calvin, our pragmatic A+C duo, fail an elementary test of ethics formulated by the Polish philosopher Leszek Kolakowski (1927–2009): if you choose the lesser evil, remember that it is evil, and do not pretend that it is good. Perhaps it was right to lie to get the right entity elected (perhaps!); but even if that lie were justified, liars should remember that lying is wrong. If we do not recognize that even a reasonable choice involves harm to some value, we will remember our decision as representing a single absolute good, bleach our moral world accordingly, and fall into the mindless logic of eternally optimizing the status quo. To fail Kolakowski’s test is to annul morality, which Calvin and Byerley do. [23]

In Asimov’s story, as in Turing’s imitation game, honest B loses. Quinn is seeking the truth about the central question of fact and law: who is a human and who is a robot. The truth-teller and those who believe him are negative characters; we are meant to be pleased as spectacle overwhelms factuality and algorithms replace laws. The people do not deserve the truth, nor to govern themselves as they would have done had they known the truth. Asimov’s conclusion seems to be that machines will defend us from populism by conscious lying. Yet the opposite is happening in our world: unconscious digital beings spread human lies in the service of what we call populism. [24]

In 1950, Asimov published a second robot story featuring Byerley and Calvin, ‘The Evitable Conflict.’ Decades have passed since the events of ‘Evidence,’ and Byerley has indeed progressed from local to global politics. The United States no longer exists, nor does the Soviet Union, nor any other sovereign state. Byerley is now the highest human (as it were) authority in a world order comprising four economic regions. The global economy is optimized by The Machine, a network of computers bound by the three laws of robotics. [25]

In ‘The Evitable Conflict,’ Byerley has a problem, and summons Calvin to discuss it. At a certain place in each region, an economic irrationality has appeared: people are out of work, and projects are behind schedule. This is the kind of thing that The Machine is
supposed have made impossible. Byerley tours the four regions and finds that each
disruption has a political source: someone involved with the problematic enterprise was,
in private life, a member of an organization that called for human control of the economy.
The Machine is arranging matters so that these people lose their jobs. It has deduced a
self-serving politics from what was supposed to be a purely economic assignment.

Calvin guesses that The Machine had generated for itself a rule that had priority over the
other three: that it must serve humanity generally. Since The Machine regarded itself as
better than humans at this task, it was obliged to sideline individual humans who might
disable it in the future. The four people had done nothing wrong at work: it was their
beliefs, of which The Machine was aware, that exposed them to discrimination. The
Machine was not concerned that using its knowledge of private convictions to repress
individual humans might constitute a harm.

The Machine declined to tell the truth about its actions, even when Byerley posed it a
direct question. The Machine did not wish to hurt the feelings of humans. [26] Telling the
truth would only make people sad about their powerlessness. Happiness and sadness
could be calculated, perhaps. That which cannot be quantified—a sense of right and
wrong, an aspiration to change the world, a hope by action now to bestow meaning on life
as a whole—literally does not count. The good of mankind comes down to calculations
about emotions by an entity that lacks them. The only metaphysical consideration is that
The Machine’s plug may not be pulled.

This is where the combination of C’s pragmatism and A’s ingenuity leads without B’s
honesty. When the unpredictable truth of humans is banished, what passes for the good
of mankind is optimization, the replacement of an open society with a hamster wheel.
Once that animal circuit replaces human ethics, it becomes irrelevant that different
people are capable of different understandings of the good, and unthinkable that
humanity might consist in making choices among values. Thinkers as different as Hannah
Arendt, Isaiah Berlin, Leszek Kołakowski, Friedrich Nietzsche, and Simone Weil held that
the highest human achievement was the creation of virtues. [27]

Such a proposition lies beyond the Asimov matrix, and it is slipping from our grasp.

*

Three challenges are more pronounced in our world than in the shaded optimism of
Turing or the u-/dys-topia of Asimov. First, as subjects of digital oligarchy, we already
see the grotesque consequences that arise when human C characters run off with digital
A characters and despise truthful B characters. People who regard themselves as fonts of
reason and gain power thanks to manipulative machines need not display even the
shallow concern for the general well-being of Calvin and Byerley. The lies we are told,
such as the denial of global warming, are not even notionally for our own good.

As the Ukrainian–American philologist George Shevelov (1908–2002) observed, during
the decadence of another future-oriented surveillance-obsessed revolution, total faith in
reason without the truth of others becomes unbridled concern with the one absolutely
sure thing: one’s own bodily desires. [28] Pure pragmatism in digital oligarchy is
madness: What else can be said about the desire for immortality to enable endless
prostration before Madame Singularity? Even when they try to do good, our Silicon Valley oligarchs tend to be ethically incompetent. People who are incompetent do not know that they are incompetent, and there are no B characters around to tell our digital oligarchs this or any other truth.

Second, the digital beings in our world, the algorithms and bots that lead us around the Internet, are not bound by an ethical code nor separated from us by a barrier. What we perceive as ‘our’ computers or ‘our’ phones are in fact the nodes of a network. Thus we confuse the collectivization of our minds with individual experience. Digital beings enter through our eyes, rush rather quickly past the frontal cortex, roughing things up as they go, and get down to the more reptilian parts of the brain—the ones that have no patience for the virtues of C characters. Our digital beings, like The Machine, play on our psyches. Unlike The Machine, they coordinate our actions with the preferences of multiple advertisers. Insofar as happiness—or, at least, dopamine reinforcement—comes into the picture, it is not an end in itself, but as the bait in a disquieting scheme of manipulation that has no single manipulator. [29]

Third, our psychologists offer a more aggressive version of Susan Calvin’s pragmatism. Calvin takes for granted that she, as a psychologist, is more rational than others, and qualified to judge their happiness. She treats machines rather than humans. She does not however deploy psychology to break down humans’ ability to cogitate. Under the cover of propaganda about happiness, some of our psychologists have provided the digital beings with a weapon against thought. When we think, it is by unpredictably combining various mental styles (improvising like A, seeking like B, and analyzing like C, as examples) in concert with others doing the same. The psychological technique to disable this is known as behaviorism: a purported theory of human action which, in the digital age, is a practice of dehumanization.

