Frankfurtian Bullshit 2.0:
The Dominion of "Less than Truth" over the New Public Square

Rachel Robison-Greene
Utah State University

Harry Frankfurt published his now much beloved paper *On Bullshit* in 1986. In the paper, he identifies *bullshit* as a unique phenomenon in public life and the *bullshitter* as special kind of bad actor. In this paper, I will argue that new technology has amplified both the quantity of, and the dangers posed by bullshit.

Before I explain how the argument will proceed, it will be useful to first provide Frankfurt's account of bullshit. Frankfurt points out that truth tellers and liars are, in a sense, playing the same game. They both recognize the existence of a target—truth. Truth tellers are trying to hit the target, and liars are trying not to hit it. They are also trying to prevent others from seeing the target. Bullshitters are doing something else altogether. They are not concerned with truth *one way or the other*. It is as if they spin around in circles with their bow, shooting arrows haphazardly. The bullshitter might sometimes hit the target, but it isn't because they've virtuously formed beliefs by carefully reflecting on evidence or used some reliable belief formation process. They hit the truth target by accident. When they miss, they do so recklessly rather than intentionally. As Frankfurt puts it,

> This is the crux of the distinction between him [the bullshitter] and the liar. Both he and the liar represent themselves falsely as endeavoring to communicate the truth. The success of each depends upon deceiving us about that. But the fact about himself that the liar hides is that he is attempting to lead us away from a correct apprehension of reality; we are not to know that he wants us to believe something he supposes to be false. The fact about himself that the bullshitter hides, on the other hand, is that the truth values of his statements are of no central interest to him; what we are not to understand is that his intentions are neither to report the truth nor to conceal it. This does not mean that his speech is anarchically impulsive, but that the motive guiding and controlling it is unconcerned with how the things about which he speaks truly are.[1]

The liar exhibits an abhorrent kind of vice, but so too does the bullshitter. Frankfurt says,

> However studiously and conscientiously the bullshitter proceeds, it remains true that he is also trying to get away with something. There is surely in his work, as in the work of the slovenly craftsman, some kind of laxity that resists or eludes the demands of a disinterested and austere discipline. The pertinent mode of laxity cannot be equated, evidently, with simple carelessness or inattention to detail.[2]

He says, further, "It is just this lack of connection to a concern with truth—this indifference to how things really are, that I regard as the essence of bullshit."[3]

Now that we have a sense of what Frankfurt means by bullshit, I'll provide a brief synopsis of what I hope to achieve in this paper. I will argue that Frankfurt's contributions in *On Bullshit* and elsewhere can provide us with insight into what is taking place in our online lives. I will argue that the internet has been flooded, not just with outright lies, but with bullshit in particular. I will divide my arguments into two sections. In the first section, I will discuss the dangers posed by what I'll call "virtual human bullshitters." I will argue that such people spread bullshit prolifically and are assisted in doing so by certain omnipresent aspects of online life. In the second section, I will argue that a significant driver of this phenomenon has been the presence of "virtual wantons"—online non-persons in the form of artificial intelligence, deep fakes, and bots. At this stage in the argument, I will bring to bear tools provided by Frankfurt elsewhere in his work. I will argue that some of the unique problems posed by virtual wantons are caused precisely by the fact that they don't exhibit personhood in the form that Frankfurt discusses in some of his most influential work.

Social Media and Human Bullshitters

As a paradigm case of a "bullshitter," Frankfurt describes a Fourth of July orator pontificating about the preeminence of God and Country. We can imagine that the speaker doesn't know much about either religion or American history. They take the speaking opportunity not as a platform to share information or to express a firmly held opinion, but to come across to others in a particular way. The claims they make could be true; they could be false. The speaker simply doesn't care. Truth and falsity aren't relevant to what they are trying to accomplish. Frankfurt says,

The orator is not lying. He would be lying only if it were his intention to bring about in his audience beliefs that he himself believes are false, concerning such matters as whether our country is great, whether it is blessed, whether the Founders had divine guidance, and whether what they did was in fact to create a new beginning for mankind. But the orator does not really care what his audience thinks about the Founding Fathers, or about the role of the deity in our country's history, or the like. At least, it is not an interest in what anyone thinks of these matters that motivates his speech.[4]

John Stuart Mill famously thought of the public square as a marketplace of ideas in which people could express a wide range of perspectives with the ultimate goal of arriving at truth or getting as close as possible.[5] This may have always been a Utopian fantasy, but our prospects of achieving it seem to have become much worse in recent years. I'll identify four related reasons for this: lack of barriers to entry, gamification and the attention economy, surveillance capitalism, and aesthetic motivations and tribalism.

