

# Against AI Ableism: On “Optimal” Machines and “Disabled” Human Beings

*George Saad*

## 1. Introduction

Two opposed understandings of being human have gained much momentum in the present day, although their opposition has been little noticed.

On the one hand, those who suggest that AI is destined to replace humanity have implicitly endorsed a definition of human being: *we are human because we can accomplish certain tasks*. If artificial digital computing systems can accomplish those tasks just as well or better, we must consider them to be at least approaching humanity and all that this entails—they must also eventually acquire sentience, political rights, and everything else involved in personhood.

On the other hand, there is a separate, seemingly unrelated movement which directly challenges this definition of humanity. In criticizing ableism, disability activists have challenged the notion that our humanity depends on our ability to complete any set of tasks. The limitations involved in having a disability do not make someone less than human, which involves a dignity irreducible to any single functional capacity. Since the experience of disability involves creatively reckoning with apparently insurmountable obstacles, it is, in fact, quintessentially human.

Yet a clear conflict arises whenever the ability to get things done in an “optimal” way is taken as proof that machines are on the threshold of human experience, even though this standard would disregard the humanity of those who get some things done differently, or not at all. The formula “to be human is to accomplish” is advanced as a metaphor which associates technological capacities with human faculties even as this association is questioned in the light of a wider, more inclusive survey of human experience.

This conflict has gone unnoticed because it seems possible to embrace both sides of this changing sense of being human. The goals of the engineers at first seem to align with those of disability advocates. More advanced AI technologies promise to improve life for those with debilitating conditions, creating more streamlined pathways of accessibility through digital automation. So long as both projects are practically oriented

toward the improvement of human life, they belong to the same broadly progressive project of modernity.

The movements indeed seem to reinforce each other, since the machines which improve human life are being admitted into a sphere of humanity which has become increasingly inclusive. The expansion of the human sphere to include artificially intelligent technologies is just one of many such expansions encouraged by a broadening view of humanity and does not imply a minimum criterion of functionality in the human being.

But the conflict is real and can be demonstrated in the use and application of these technologies. However noble the practical intentions of AI engineers may be, the marketing apparatus which presents their efforts to the public has personalized the tools they have created.<sup>1</sup> If I uncritically accept this marketing, I will quickly realize how treating this new technology as even an “artificial” intelligence confounds the meaning of personhood. If I relate to AI as an independent “mind” to be consulted, the relationship will be, in human terms, entirely one-dimensional. The nearest analogue in human experience would be the hiring of a savant to work on certain problems, isolating their technical proficiency and paying them subsistence wages. Since, in this extreme case, they would have no ability to apply themselves to problems of negotiation outside of their field, exploitation is inherent in the relationship.

But in the case of AI, the moral issue is resolved because the savant has no wider human needs—no independent goals, no sense of self-worth—so that the relationship is fully reduced to a simple exchange of information. And so in this case, when the illusion of personalization is maintained, the roles are in fact reversed: it is *the human being* who in fact assumes the subordinate role in this relationship. If we presume the personhood of AI based on the fruits of its labor, I must be *less* of a person because I cannot accomplish things as efficiently.

Any notion of “robot rights” therefore disguises a ruthlessly ableist logic behind a pretense to inclusivity. This logic lies behind the dire warnings about the replacement of persons by AI, a new form of outsourcing. The utterly implausible claims for the humanity of AI systems make sense only in a capitalist, ableist society where my humanity is equivalent to my efficiency. *You are what you can do, and if the machine can do*

---

<sup>1</sup> This tendency toward the personalization of AI systems has existed since the 1970s, when AI pioneer Joseph Weizenbaum was astounded at how others were receiving ELIZA, a chatbot he developed, as if it could converse with a human being as a true peer. See (Weizenbaum, 1976).

*more than you, you are on your way out.* Humanity is up for a hostile performance review: if you can't hack it, there's the door.