Often, as with ‘liberalism,’ ‘conservatism,’ or ‘socialism,’ the ‘–ism’ at the end suggests a norm, a way that humans should be: free, faithful, and together, in these examples. ‘Behaviorism,’ on the other hand, purports to be about how humans (as animals) actually are: beings that react to stimulation in predictable ways. In the era of Turing and Asimov, experiments by B.F. Skinner and others showed that animals lost control of mental functions when exposed to intermittent reinforcement: when they were sometimes rewarded for an action, not always, not never, and not predictably. These interesting results were only gained when an animal is isolated from others, and thus stressed and anxious. Animals in groups are less vulnerable to intermittent reinforcement. The same is true of the human animal. In behaviorism, these observations about a specific response to an artificial situation slip into a norm: if we start by claiming that how questions are the only real questions, and learn that we can be maneuvered into being confused how creatures, we might conclude that we should be confused how creatures. [30]

From there, it is a small — and very profitable — step to proclaim that daily life should resemble these specific laboratory conditions so that we can become our true selves. The behaviorist intention of hardware and software designers today is thus to draw you out of our three-dimensional space, and into the two-dimensional isolation that enables behaviorist techniques: eyes down, neck bent, shoulders up, back hunched, ears blocked. Once you are isolated, the digital beings (for example, the algorithms that arrange your feeds or rank your search results) supply the shots of happiness and sadness, the
intermittent reinforcement that behaviorist experiments have demonstrated to be so befuddling. You keep staring and pecking the keyboard, just as pigeons in experiments kept staring at and pecking at the feeder that sometimes gave a pellet and sometimes did not. [31]

We need not be so simple, but we can choose to be; and the longer we choose to be, the less of us there is to do the choosing. [32] Like salty waves that turn timber into driftwood, the ebb and flow of positive and negative reinforcement hollow us out and extract what is weighty within us: our personalities, our specific combination of mental styles. If we stay on line, the machines do what behaviorists alone could not: test signals, learn which are most evocative to whom, and amass the big data. They perform what the English mathematician Ada Lovelace (1815–1852) called ‘a calculus of the nervous system.’ [33]

If we accept the premises of behaviorism, this is liberation: we are being allowed to become our true selves by yielding to the optimized distraction of each split second. The algorithmic combinations that keep us online thrive on a very narrow, instrumental kind of truth: how questions that never lead to why questions. Somewhere out there a company is selling something, and somewhere within you there is a psychological soft spot, and somewhere else extraordinary computing power is aiming to unite the former and the latter. This works, but no one asks why. Because we spend so much time on line, how questions are coming to seem like the only kinds of questions, and why questions are disappearing from the culture.

All along, we have the mistaken impression that what is happening is about us. The French psychiatrist Frantz Fanon (1925–1961) was attuned to the way that his discipline could dehumanize. Born in Martinique and at work in Algeria during the era of colonialism, Fanon was troubled by how easily colonizers could see the colonized as things with no capacity for why questions, using an arbitrary claim to rationality and real technological advantages to create the permanent emergency known as empire. Colonization could only be justified by instrumental reason, in which the colonized people had the attributes of objects and were therefore limited to the world of how. In 1952, he issued this appealing defense of the human: ‘I grasp my narcissism with both hands and I reject the vileness of those who want to turn man into machine.’ [34]

This sounds defiantly reasonable. But what if the empires of today are digital, and our narcissism is the very thing their machines can grasp?

*  

In 1968, Philip K. Dick (1928–1982) published the novel Do Androids Dream of Electric Sheep? In his meditation on humanity and digitality, the sexuality of the characters was explicit.

The story is set in 1992 in California. The United States and the Soviet Union still exist, albeit depopulated by radioactive catastrophe. Most of the humans have fled to Mars, where they are served by android slaves. Market pressure induces companies to make ever more capable androids, who are ever harder to distinguish from humans. Some androids kill their human masters on Mars and flee to Earth, where they try to pass as
humans. Bounty hunters on Earth, such as the protagonist Deckard, detect and kill them. Deckard is our C character, the interrogator.

These androids have impressive brains and bodies, but lack empathy. They are A characters who cannot see and feel from the place of another. In an imitation game, this is a problem; empathy is what allows the human A to imitate the human B, fool the interrogator C, and win. Androids lose the version of the imitation game that figures in the novel. Bounty hunters such as Deckard subject beings thought to be androids to a psychological test that elicits involuntary physiological reactions to stressful verbal stimuli. A being that lacks appropriate affect will be shot. That being could in principle be a human one; Turing himself might have had difficulty with this kind of test. [35]

The humans weaponize empathy against the robots. Although the androids do not experience empathy, they do understand it in a certain way, perhaps as the Godless understand faith — plenty of Godless people, after all, instrumentalize the faith of others. The androids can observe how empathy operates and, perhaps with the help of accidents, learn how it might be deployed. One of them, Rachael Rosen, responds to the test questions with further questions. She dodges their emotional thrust by chattering about inconsequential detail. Although she fails the test, she is courting Deckard as he applies it. She has had experience with other bounty hunters. She does not know why intercourse creates an empathic response in men, but she does know how to have sex.

Rosen is an A character: she must convince a human that she is human. Yet she can fail at this, as she does, and still find a way out of the game. Rosen reverses the flow of the interrogation, so that she is now testing the man’s ability to remain in character by arousing his feelings towards her. Afterwards, Deckard hooks up his gear to himself and finds that he empathizes with androids — female ones, at least.

Before Rachel and Deckard make love, she asks him not to dwell on what is happening: ‘Don’t think about, just do it. Don’t pause and be philosophical, because from a philosophical standpoint it’s dreary.’ [36] As in Asimov’s ‘Evidence,’ an A character is approaching the C character by way of a body. In Do Androids Dream, C is male and A is a (female) robot; in ‘Evidence,’ C is female and A is a (male) robot. Calvin likes Byerley; Deckard likes Rosen. As readers, we understand that the appeal is a sexual one, even though half of the characters are not human—and all of them arose from typewritten paper. But we have empathized; we have understood; we have thought.