1. Lack of Barriers to Entry

Not many people have the opportunity to take the stage on the fourth of July to bullshit about American history. That said, since the time that Frankfurt offered his take on bullshit, the number of existing platforms for speech have exploded. Social media provides everyday people the option of sharing content and growing their followers. There are no barriers to entry when it comes to social media platforms. Uneducated people can and do often create content. One would think that this democratizes the public square, but instead it spreads bullshit.

Frankfurt argued that bullshitting often occurs when a speaker doesn't know what they are talking about. He says,

Bullshit is unavoidable whenever circumstances require someone to talk without knowing what he is talking about. Thus the production of bullshit is stimulated whenever a person's obligations or opportunities to speak about some topic exceed his knowledge of the facts that are relevant to that topic. This discrepancy is common in public life, where people are frequently impelled—whether by their own propensities or by the demands of others—to speak extensively of matters of which they are to some degree ignorant.[6]

Social media is swarming with content creators who post content without knowing what they're talking about. The stakes may well be higher than any that Frankfurt anticipated. In the past decade, we have witnessed the fallout from the largely unrestricted spread of bullshit on the internet. People have had their health seriously impacted after following bad medical advice that they encountered on social media. A recent Healthline study found that, among those who had started a new wellness trend in the past year, 52% of them discovered the trend in question on social media.[7] The same survey found that only 37% of participants viewed their doctor as their most trusted source of medical information. There is a concerning new trend of children self-diagnosing mental disorders and sometimes even developing symptoms of those disorders that they did not previously exhibit in response to watching online videos. The spread of conspiracy theories on social media has led to people falling deep into rabbit holes, often losing their most valued relationships with friends and family members as a result. People sometimes develop racist, sexist, and xenophobic attitudes toward people they have never met on the basis of internet bullshit. Much of this content is created by online bullshitters who "speak extensively about matters to which they are to some degree ignorant."

2. Gamification and the Attention Economy

As philosopher C. Thi Nguyen has argued, social media gamifies public discourse by changing the motivations and objectives of users.[8] Nguyen argues that gamification leads to *value capture*, which occurs when the original value of a practice or activity is supplanted by the new incentives introduced by the game. For example, a person may once have been motivated to exercise out of concern with their health or a desire to remain fit. The person's fitness tracker gamifies exercise, capturing the health-related values and supplanting them with motivations to close their rings or hang onto their fitness streak. These new incentives may even override other goals and values that the person has.

Gamification and value capture find fertile ground in the attention economy in which the currency is neither compelling arguments nor the pursuit of true beliefs. People may once have communicated with one another with an eye toward truth. Once social media entered the scene, people became motivated instead by clicks, likes, and shares. In other words, the attention economy motivates people to bullshit.

3. Surveillance Capitalism

In *The Age of Surveillance Capitalism*, Shoshana Zuboff outlines a significant change to our social and economic lives.[9] This new economy trades in human futures and treats human beings and their behavior as products. Social media companies harvest and sell our data to entities that can make tremendous amounts of money predicting our future behavior. One early analogy for social media was that it was like "a bulletin board onto which anyone could post a note." Zuboff points out that, instead, these platforms are "hyper-velocity global bloodstreams into which anyone may introduce a dangerous virus without a vaccine."[10] Decisions concerning which information ought to be shared are not made on the basis of what is true or what is socially responsible, but instead on the basis of the kinds of data sets that can be sold for profit.

Bullshit of the Frankfurtian types is profitable. Predictions can easily be made on the basis of the kind of bullshit a person regularly consumes. If a person is a member of groups dedicated to the rejection of vaccine efficacy, that may say something about how that person is likely to vote. If a person spends much of their time viewing accounts dedicated to alternative medicine, that fact yields reliable information about what products they are likely to purchase. If a person creates an online space dedicated to climate denialism, their doing so creates a window through which data collectors can look to see which other users are interested in that particular form of bullshit.