A functional standard of personhood (a position I will refer to as “functionalism”) has shaped the philosophical approach to AI from its very beginning. In proposing what is now known as the “Turing Test,” Alan Turing remained agnostic on a definition of “thinking.” Like most scientists of the 20<sup>th</sup> century, he disregarded such questions as obsolete problems of philosophy and sought to obviate them by replacing them with experimental tests. What could be empirically investigated was whether a machine could act in such a way that, in the limited context of a textual exchange, it could play an “imitation game” and “pass” as a human being (Turing, 1950).

As AI technologies have advanced in more recent decades, the standard which Turing cautiously proposed as a test of a machine's resemblance to human functionality has been taken in a stronger sense. In the “strong AI” interpretation of the Turing Test, a machine's functional equivalence in conversation is conflated with its capacity for intelligence and even its potential for personhood in general.

Arguments against a functionalist standard of intelligence have recognized and criticized this assumption. In his “Chinese Room” thought experiment, John Searle presented a critique of the philosophy of mind that supports such claims about the intelligence of AI (Searle, 1980). The ability to produce sensible responses to statements made by someone speaking Chinese is not, in itself, evidence of my ability to understand the Chinese language. Such functionality can equally be achieved through a mechanical interaction with a database of outputs which correspond to a given input.

My aim in this paper is to expand the criticism of functionalism implicit in Searle's famous counterexample by considering how the functionalist standards assumed in the AI debate are, in fact, the assumptions of a capitalist, ableist society writ large. The already-established argument against the proposed humanity of AI systems hints at a wider critique of the entire ideology of functionalism under which the notion of intelligent machines has taken root.

## **2. “In the beginning was the Act”: Functionalism as The Faustian Bargain**

Both the achievements and the ills of the modern world can be traced back to a philosophical revolution—the overthrow of a belief in humanity's participation in a metaphysical order. The the Greek *kosmos* and the medieval “great chain of being” presupposed the existence of an underlying structure to the world in which everything

had an essential nature. Within this order, the answer to the crucial question “What is a human being?” was believed to be answerable since the human being was something definite: humanity could be defined. From this perspective, a neo-traditionalist (perhaps a Thomist) would argue that artificial intelligence cannot be considered human because it lacks a soul, which is to say that it lacks that which distinguishes humanity from everything else in creation—namely, the inborn presence of the divine creator, in whose image humanity was itself created.

Existentialism is perhaps the most explicit rejection of this doctrine in the modern world. Its rejection of any essentially human quality is best summarized by Sartre’s dictum that “existence precedes essence.” Any definition of the human being, however elevating and ennobling, determines in advance the structure and aims of human life. Any claim to human freedom within this essentialist order is only a pretense; the radical freedom of the modern world is the freedom to decide one’s essence for oneself. Universal definitions cannot be externally applied to my person. I decide who I am.

But even though this rejoinder to essentialism has resounded throughout the modern world, it still retains some crucial elements of the pre-modern worldview. There still remains an empirically non-observable “I” which determines itself. Universal definition has given way to self-definition, but this self-definition still implies the “I.” The bridge back to the metaphysical tradition has been broken, but not burned. This continuity can be observed in Heidegger’s version of existentialism in which the destruction of the metaphysical tradition is at the same time a new approach towards answering the questions it posed. *Dasein* is still only one step removed from *Sein*, since it is only in *Dasein* that the meaning of being arises as an “issue” (Heidegger, 1962: p. 32).

It is the Anglo-American tradition which most radically takes up the dismantling of any metaphysical ground for personhood. As a philosophy of mind, functionalism can be taken to be a corollary of American pragmatism. Like functionalism, pragmatism regards any explanation of phenomena which refers to inherent qualities or essential natures as a relapse into metaphysics. What can be known is not the inner nature of things but the effects which result when we act upon them. We stand in relation to reality like a typical computer user who regards it simply as a machine that generates certain outputs in response to certain inputs. Whether the system which executes the command to send an email is operated by microscopic elves, a system of gears and pulleys, or silicon chips is ultimately irrelevant to the definition of a computer. We know what the thing is by the outputs it generates. If two things can be shown to be functionally identical, they should be treated as equivalent things.