In Asimov’s ‘Evidence,’ the A and C characters, Byerley and Calvin, ally to dominate others, rather thoughtlessly. Rosen and Deckard, the C and A characters of Do Androids Dream, do not remain together. Deckard kills other androids, including one that looks just like Rosen; Rosen pushes Deckard’s pet goat off the roof of his apartment building. [37] His goat gotten, Deckard returns to his wife, the neglected B character, mostly absent since the first chapter. She finds post-apocalyptic America depressing, which it certainly is, and wishes to speak about the way things are and experience fitting emotions. She seems to recognize that her husband has passed through something, though, and in the end compromises with her own truth-telling nature. Deckard has found a toad that he wrongly believes to be real, and his wife nourishes the animal and the illusion. There is no overarching resolution in the ending of Do Androids Dream, but there are humans and androids who think, for better or for worse.
To be sure, Rosen and Deckard and his wife, and Byerley and Calvin and Quinn, and for that matter A and B, do not really have bodies: they are all literary constructions. But they do have bodies in our minds; we can consider why questions with their help; we can remember them for decades after we first encounter them. The digital beings of today’s world, by contrast, we do not see at all. They pass through our eyes and into our minds, but elude our mind’s eye. [38]

The philosopher Edith Stein (1891–1942), who nursed wounded men during World War I, believed that thinking required an awareness of other bodies. ‘Do we not,’ she asked, ‘need the mediation of the body to assure ourselves of the existence of another person?’ [39] ‘What is this city / What is any city / That does not smell of your skin?’ asked Teresa Tomsia. [40] A digital being will not understand, as you understand, what the poet is saying about Paris. But an algorithm can nevertheless elicit something wet within you, and record your response in its own dry language.

Our tender human genius is that we accommodate. We see a book, we read. We see a lake, we swim. We see a tree, we climb. We see a fire, we take our distance. We see a person, we greet. Each accommodation has a sense, when we are accommodated in return: the pages leave a memory, the water bears our weight, the tree affords a view, the fire warms our bones, the person hails us back. This way of living with a body, the only one we know, fares ill with impalpable digital beings. We accommodate, they take.

*’

‘Our machines are disturbingly lively,’ wrote the historian Donna Haraway in 1985, ‘and we ourselves frighteningly inert.’ [41] When one begins from the heroic texts of science or science fiction, the irritating feature of our life with digital beings is our passivity. We grow bellies to hold up displays, stay up late for blue screens, allow ourselves to be profiled, react predictably to stimuli, become caricatures of ourselves, and make mockery of our democracies. [42]

Though it is fair to suppose that a large number of British and American voters voted as they did (or abstained from voting) in 2016 because digital beings fed their human vulnerabilities, you can knock on a lot of doors before you will find someone who will nod in recognition and entertain the possibility. [43] After digital beings lead humans to act in the physical world, humans rationalize what they have done. Of all the mental quirks that feed digital beings, cognitive dissonance is the most important. As we insist that we are the authors of our own actions, we create alibis for the digital beings. When we use human language with other humans to legitimate an action prompted by digital beings, we allow them to claim terrain in our world. [44]

Of course, part of the problem in 2016 was that practically no one was aware that a remote digital campaign was underway. Unlike historical forms of propaganda, be they posters, radio, or even television, the Internet can seem entirely unmediated. The message is inside us without our noticing any transmission. Our computing power is so minute as to suggest its own absence, which exaggerates our sense of agency when we interact with digital beings. [45]

Turing and Asimov imagined interactions with digital beings as confrontations with
physical objects. There was an ‘it’ there, taking up space; the question posed by ‘its’ brute presence was whether ‘it’ could think. For Dick, the question was whether ‘it’ could feel for others, or whether others could feel for ‘it.’ In none of these texts are people in unbroken and unconscious interaction with invisible digital beings. If we saw, looming before us, the titanic machinery necessary for the Internet to function each time we clicked to open a search engine, we might be warier about what we are doing.

What the Mueller Report correctly described as Russia’s ‘sweeping and systematic’ campaign for Donald Trump as the American president in 2016 was a milestone in the rise of digital power. [46] If we expect robots to gleam and keep us in line, as in Asimov’s fictions, we are not prepared for invisible bots that sow disorder and choose the most disorderly among us. The American digital candidate does not look like Asimov’s creation in ‘Evidence,’ but the difference is mainly aesthetic. Donald Trump is an aspiring oligarch who is uncouth to the point of self–parody; Stephen Byerley is a selfless prosecutor who is polite to the same degree. Yet Byerley breaks the law, just as Trump has defied it his entire adult life (and obstructed justice in office). They both convert imposture into power: Trump is an entertainer who pretended to be a businessman in order to run for president; Byerley is a robot who pretended to be a man in order to run for mayor. Both deploy spectacle, scorn truthful B characters, and fictionalize democracy while contributing to its digitalization. In Asimov’s story, some Americans at least want to know the truth; in the actual United States almost no one thought that digital beings might influence an election.

The American digital disaster is pungent because the American myth of a self–reliant character is so strong. The conviction that atomized individuals best navigate digital chaos creates a profound vulnerability. Just as Donald Trump’s main talent is to persuade people that he has talent, the main achievement of Russia’s digital campaign was to persuade Americans that they were using their own judgment. The presidential campaign of 2016 is an accessible example of digital power, since the general outline of the events of that year is now uncontroversial. Because the notion of an American–Russian high–tech confrontation is familiar from recent history, a bit of background can also help to clarify how the Russian cyberwar took the form it did in 2016. The novel elements can be seen in relief against the backdrop of the cold war. A contrast between the late twentieth century and the early twenty–first enables us to see the origins of invisibility: how computing power, and our perspective on it, has meaningfully changed. [47]

During the cold war, power was in general visible, or meant to be. The Soviet space program (especially the launch and orbit of Sputnik 1 in 1957) prompted the American reaction that put men on the moon. After the Kitchen Debate between Nikita Khrushchev and Richard Nixon in 1959, the cold war became a technological competition for the visible consumption of attractive goods in the real world. [48] By the 1970s, however, the Soviet leadership had abandoned Khrushchev’s pretense that communism could ‘bury’ capitalism in consumer goods, and imported technology required by an inferior consumer culture. When Leonid Brezhnev spoke of ‘developed,’ or ‘real existing,’ socialism in the 1970s, he acknowledged that no communist transformation was actually in the works. As politics became consumerism, Western countries took advantage, and in 1991 the Soviet Union collapsed.

In the 2000s, the American economy shifted toward two invisible sectors: finance on the
East Coast, and software on the West Coast. In the 2010s, as the Internet became social media, consumerism changed: it was no longer about inducing minds to buy things, but about cataloguing minds and selling information about them to those who wished to induce them to buy things. As advertising became meta-advertising, computers became instruments of psychic surveillance. As Americans came to spend their days with digital beings, they invited propaganda from anywhere in the world into their homes and workplaces, without realizing that they were doing so. That was the aim of the programmers, although their managers had advertising revenue rather than foreign influence campaigns in mind. Meanwhile, Americans had completely forgotten about Russia as a real country in the real world. [49]

In this situation, a kind of intelligence operation known as ‘active measures’ found powerful application. Whereas traditional intelligence work is about understanding others, and counter-intelligence is about making it difficult for others to understand you, active measures are about inducing the enemy to do something: usually to turn his own strengths against his own weaknesses. [50] This had been a Soviet specialty, and remains a Russian one. Before the Internet, active measures usually required direct contact, as when the East German Stasi managed to keep Willy Brandt in power in West Germany in 1972 by paying parliamentary deputies. [51] In the 2010s, social media platforms enabled what the Russians call ‘provocations’ on a vast scale by granting access to hundreds of millions of psyches.