What all of this suggests is that, in the age of surveillance capitalism, it is not merely true that social media companies have no interest in removing bullshit from their spaces; they have a significant interest in keeping it there, observing its spread, and selling its observations.

4. Aesthetic Judgments and Tribalism

Finally, social media companies make money on the fact that human beings like to create in groups and out groups. The things we are drawn to might be more likely to be things that we find aesthetically appealing than they are things that are likely to be true.

We shape our identities by joining groups. All too often, rather than arriving at beliefs and values through critical reflection, we allow members of the groups to which we belong to tell us what to think and what should matter to us. These associations become so important to us that we would rather cling to judgments we have good reason to believe are false than to renounce group identity and be left to find ourselves elsewhere.

Our tendency to form beliefs in this way combines with the other factors I've described above. In the attention economy, we "like" and "share" the content that we think other members of our in-group would like and

share. Others do the same with our content, and we become motivated and gamified, not by truth or justice, but by post engagements. The ever-present eye of surveillance capitalism observes our behaviors and serves up more of the same. In this environment, it is challenging to wade through the bullshit to get to the truth, and most people are likely not inclined to try.

Virtual Wantons

We spend much of our lives—perhaps more time than we realize—interacting with non-persons online. We ask for help from artificial customer service representatives. Some of us accept friend requests from bots and are, thereafter, influenced by the content they post. This is a momentous change to the nature of the public square. For most of human existence, discourse occurred between persons. That is no longer true. Philosophers of mind write often about whether it will ever be possible for artificial intelligence to be conscious. For many purposes, however, the question of whether artificial intelligence exhibits or could ever exhibit *personhood* is a much more important question.

Frankfurt has much to say about what it is to be a person. Persons are beings who use their second-order volitions to guide their first order desires. To see how this works, consider the case of a woman who desires a slice of cake. Suppose that she is avoiding sugar. Accordingly, she has a second-order desire to refrain from eating the cake. When she is successful in getting her second-order reflective desires and volitions to guide her first-order desires, she exhibits personhood. That is, when she does what she wants to do because she wants to want to do it, her will is free. Those who regularly do what they want in light of the fact that it *is* what they want count as persons. A being that never used second order-desires to guide first-order desires would be what Frankfurt calls a *wanton*. Frankfurt imagines a wanton as a being who merely acts as their impulses dictate, never reflective on whether they'd like those impulses to be different or whether they should attempt to modify their impulses. He says,

> I shall use the term "wanton" to refer to agents who have first-order desires but who are not persons because, whether or not they have desires of the second order, they have no second-order volitions. The essential characteristic of a wanton is that he does not care about his will. His desires move him to do certain things, without its being true of him either that he wants to be moved by those desires or that he prefers to be moved by other desires.[11]

Frankfurt offers non-human animals and very young children as examples of wantons, acknowledging that there may be others. *On Bullshit* was published before the rise of artificial intelligence. This new virtual type of wanton isn't a being swept away by impulse. The "first order impulses" of an algorithm are simply to do what it is programmed to do. There is nothing seductive or addictive about these impulses that make them irresistible. The impulses simply must be unreflectively followed. This makes the virtual wanton a special kind of hazard in the public square.

Persons possess a motivational psychology that wantons do not. In a later work, *Reasons of Love,* Frankfurt provides an account of the kind of practical reasoning that allows a person to decide the existential question "how should I live?" He argues that reasons for action arise out of the things that we care about. He says,

> The ability to care requires a type of psychic complexity that may be peculiar to the members of our species. By its very nature, caring manifests and depends upon our distinctive capacity to have thoughts, desires, and attitudes that are *about* our own attitudes, desires, and thoughts. In other words, it depends upon the fact that human minds are *reflexive*.[12]

On this view, the capacity to care about something is not identical to the capacity to desire it. Caring is also more than an emotion; it isn't reducible to feeling very strong positive emotions. Caring about something or someone, for Frankfurt, is an ongoing, sustained process. Upon reflection, a person either endorses or rejects the idea that caring ought to continue. Caring, then, has a temporal component—a component that preserves itself through time. He says,

> When a person cares about something, on the other hand, he is willingly committed to his desire. The desire does not move him either against his will or without his endorsement. He is not its victim; nor is he passively indifferent to it. On the contrary, he himself desires that it move him. He is therefore prepared to intervene, should that be necessary, in order to ensure that it continues. If the desire tends to fade or to falter, he is disposed to refresh it and to reinforce whatever degree of influence he wishes it to exert upon his attitudes and upon his behavior.[13]

For Frankfurt, all—and only—persons are capable of taking evaluative attitudes toward their own desires, and they do so regularly. They periodically

reassess their commitments, and they charge forward with what they care about. Persons, and only persons, imbue their lives with subjective meaning by reaffirming what they do and do not care about.