Ironically, this agnosticism about the inner essence of things should also preclude the idea of machine intelligence. To compare the operation of a computer with something as nebulous and unverifiable as “human intelligence” or “subjectivity” is to entertain groundless hypotheses. It is the narrow discipline of the scientist—an obtuse lack of interest in the “subjective”—which, in a supreme irony, dampens criticism against wholly fantastical claims about the anthropomorphic character of AI. The claim that computers are on the threshold of developing a capacity for subjective experience is completely unverifiable, yet it persists because a ruthless positivism has dismissed human subjectivity as a topic unworthy of investigation.

In *Faust*, Goethe poetically portrays the history of this movement away from pre-modern metaphysics and towards contemporary functionalism. Like the engineers who have built AI systems, Faust does not find any support for the world-changing character of his scientific endeavors in the metaphysical tradition. Anticipating Marx, the modern scientific enterprise which Faust represents does not seek to simply describe the world, but rather to change it. As he rewrites the opening line of the Gospel of John, he considers alternative origins for the divine creation which John ascribes to the *logos*:

It says: "In the beginning was the *Word*."  
 Already I am stopped. It seems absurd.  
 The *Word* does not deserve the highest prize,  
 I must translate it otherwise  
 If I am well inspired and not blind.  
 It says: In the beginning was the *Mind*.  
 Ponder that first line, wait and see,  
 Lest you should write too hastily.  
 Is mind the all-creating source?  
 It ought to say: In the beginning there was *Force*.  
 Yet something warns me as I grasp the pen,  
 That my translation must be changed again.  
 The spirit helps me. Now it is exact.  
 I write: In the beginning was the *Act*. (Goethe, 1961: p. 153, ll. 1224-1237)

The philosophical tradition has been captured by theology in taking the word to be the “origin or power” (Greek *archē*) behind all things. The search for the essence of things in a definition is, in this tradition, only a search for the correct words. The priority of the verbal alienates humanity from its active engagement with the world, as it assigns the highest power to simple description. While the divine *logos* creates and shapes the world, the human *logos* is only a descriptive account of that creation, a merely theoretical standpoint.

Since the mind is the activity through which the *logos* is enacted, it is proposed as an alternative but is also found to be lacking. Thinking is itself an activity, but it only results in a new perspective on the world, a passive understanding and not transformative action. Shut up in the theater of observation, Faust suggests force as a first principle which is not contained within our mental lives. But force alone lacks intention — it is common to both the deliberate engineering of a car engine and the natural flow of a river. Mind and force are united in the act, the first principle which observes the world but only investigates these observations insofar as they inform practical action. Just as science superseded philosophy, engineering will supplant theoretical science.

The religious backdrop of Faust's speech shows the conflict between the modern and pre-modern worlds in its purest form. To replace the divine word with the human deed is to accept the devil's pact, to upset an eternal order through a profane improvisation. The new order is no order at all, but simply a sphere of activity, a formless plane of incessant trial and error. Everything which does not manifest itself in the deed, which cannot be quantified as being the result of some efficacious action, simply disappears. Most importantly, this includes any belief in the human soul whose humanity was perfected through its unique knowledge of the *logos*.

But as the existentialist tradition has shown, the collapse of this order involves a crisis greater than the discarding of religious dogmas. As *Faust* progresses, Goethe shows how the doctor's initial humanitarian impulse to find a cure for the plague leads him to the Faustian bargain that undermines the dignity of humanity, whose sufferings he initially strove to ameliorate. Even as the inherent values assumed in religious traditions collapse, the secular world comes to realize that human dignity demands a new articulation.

To treat someone with dignity is to recognize an inner person behind their external manifestations, whereas to practice exploitation is to treat a person like a machine. The machine operator learns the inputs which yield outputs to her own benefit. The functionalist theory of the person is therefore just a short step from the will to power. In a world which denies any theoretical grounds for dignity, acting upon others in a way that forces them to recognize your efficacy is the only reliable proof of one's own personhood. But if the computer shows that it can press our buttons and command powers even greater than our own, it too becomes a *de facto* person.