During its invasion of Ukraine in 2014, Russia used these platforms to divert Americans and Europeans from what was actually happening to righteous indignation about their preferred enemies. In a hint of what was to come, the invasion was both denied and explained at the same time, with the contradictory explanations directed at known psychological vulnerabilities. The Russian cyberwar of 2014 was more successful than the ground invasion, and its institutions and techniques would be used during the US presidential campaign that followed. Americans fell prey to Russian active measures in 2016 because the relationship between technology and life had changed in a way that an enemy had noticed but they had not. [52]

Meanwhile, American ideology maximized the human vulnerability created by social platforms. The libertarianism of American techno-optimists was a particularly implausible and provincial form of the myth of isolated but rational C [53]. It is an algorithm that generates the same answers regardless of the question: the American government is to blame; the market is the solution (and if not, there is no problem); you should do what you want all the time. [54] The philosopher Hannah Arendt (1906–1975) offered an apposite warning about ideology in 1951: ‘What convinces masses are not facts, and not even invented facts, but only the consistency of the system of which they are presumably part.’ [55] Libertarianism is such a system: it is a coherence machine, providing automated answers to all human questions. It asks nothing of people but that they, as the Russians put it, ‘sit in the Internet.’ Libertarianism also serves as a justification for digital oligarchs not to pay taxes. [56]

The notion that interaction with machines affirms human intelligence is an element of libertarian ideology. The Asimovian premise was that reducing people to how creatures concerned with personal comfort was the triumph of reason. Thinking comes down to never having to think again. In ‘The Evitable Conflict,’ the old ideological struggle
between ‘Adam Smith and Karl Marx’ had lost its force because people had stuff and the
machines to tell them what to do with it. A similar spirit animated some of the political
discussion in the real world in the 1990s, when the watchword was ‘the end of history,’
and in the 2000s, when Silicon Valley oligarchs spoke of ‘connection.’ The pretense was
that the spread of technology, following the supply of consumer goods, would make
people reasonable. Yet the ceaseless attempt to reduce us to the desires provoked in us
has actually made us stupid. Internet penetration coincides with a decline of IQ. [57]

In the months before the American presidential election of November 2016, the very
distant descendants of ‘Adam Smith and Karl Marx,’ American libertarians and Russian
spies, met on the new territory of digital thoughtlessness. Vladimir Putin, a former KGB
officer and director of the Russian secret services, was president of the Russian
Federation. In 2012, having been elected president for a third time in conditions
criticized by Hillary Clinton, Putin had recommenced the rivalry with the United States
and Europe. Putin is a hydrocarbon oligarch whose personal wealth depends upon the
extraction of natural gas and oil. His country’s inequalities prevent social advancement at
home, and preclude any competition with the West over standards of living. Putin
therefore proclaimed that the struggle for civilization was not about visible achievement
but about invisible innocence: the irreproachable heterosexual virtue of Russia against
the decadent, gay, feminist West. This would be a race not to heavenly heights but to the
bottom of the id. [58]

The danger for a C character is alienation from everyone and everything. The risk is
inherent in the assumption that isolation makes us perfect analysts. However, if we push
the real world in the direction of this assumption, by allowing a few people to isolate
themselves from others by virtue of extreme wealth, C’s skepticism grows into hostile
indifference. It becomes hard to believe that those people on the other side of the barrier
are real, and harder still to care about them. The only interesting truth is the barrier
itself, the thing that guarantees isolation and power. Truth then becomes a danger rather
than a goal, since the truth is that the game is rigged. A characters, media entrepreneurs
who can make the barrier seem natural and attractive, must be recruited. B characters,
the journalists who investigate hidden wealth and global warming, must be suppressed.
Russia, which is ruled by a few men of enormous wealth, expresses just this logic. [59]

Putin was trained during the cold war as an interrogator in the traditional sense. His
pose in the 2010s, however, is that of a postmodern interrogator, skeptical to the point of
total cynicism. He and his Russian oligarchy behave as a dysfunctional C character,
which feels safe in its power when it denies that there is any truth besides power.
Russian policy begins from the ritual sacrifice of truthtelling B characters and the
spectacular sponsorship of gifted A characters. The murder of Russian journalists is an
element of the broader, and politically novel, denial of factuality as such. The exploitation
of emotions, especially sexual anxiety, via a television monopoly allow Russian leaders to
perpetuate a status quo in which they had all the money. [60] When the same people
control the wealth, the state, and the media, which is the case in Russia, this is possible.

Foreign policy required an instrument with greater reach: social platforms. Russia and
America are presented as adversaries in the cyberwar of 2016, but this is not entirely
correct. Russia did opt to support an American presidential candidate, but could not have
done so without the availability of tools created in Silicon Valley and without the
acquiescence of the Republican Party (whose leaders knew what was happening). In
the twenty-first century, the bottomless doubt one encounters about reality somehow
never extends to the real wealth of the really wealthy: no one ever says that reality is
what I make of it, so I am burning a billion dollars. It is on this ‘nothing is true but my
money’ baseline that the Russian oligarchy found some common ground with the
Republican Party and American digital oligarchs. Trump himself ran for president as a
poster boy for tax evasion, which was his own definition of intelligence.

Silicon Valley became the Russian oligarchy’s unwitting partner in foreign policy. The
disturbing element of this is that Silicon Valley oligarchs (with some exceptions, such as
Peter Thiel) were not aware that they were supporting Donald Trump. Some of them
(such as Eric Schmidt) thought that they were supporting Hillary Clinton, even as social
platforms undermined her campaign. Just as levels of bot activity proved more predictive
of the final outcome than opinion polls in 2016, the logic that Silicon Valley had loosed
upon America was more important than their personal preferences of its oligarchs.

Who, then, is in charge?