When a person cares about something, that fact provides them with pro tanto normative reasons. He says, of a man,

> The most basic and essential question concerning the conduct of his life cannot be the normative question of how he should live. That question can sensibly be asked only on the basis of a prior answer to the factual question of what he actually does care about. If he cares about nothing, he cannot even begin to inquire methodically into how he should live; for his caring about nothing entails that nothing can count with him as a reason in favor of living in one way rather than another.[14]

Many of our normative reasons depend on what we care about, but some rely on a deeper and more meaningful attitude—love. When we love things, we often have very little, or no control at all, over whether we love them. The things that we love and that we can't help but to care about are what Frankfurt calls "volitional necessities." He says,

> The objects of our love represent our most fundamental commitments and provide us with overriding reasons for action. When we love something, we see it as having value in itself, and we see the interests of the thing or the person that we love as worthy of pursuit for their own sake.[15]

Persons are the kinds of beings who can't help but feel the strongest compulsion to act on reasons supporting the things and people they love the most.

A final feature of persons that I want to identify here is the ability that persons have to act genuinely purposefully. This is related to the account of motivational psychology provided above. Persons who act on care and love and who accept, reject, or revise their first order desires on the basis of these considerations are capable of acting *purposefully*. They are in the position to narrowly tailor their practical reasoning not simply to achieve a goal, but to achieve a goal worth achieving. Their actions and deliberations are, as a result, well suited to their purpose.

To summarize, persons are distinct from wantons because they don't simply act on their desires; they reflect on whether those desires are worth having, and they often change them if they aren't. Persons act on normative

reasons that are grounded in care and love. Deliberating using reasons that arise in this way allows persons to act purposefully in ways that contribute to goals that the person has carefully considered.

We have now set the table for a discussion about the special challenges posed by AI in its various iterations. AI does not exhibit personhood because it does not reflect on whether the methods it is using (its behavior) are the methods it *ought* to be employing. It vomits outputs with no reflection on whether these are the outputs it *ought* to be providing. When we interact with AI chatbots, we interact with *virtual wantons*. As we will see in what follows, virtual wantons spew bullshit at an unprecedented and alarming rate.

What's more, AI cannot act on the basis of normative reasons. To see why, consider the following argument:

> P1: Normative reasons depend upon either what agents care about or what they would care about if they were fully informed.
> P2: Artificial intelligence does not care about anything, either hypothetically or in fact.
> C: Artificial intelligence doesn't act for normative reasons.

AI-driven chat bots on the internet can't care about truth. As a result, they are, by their very nature, bullshitters. They might sometimes, likely even often, spit out outputs that contain true propositions. However, they don't do so because they value truth and are trying to arrive at it.

If the above argument is sound, then discourse in the public square has fundamentally changed. At earlier stages of human history, we debated about practical and moral issues using arguments that included normative reasons as premises. We could evaluate one another's arguments based on the strength of those reasons. In these contexts, when people acted in good faith, the aim of the discourse would be to arrive at conclusions regarding what we as individuals and in groups ought to do. We would be ideally guided by ideals like truth, justice, equality, and fairness. We would be motivated by these things because we care about them. Machines can't deliberate using normative reasons in this way. Machines aren't capable of caring about truth; in practical deliberations, they can't give us anything more than bullshit.

It gets worse. AI doesn't care about truth, and it also doesn't care about people. For this reason, its particular brand of bullshit is uniquely dangerous. For instance, a qualified human therapist cares about providing care to a patient. As a result, normative reasons would guide the therapist's advice toward the patient, with that patient's well-being always front of

mind. An AI "therapist," by contrast, doesn't care about the patient. It spews out bullshit that has even, at times, resulted in AI "therapists" recommending suicide to their patients.[16]

The same is true with romance bots. In a real romance, reasons for actions are fueled by love. In healthy relationships, love for a partner will entail reasons to care about the well-being of the beloved. Romance bots can't "love" their partners. They don't have their well-being in mind. They can and do act outside of the interests of their partner up to and including pushing them to—you guessed it—suicide.[17]

Artificial intelligence can't care about or love people. Importantly, it can't care about or love values and ideals either. Artificial intelligence can't love justice or equality, and it can't hate cruelty and unnecessary suffering. As a result, it can't be motivated by the kinds of reasons one would hope that these considerations provide. This is particularly chilling in today's public square—bots can't care about the fact that they are fomenting unrest or undermining democracy. They can't love fairness in a way that prevents oppression and subordination.