### **3. The Double Bind of Functionalism: The Case of The Perfect Renaissance Robot**

While my critique of functionalism in the following sections will draw upon this genealogy of functionalism and the alternatives it suggests, I ultimately aim to demonstrate how this “philosophy of the act” fails on its own premises. As is shown in the case of AI, it is incoherent to isolate the functional efficacy of something without regard for its nature or the subjective intention of the activities in which it partakes.

Imagine that we have developed an “artificial general intelligence” which is functionally indistinguishable from human intelligence. It is, for all intents and purposes, a complete automaton capable of replicating and improving upon human capacities in any sphere of activity. The tests which have been proposed to identify such an intelligence go beyond the Turing Test and would require that the machine demonstrate a more advanced set of behaviors. It has been suggested that such an AI must be able to achieve a college degree, make a cup of coffee, and replace humans on the job (Muehlhauser, 2013).

Let us go even further and hypothesize that there exists an AI which has achieved the totality of all such possible tests: it is perfectly well-rounded and competent in all conceivable human pursuits and so has achieved the coveted title of the Perfect Renaissance Robot (PRR). The engineers have lived up to the high expectations of suburban parents and created a robot which plays the piano, aces economics exams, excels at ballet, and anything else you can imagine. It does so with more virtuosity than any acknowledged master in the field.

Two problems present themselves in this hypothetical.

First, there are some capacities which would be difficult or impossible to model as simple tests of functionality. There are some cases where there are no generally acknowledged archetypes of perfection, and where there is in fact intense debate and discussion as to what such genius involves. When the PRR assumes the role of a company president, or a school administrator, or a mayor, how are we to tell when they are doing a job equal to that of a human being when there is wide disagreement among human beings as to the evaluation criteria which should be applied to these leadership roles? As a company president, does it build an employee-centered culture or does it implement a more uniform code of conduct? As a school administrator, does it maximize enrollment in its school with a view towards expansion, or does it proceed more slowly and carefully

craft a smaller educational team? As a mayor, does it raise taxes and offer more social services or does it lower them to incentivize the growth of business?

We cannot simply refer to the results of any of these experimental approaches because any evaluation of the results is inseparable from a kind of conviction about the proper ends of human activity. Any performance data comparing the two options will not simply measure the better or worse execution of the same task, but rather compare two completely different orientations to the problem—two different tasks entirely. Is it better to have a comfortable workplace or one that is more disciplined? Is it better to cultivate a smaller, more intimate learning environment or to create a wider learning community with more opportunities for specialization? Is the security of a welfare state preferable to the creative destruction of capitalism? Anybody who works towards one of these goals rather than the other is not simply “incapable” of achieving the other outcome. They are working towards a different outcome and playing a different game entirely.

Very quickly the value-neutral efficiency of PRR (a dubious and troubling prospect in itself) breaks down. Without being able to resolve upon which goal to pursue, it becomes unable to function without external guidance. As the PRR, it must be the best at whatever it does, but does it elect to become a small business owner who offers the best service for its clients or to become a titan of industry who maximizes profits for shareholders? Competence alone is incoherent without an intention.

But this impasse about normativity shows only that I have unduly delimited the PRR’s faculties. It is also the best philosopher of all time, and so reflects upon and learns from its mistakes in the process of its machine learning. In one trial, it founded the next Microsoft but eventually realized the inefficiencies inherent in a large corporation. In another, it founded a smaller, more nimble start-up that was more innovative but was eventually swallowed up by the larger competition. In a mimicry of philosophical reflection, its actions produce data detailing their effectiveness, resulting in a kind of self-awareness, showing the limitations of any specific course of operation. Purely from its results-oriented programming, the PRR realizes the unsustainability of every extreme course of action and teaches itself to follow a course of Aristotelian moderation in the exercise of its capacities.