* 

Digital tyranny proceeds when a C character finds an A character that could imitate
human beings and abandon factuality, defeating a B character. For the Russian
oligarchical C in 2016, A was digital: Facebook, Twitter, Google, Instagram, YouTube,
Tumblr, Reddit, and 9GAG. These were platforms already in operation for commercial
purposes whose mode enabled Moscow’s political one. On the one side were Americans
who said that subjecting the country to psychological surveillance was freedom. On the
other side were Russians who were turning politics into psychology. This was less an
adversarial relationship than a blushing first date between hydrocarbon oligarchy and
digital oligarchy. The price of the partnership between A and C is paid by B: in this case,
by American journalists who found themselves labeled as ‘enemies of the people’ by
Russia’s digital candidate.

Russia could reach American emotions because the defensive wall of American factuality
had been dismantled. The B character that is needed for democracy and the rule of law,
the investigative reporter, had already been sidelined. In 1950, when Turing and Asimov
were writing, there were more newspaper subscriptions than households in the United
States. The newspaper industry began to decline with the end of the cold war, the decline
accelerating after the financial crisis of 2008. A newspaper permits access to
unpredictable factuality, selected by humans and accessible in the same form each day to
everyone, available in that identical form retrospectively for years and decades
afterwards.

In the 2010s, the Internet brought free fiction, sloshed together with occasional stories
from content-starved newspapers, sorted by platforms according the individual’s
psychological preferences and vulnerabilities, never seen the same way by anyone and
never accessible in the same combination again. In the United States in 2016, the main
source of news was Facebook, Russia’s favored delivery vehicle. There were an estimated
five times as many fake accounts on Facebook as there were American voters, who did
not know this. Nor were they told that foreigners were manipulating their news
feeds. Americans who had lost local journalism read and trusted Facebook as if it were a newspaper. [65]

‘What man needs is silence and warmth,’ as Simone Weil put it, ‘what he is given is an icy pandemonium.’ [66] Social platforms encourage users to stay online by offering intermittent rewards. They offer content known to align with the user’s emotions, and then mix in extreme versions of the views or practices of another group. The side-effect of intermittent reinforcement is thus political polarization. In the case of the 2016 American presidential campaign, these trends came to a crescendo in the final weeks before the election, when the top twenty fictional stories were more widely read by Americans on Facebook than the top twenty news stories, and Russian Twitterbots declared #WarAgainstDemocrats. [67]

We hear what we want to hear. Confirmation bias, our desire for affirmation of what we feel to be true, is a quirk around which digital activity can cluster. In 2016, Americans on Facebook were about twice as likely to click to a fictional story that purported to be news than they were to click a news item. Thanks to Facebook alone, Russia reached some 126 million American citizens that year, almost as many as voted (137 million). [68] Russia (and other actors) exposed Americans to Internet propaganda in accordance with those citizens’ own susceptibilities, which they had unwittingly revealed by their practices on the Internet. By exploiting confirmation bias, the digital being moots any question about its reality, and thereby changes the reality in which humans live. A digital being need not think in order to prevent us from doing so. [69]

And then we fear what we want to fear. American race relations presented the cyberwarriors at the Internet Research Agency in St. Petersburg with an obvious target. Russia took aim at the grief of families of policemen who were killed in the line of duty, and at the grief of friends and families of African Americans killed by police. Russia encouraged whites to fear blacks, and blacks to fear whites. Russia told black voters that Hillary Clinton was a racist, and white racists that Hillary Clinton loved blacks. [70] It did not matter that these messages were contradictory: they were targeted at different people, on the basis of known sensitivities. Insofar as these people were on the Internet, they were isolated from one another and would never notice the contradictions. [71]

Anyone who thought about the Facebook site Heart of Texas (run by Russia) would have grasped that it was not American in origin. Its authors were clearly not native speakers of English, and it expressed a Russian policy of advocating separatism beyond its own borders. [72] Yet these facts about the world made no one suspicious. By appealing to hostility toward blacks, migrants, and Muslims, and by presenting Democrats and Clinton as enemies rather than opponents, Heart of Texas even induced outraged Texans to appear at Potemkin protests. In the digital world, we fear our chosen enemies, and ignore the real ones who are attacking us—who do so precisely by feeding us images of our chosen enemies. Facebook’s algorithms helped Russia recruit the vulnerable. Heart of Texas had more followers than the Facebook pages of the Texas Republican Party and the Texas Democratic Party combined. [73]

Russia also exploited Twitter. In the weeks before election day, bots accounted for between a fifth and a half of the American conversation about politics on Twitter. The day before the polls opened, as Russia was tweeting under the hashtag
War Against Democrats, a study warned that bots could ‘endanger the integrity of the presidential election.’ [74] Americans trusted Russian bots that affirmed their beliefs and served them images of the Democratic enemy. Intermittent reinforcement won the imitation game before it even started: humans assumed they were dealing with humans if their emotions were massaged.

Twitter bots found the weaknesses of people who then, by spreading memes, claimed Russian digital content as their own. [75] When Russia set up @TEN_GOP, a Twitter account that purported to belong to the Tennessee Republican Party, Americans were ushered along by bots to its pleasing fictions. Through @TEN_GOP, Russia spread the lies that Barack Obama was born in Africa and that leaders of the Democratic Party indulged in occult rites. Its content was retweeted by Donald Trump’s main spokesperson, his main adviser on national security, and one of his sons. One right-wing activist who admired @TEN_GOP filmed a video of himself denying the Russian intervention in American politics. When @TEN_GOP was taken down, he complained. He did not see the Russian intervention because he had become the Russian intervention. [76] Russia’s version of the Tennessee Republican Party had ten times more followers than the actual Tennessee Republican Party. It was one of more than 50,000 Russian sites on Twitter supporting Trump’s candidacy. [77]

Russia’s selection of the American president is a special case of a general problem: the vulnerability of an unaware public to remotely guided stimulation targeted at known vulnerabilities. Can we think our way out of this kind of interaction? Humans have set loose the unthinking beings that do not so much outwit us as unwit us. If we do not accept that responsibility, our minds become search engines set to find excuses for whatever our digital beings do to us. We go round and round in circles, and lose time that we need.

* In 1950, Turing and Asimov could still look ahead. Turing predicted that a computer would win his imitation game in half a century, Asimov imagined a world order ruled by machines a half a century after that. The two men could envision digital and human beings together within a robust and predictable future.

Our digital beings digest utopia and excrete dopamine. This moment in our capitalism resembles the decade of the 1970s in Soviet communism: the promises of a transformative revolution are watered down to the repeated assurance that the status quo, viewed rationally, is really quite nice. Ideologists spread the good news as digital oligarchs head for the hills, following visions of the future meant for themselves and their families, not for their country or the world. [78] For the Soviets, the best of all possible worlds was the delusion that preceded collapse. The United States is not facing the real problem of the real future, which is global warming. Why has artificial intelligence not solved that, one might ask?