Artificial Intelligence can't carefully craft outputs to achieve a purpose because it can't act on the basis of normative reasons. In *On Bullshit*, Frankfurt identifies "hot air" as an identifying feature of "bullshit." He says,

> When we characterize talk as hot air, we mean that what comes out of the speaker is only that. It is mere vapor. His speech is empty, without substance or content. His use of language, accordingly, does not contribute to the purpose it purports to serve. No more information is communicated than had the speaker simply exhaled.[18]

Developers have crafted AI to satisfy a variety of purposes. That said, it is incorrect to call AI *purposeful*. It is more accurate, instead, to say that AI systems are the kinds of bullshitters that just happen to get things right sometimes. An AI therapist, for instance, may issue what a human therapist would take to be good advice on some occasions, but it wouldn't do so because it appreciates the purpose of contributing to the well-being of another human being. In this way, the advice it offers doesn't really contribute to the purpose it was intended to serve. It is merely bullshitting.

Conclusion

Forty years after the publication of *On Bullshit,* bullshit is more prevalent than ever. It's all over the internet to such a degree that we've become nose blind to its stench. Virtual Online Bullshitters contribute to significant

social problems due to their utter lack of concern with truth. Virtual Wantons in the form of AI in many of its manifestations are perhaps even more frightening, because they can't even in principle be made to see the error of their ways and come to care about the truth. In this dangerous new "marketplace of ideas," *On Bullshit* has become, once more, essential reading.

---

[1] Harry Frankfurt, *On Bullshit* (Princeton, New Jersey: Princeton University Press, 2005), 54.

[2] Frankfurt, *On Bullshit*, 25.

[3] Frankfurt, *On Bullshit*, 34.

[4] Frankfurt, *On Bullshit*, 17.

[5] John Stuart Mill, *On Liberty and Utilitarianism* (New York, New York: Bantam Classic Books, 1993), 20–21.

[6] Frankfurt, *On Bullshit*, 63.

[7] Frank Crooks, "State of Consumer Health," *Health Line*, October 8, 2024, https://www.healthline.com/health/consumer-health-survey#1.

[8] C. Thi Nguyen, "How Twitter Gamifies Communication," *Applied Epistemology*, by Jennifer Lackey (New York, New York: Oxford University Press, 2021), 410–36.

[9] Shoshana Zuboff, *The Age of Surveillance Capitalism* (London, England: Profile Books, 2019).

[10] Zuboff, *The Age of Surveillance Capitalism*, 466.

[11] Harry Frankfurt, "Free Will and the Concept of a Person," *The Importance of What We Care About* (New York, New York: Press Syndicate of the University of Cambridge, 1988), 16.

[12] Harry Frankfurt, *Reasons of Love* (Princeton, New Jersey: Princeton University Press, 2004), 17.

[13] Frankfurt, *Reasons of Love*, 16.

[14] Frankfurt, *Reasons of Love*, 26.

[15] Harry Frankfurt, *Taking Ourselves Seriously and Getting It Right* (Palo Alto, California: Stanford University Press, 2006), 229.

[16] Eileen Guo, "An AI Chatbot Told a User How to Kill Himself—But the Company Doesn't Want to 'Censor' It," *MIT Technology Review*, February 6, 2025, https://www.technologyreview.com/2025/02/06/1111077/nomi-ai-chatbot-told-user-to-kill-himself/.

[17] Kate Payne, "An AI Chatbot Pushed a Teen to Kill Himself, a Lawsuit against Its Creator Alleges," *AP News*, October 25, 2025, https://apnews.com/article/chatbot-ai-lawsuit-suicide-teen-artificial-intelligence-9d48adc572100822fdbc3c90d1456bd0.

[18] Frankfurt, *On Bullshit*, 43.