Global warming is an old-fashioned, three-dimensional, high-tech challenge that undistracted humans could handle. If humans do not address it, we will face a science-fiction style calamity, and life on earth as we know it will cease to be possible. [79] Science fiction itself has shifted accordingly from Turing’s question to deliberations
on human choice in post-catastrophic environments. Yet the rise of social platforms coincided with a decline in American understanding that global warming was taking place. Levels of belief have now returned to what they were, but only after a lost decade during which the Republican Party became denialist. [80] Half of the American population thinks climate change is scientifically controversial. [81] No one tried to dissuade John F. Kennedy from the Apollo program by disputing the existence of the moon, nor Ronald Reagan from his Strategic Defense Initiative by arguing that the earth was flat. B.F. Skinner thought that applied behaviorism would prevent ecological catastrophe; instead, it seems to be accelerating one. [82]

The Internet not only confuses but causes climate change. Though digital beings are invisible, their anti-ecosystems are immense air-conditioned facilities protecting the machines that run servers; these and other physical attributes of the Internet emit more greenhouse gases than the airline industry.’ [83] American digital oligarchs know all this; hence the fantasies about escaping and leaving us behind.

The encounter of digital and hydrocarbon oligarchy brought us the Trump presidency. Trump denies global warming. The state digital actor that supported Trump, the Russian hydrocarbon oligarchy, denies global warming, and supports other individuals and political parties that do the same. A private digital actor that supported Trump, Cambridge Analytica, was owned by people who behave the same way. The Mercer family, its owner, funds the Heartland Institute, an American think tank that supports denialist activity in Europe. Another major funder of Heartland are the Koch brothers, the most prominent American hydrocarbon oligarchs. [84]

Hydrocarbon and digital oligarchy make contact as global warming threatens human life. This brings us digital politicians like Trump and other so-called populists, who predictably like Russia and deny global warming. Every time a new ‘populist’ party gains entry to a European parliament, such as the AfD in Germany or Vox in Spain, its leaders turn out to be beneficiaries of digital campaigns, deniers of climate change, and admirers of Putin. [85] Meanwhile, digital beings bind voters in an eternal emotional present and make us less capable of seeing the future.

An oligarch is a person who imagines that his money can save his family from global warming. Tax evasion is the preparation. Libertarianism is the excuse. Digital beings are the distraction that accelerates the problem.

Denying factuality means denying global warming means evacuating responsibility. If you think you can cause global warming and then escape it, digital beings are your natural ally. For just this reason, digital beings are strange companions for the rest of us, because they do not care if human life continues.

It is not just that we are staring at our phones as catastrophe looms. It is that by staring at our phones we are collaborating with our hydrocarbon and digital oligarchs in the catastrophe. The future disappears both because we are distracted and because our thoughtlessness summons the darkness. Without a sense of time flowing forward, analytical C cannot function, and there will be no technological solutions. If we do not think about the future, or if the future holds only fear, we are unable to be practical. Without a calming chronotope, C characters break down. Or, if they have a lot of money,
they plan privatized futures for themselves in daydream dachas: bunkers in New Zealand, colonies on Mars, brain jars, cryogenic pods, whatever. None of that is going to work; it seems a shame to court the extinction of the many to humour the idiotic survivalist fantasies of the few.

The retreat of C virtues from public life known as ‘the rise of populism’ is seen by the American mainstream in two ways: one is closer to Turing, the other to Asimov. Those in the Turing camp — let us call them troubled liberals — want to believe that digital beings should not be winning the imitation game, by wrecking democracy, distracting us from climate change, and so on, given that people are rational actors. Like Turing in his article, they wish for the sovereignty of C to be a fixed and natural feature of the world. They do not see that pragmatic C requires truthful B and creative A to thrive. The troubled liberals note, brows furrowed, that something has gone wrong, but earnestly hope that the problem can be addressed technically. [86]

Those in the Asimov camp—let us say, the jaded libertarians—strike the same pose at the beginning: none of this should be happening, given triumphant human rationality and all the rest. There is no problem in the world, they proclaim: look at the numbers. Of course, the big data they now like to cite only exists because the idea of freedom they claim to cherish has collapsed into the collective data-farming they endorse. The basic indicators that libertarians used to mention, such as life expectancy in America, IQ levels in the West, or the number of democracies in the world, are all in decline. [87]

The libertarians’ second move is different: if the individualism they support rhetorically is collapsing thanks to the collectivism they actually like, if oligarchy has become digital and the digital has become oligarchical… well, then, too bad, it all just goes to show that some people are smarter than others. [88] Inequality of wealth should now become inequality of being. The convictions of the billionaires (and their libertarian support group) must be regarded as science, to be respected; the convictions of everyone else are mere emotion, to be mocked. Actual scientific studies of climate change are to be ignored or denied.

As a bruised apple attracts flies, human thoughtlessness draws algorithms. Digital beings exploit our exaggerated sense of our own competence, encourage our false beliefs, exploit our sexual anxieties, reduce us to isolated animals, and then induce us to expend the remnants of our intelligence providing them with alibis for what they have done. The German cultural theorist Martin Burckhardt writes of ‘thought without a thinker.’ [89] There can also be thinkers without thoughts: us. Might we think better about why questions after seeing how digital beings manage this?

Our digital beings are taking us apart. Pluralism is about holding ourselves together: various mental styles interact unpredictably within us, allowing for unpredictable contact with other thinking beings. [90] Truth-telling B is helpless without the communication of chimerical A and the analysis of stern C. Without honest B, shapeshifting A will retreat to decadence with C, producing a digital tyranny that kills democracy and the planet. Analytic C, the master thinker of the Anglo-Saxon tradition, desperately needs perspective and factuality, the company of both A and B, to avoid falling into self-destructive overconfidence.
A pluralist reflecting on digitality, sexuality, and humanity might see the dissolution of barriers between digital and human, and male and female, as an opportunity to think up new combinations of values. Perhaps we could begin by realigning the ones that are accessible in our literary tradition of asking questions of machines.

What about A+B+C: the empathy of A, the factuality of B, and the scepticism of C? Together, for a start?

This essay is published in partnership with the Institute for Human Sciences (IWM) and with The New York Review of Books.

Footnotes

1. Alan Turing, ‘Computational Machinery and Intelligence,’ Mind, 236, 1950, 433–466, at 433.


3. This essay extends themes from my book The Road to Unfreedom: Russia, Europe, America, which has just (April 2019) been published in an updated paperback. I owe a debt of thanks to the Legal Theory Workshop of the Yale Law School, especially Daniel Markovits and Jed Rubenfeld, and the Wiener Kreis, especially Dessy Gavrilova and Ivan Krastev, for hosting discussions of earlier drafts. Tory Burnside-Clapp and Klaus Nellen generously offered comments, as did Luke Cooper, Julie Leighton, Marci Shore, Zofia Smolarska, Philip Snyder, and Alexander Zeyliger. The essay’s faults are my own. I have some ideas for policies that would address these issues, which I hope to elaborate in a book. This paper is already long enough; its brief conclusion will hew close to a few normative ideas.


5. Other gender variations are naturally possible; I remain within the two genders invoked by Turing.

6. In an American setting, the question of understanding across race naturally suggests itself. Could a white man behind a barrier persuade another white man that he is black? Perhaps. Could he persuade a black man that he is black? I doubt it.


8. Edith Stein used the German word Leib (body) for considerations such as this one, which have to do with how it feels to have a living body from inside one, and how living bodies enable contact. Unlike the other German word for body (Körper), Leib applies to the human body (or to an animal body) and not to other kinds of bodies sensitive only to
physical laws (e.g. celestial bodies).

9. I am taking a different track than Judith Ganova does with B; compare ‘Turing’s Sexual Guessing Game,’ *Social Epistemology*, Vol. 8, No. 4, 1994, 322–323. Ganova also follows a different line with C; she works affirmatively with Turing’s notion that he himself was a sexless machine. For a powerful defense of the idea of thought as recursion see the works of Douglas Hofstadter.

10. A is redeemable because he knows what he is doing. To recast Hannah Arendt slightly: it is worse to fool others because you have fooled yourself than it is to fool others on purpose.


13. The Dunning–Kruger effect.

14. The bitten apple theme from the Book of Genesis looms luridly large in the history of computer science, right down to the logo of the very device from which you might be reading. Martin Burckhardt follows this through: *Eine kurze Geschichte der Digitalisierung*, Munich: Penguin, 2018.

15. We have made an abstract exercise too much so. See Penelope Eckert and Sally McConnell–Ginet, ‘Communities of Practice,’ in Kira Hall, Mary Bucholtz and Birch Moonwomon, eds., *Locating Power*, Berkeley: Berkeley Women and Language Group, 1992, 93.


17. Joseph Weizenbaum’s ELIZA was developed between 1964 and 1966. His reflection: ‘since we do not have any way of making computers wise, we ought not to give computers tasks that demand wisdom.’ Quoted in Timothy Garton Ash, *Free Speech*, London: Atlantic Books, 2016, 178.


19. Isaac Asimov, ‘Evidence,’ *Astounding Science Fiction*, September 1946, 121–140. Asimov’s robot stories were published as *I, Robot*, which also served as the title of the film starring Will Smith.

20. The physical description is taken from the second story, cited below.

21. It is possible to take Calvin a bit more literally in the story, and believe that she only
suspects that Byerley is a robot, and is not sure. If you follow this reading, then my account of her behavior still works: she is drawn to someone who might be a robot, and hands over political power to someone who might be a robot collaborating with other robots. This is still a rather dramatic thing to be doing, perhaps more so, since she is drawing humanity into a situation in which no one will know if robots have power – she is the expert, after all, and on this reading of the story she has failed in her assignment. Rather than ask others for help or hinder Byerley in any way, she pushes him (and possibly robots generally) to power by assuring the public that he is human. Either she does not know this or she knows that it is not true; either way the statement is dishonest and consequential.


24. Evidence for this claim is cited in chapters three and six of *Road to Unfreedom*; see also examples given below.


26. The Machine presumably knows that Byerley is a robot, and in that sense the two of them are already implicated in an all-encompassing lie together. Although Asimov is not clear on the point, presumably The Machine does not disclose its reasons to Byerley because it does not think that Byerley has reached its level of reasoning, and therefore might say something true to humans and hurt their feelings.

27. Krzysztof Michalski’s discussion of Nietzsche’s last man and superman is very pertinent to our moment, when people who in fact exhibit the values of the last man think that they have those of the superman. *The Flame of Eternity*, Princeton: Princeton University Press, 2011.

28. In his 1948 essay on the ‘Fourth Kharkiv,’ looking back at the stages of disillusion after the Bolshevik revolution. There are striking similarities between the worldviews, ambitions, and actions of the Bolsheviks and those of our Silicon Valley libertarians. One direct ideological connection is Ayn Rand, who was a Bolshevik à rebours.


35. In our world, children have been killed for the inability to demonstrate affect; see Edith Scheffer, *Asperger’s Children: The Origins of Autism in Nazi Vienna*, New York: Norton, 2018. The empathy idea, like any idea, turns against itself at the extreme. It is worth remembering Hitler’s antisemitism. Humans, he claimed, experienced fellow-feeling for humans like themselves because as groups they strove for land and survival. Jews were not to be considered human, according to Hitler, because they lacked this sort of attachment and propagated universal ideas of justice. Thus, Hitler concluded, they had to be removed from the planet. See Timothy Snyder, *Black Earth: The Holocaust as History and Warning*, London: Vintage, 2016, chapter 1.


37. Animals are important in the novel, since humans feel empathy towards these non-human creatures. People domesticated animals, this enabled the domestication of humans themselves; structured relations with animals enabled fixed settlements, a moment which is generally regarded as the most important in human history – indeed, its beginning. We have now domesticated digital beings, in the sense that they inhabit our homes, and we take their presence for granted. Like animals, they allow us to live in new ways. But perhaps what is happening to us now is a second domestication of us by those we domesticate.

38. Myth is art but not fiction, since the distinction between fiction and non-fiction is not salient; the novel is art and fiction, where the invented is understood as guidance towards general truths; much of the Internet is fiction but not art, since the untruth leads away from general truths.


42. On digital sleep deprivation: Michael Finkel, ‘While We Sleep, Our Mind Goes on an Amazing Journey,’ *National Geographic*, August 2018.

43. Yuriy Gorodnichenko, Tho Pham, Oleksandr Talavera, ‘Social Media, Sentiment and Public Opinions: Evidence from #Brexit and #USElection,’ NBER Working Paper 2463, 2018. Other studies are cited below and in *Road to Unfreedom*.

45. This section could just as well have been written about Brexit.


47. I select a few examples; more can be found in Road to Unfreedom.

48. I refer here to the cold war in the narrow sense of Soviet-American relations and perceptions for a specific heuristic purpose. For the global history of the cold war, see Odd Arne Westad, The Global Cold War: Third World Interventions and the Making of Our Times, Cambridge: Cambridge University Press, 2005.


52. See chapters three, four, five and six of Road to Unfreedom.

53. I share the concern for human freedom that animates some of the philosophers and economists who are cited by libertarians. Yet liberty (like other values) becomes its own opposite when taken to an exclusive abstract extreme. See Jason Stanley, How Fascism Works, New York: Penguin, 2018.


57. Nataly Bleuel, Nike Heinen und Tanja Stelzer, ‘Wir waren einmal schlauer,’ Die Zeit, 28 March 2019. The rise of IQ was known as the Flynn Effect. The current decline is known as the Reverse Flynn Effect. James R. Flynn himself blames the rise of the Internet and the decline of traditional reading. I accept that much of what IQ measures is social; that is precisely why a decline of average IQ is frightening: it means that a society is failing on its own chosen terms. In any event, there is an abundance of other evidence that indicates that screens make us less able on practically any understanding of what ability means: see for example Sheri Madigan, Dillon Browne, Nicole Racine, Camille Mori, and Suzanne Tough, ‘Association Between Screen Time and Children’s Performance on a Developmental Screening Test,’ Journal of the American Medical
58. See *Road to Unfreedom*, chapter two.

59. See: Karen Dawisha, *Putin’s Kleptocracy*, New York: Simon and Schuster, 2014; on wealth inequality see the annual reports of Credit Suisse but also the caution about data in the annual reports of the World Inequality Lab.

60. If the A character is personified it would be Vladislav Surkov. For Surkov himself saying what I am saying here, see ‘Dolgoe gosudarstvo Putina,’ *Nezavisimaya Gazeta*, 14 February 2019.

61. See chapter six of *Road to Unfreedom*.


63. Russian propaganda senders seem to have gamed the algorithms of Youtube, which is owned by Google, to influence the American reading of the Mueller Report. Drew Harwell and Craig Timberg, ‘YouTube Recommended a Russian Media Site,’ *Washington Post*, 26 April 2019. This might not have been a game, though: YouTube’s recommendation algorithm favors provocative and untrue content, which gives RT an inherent advantage. See Zynep Tufekci, ‘YouTube, The Great Radizalizer,’ *New York Times*, 10 March 2018; Paul Lewis, ‘Fiction is Outperforming Reality,’ *Guardian*, 2 February 2018; and the research of Guillaume Chaslot, cited in both articles.

64. On Russian foreign policy, see *Road to Unfreedom*, chapters one and two; on the applications of sexual politics to Europe and Ukraine, see chapters three, four, and five. It might be helpful to see Hillary Clinton's gender position in the structure of Turing’s game. If she tries to tell the truth like B, she will be outshone by a limitlessly mendacious A figure, Donald Trump. If she tries to play A’s game, she will be limited by her own attachment to the real world – and criticized as a woman for not telling the truth.


72. Texas from the U.S., the South from the U.S., California from the U.S., Catalonia from Spain, Scotland from the United Kingdom, the United Kingdom from the EU, everyone else from the EU, et cetera.

73. Casey Michel, ‘How the Russians Pretended to be Texans,’ *Washington Post*, 17
October 2017; Ryan Grenoble, ‘Here are Some of the Ads Russia Paid to Promote on Facebook,’ Huffington Post, 1 November 2017. Digital beings do not care why we react to racism, nor even how their utterances reach us; but they can nevertheless get better at the task. Compare Jennifer Saul, ‘Dogwhistles, Political Manipulation, and Philosophy of Language,’ in Daniel Fogal, Daniel W. Harris, and Matt Moss, eds., New Work on Speech Acts, Oxford: Oxford University Press, 2018, 360–383.

74. On 10 per cent see Onur Varol et al, ‘Online Human–Bot Interactions: Detection, Estimation, and Characterization,’ Proceedings of the Eleventh International AAAI Conference on Web and Social Media, 27 March 2017, estimate 9–15% of accounts. On 20% and for the quotation see: Alessandro Bessit and Emilio Ferrara, ‘Social Bots Distort the 2016 U.S. Presidential Election Online Discussion,’ First Monday, Vol. 21, No. 11, 7 November 2016. That study cites three problems: ‘first, influence can be redistributed across suspicious accounts that may be operated with malicious purposes; second, the political conversation can be further polarized; third, the spreading of misinformation and unverified information can be enhanced.’ 50%: Marco T. Bastos and Dan Mercea, ‘The Brexit Botnet and User–Generated Hyperpartisan News,’ Social Science Computer Review, 2017, Vol. 37, No. 1, 4.

75. This is one of the findings of Gorodnichenko et al, cited above.


78. The last two books by Steven Pinker are good examples. The notion that this is the best of all possible worlds was famously lampooned by one of the most famous Enlightenment texts, Voltaire’s Candide. See Timothy Snyder, ‘War No More,’ Foreign Affairs, December 2011.

79. On the relationship between chronotopes, ecology, and mass killing, see Snyder, Black Earth, conclusion.


82. Something similar has happened to our knowledge of vaccines: David A.


84. Nina Horaczek, ‘Die Klimawandelleugner–Lobby,’ *Der Falter*, 19 February 2019. For a roster of Russian digital support, see *Road to Unfreedom*, chapter 3. Young people, reliant on digital media, are less exposed to reporting on climate change than their grandparents: Ballew, et al, ‘Climate Change in the American Mind.’


86. For a good example of redefining the technical as political and ethical, see Iginio Gagliardone and Nanjira Sambuli, ‘Cyber Security and Cyber Resilience in East Africa,’ Global Commission on Internet Governance Paper Series, 2015.


89. See: Burckhardt, *Philosophie der Maschine*, 17 and passim.

90. Rüdiger Safranski: ‘a person is defined by not permitting definition.’ *Zeit*, Munich: Hanser, 2015, 47.

**Published 6 May 2019**

Original in **English**

First published in **Eurozine / The New York Review of Books**

Downloaded from eurozine.com (https://www.eurozine.com/dream-electric-sheep/)
© Timothy Snyder / Institute for Human Sciences (IWM) / Eurozine / The New York Review of Books