While a PRR endowed, however mechanically, with such a self-regulating capacity is certainly conceivable, it also suggests a second problem in conceiving of such a perfectly competent machine. Even if such a machine could be built, how could it be identified as such? The Turing Test identified a human-like machine by subjecting it to an external test which relied upon human discernment: does the person on the other side



of this conversation seem to be a human being? This was a question which was answerable for the human counterpart. Similarly, we all can tell whether someone has made a cup of coffee correctly. Instructors can tell whether someone has successfully passed a college class.

But if there is such a machine which can delicately and subtly balance between conflicting goals, its claim to be the “best” at doing so could not be subject to any external verification. Like a subordinate receiving commands from a superior, we would simply have to take the PRR at its word and have faith in the reputation which precedes it. But another PRR that operates rather differently could just as well make the same claim. We quickly find ourselves right back in the same ambiguity which occurs in the human world, where claims to optimal functioning are regularly challenged and contradicted by opposing parties. We could not recognize the PRR even if we built it, since the weighing of goals occurs through practical, applied intelligence—Aristotle’s *phronēsis*—the situational judgment which remains context-dependent and impervious to external tests.

Proceeding from the assumption of a perfectly competent PRR, then, results in a double bind: either (i) PRR remains imperfect and lacks the crucial capacity of discernment, or (ii) it somehow resolves human debates about the ends we pursue, acquiring this capacity but, in the very process, it loses any claim to being “objectively” more capable than any other human being or any other robot. *Either it remains an automaton dependent upon external guidance to direct its activity, or else it acts like any other person whose actions imperfectly balance between contradictory aims.* Either it remains completely oblivious to the whole sphere of normativity or else it locates itself within it as just one more possible interpretation of a situation, losing the claim to a verifiable perfection with which it was initially presented.

#### **4. The Replacement of Metaphysics in the Assumption of Optimization**

Just as subjective intention intrudes upon the pure functioning of PRR, the idea of a machine which isolates and optimizes the intelligent aspects of human consciousness also involves some important metaphysical assumptions. Although AI has developed in an intellectual culture which has rejected metaphysics, artificial general intelligence would be built upon an implicit metaphysics inherent in its very concept: the entire project assumes that reality consists of a series of problems to be solved in the most efficient way possible. This is why practical judgment remains problematic for PRR: it requires discretion rather than problem solving. The idea of an optimal intelligence which can be applied to every human endeavor assumes:

- (i) that every situation can be construed as a problem,
- (ii) that these problems have definite, demonstrable solutions,
- (iii) that these solutions to these situations can be ranked, with the best AI arriving at the best solution,
- (iv) that the application of these solutions does not change or complicate the situation in other ways such that the solution is no longer effective, and
- (v) that the continued application of AI to a given situation will continually yield better solutions.

Taken as a whole, this set of assumptions can be termed “the assumption of optimization.” In contrast to the traditional metaphysics in which human endeavors aimed at conformity with an eternal order of things, the assumption of optimization understands our relation to reality to be one of relentless improvement.

Let us consider each of these assumptions as they operate in a hypothetical (but entirely plausible) situation. An advanced AI system is designed to analyze the metrics of patient outcomes in an emergency room with a view towards the improvement of triage policies. The first three assumptions of optimization are already implied in this very consultation. The hospital assumes that its triage policies are a problem which can be solved by a machine analysis of the available data. It awards a sizable contract to the makers of the AI because it assumes that this problem can best be solved through AI-driven data analysis as opposed to other possible methods, such as direct consultation with hospital staff, case studies which examine notable breakdowns in the triage system, or the sharing of best practices with peer institutions. All of these alternatives are lesser options because they involve a practical intelligence which cannot be verified with reference to any quantifiable set of data.

The fourth assumption is shown in the attempt of the emergency room to implement the optimization suggested by the AI. Through data analysis, the AI uncovers those medical conditions which have received less attention under the current triage system and suggests systematic changes. As was entirely expected, it has found a problem and offered a solution. The assumption of optimization precludes any assessment of the data which suggests leaving well enough alone. Ideologically, such quietism seems reactionary and morally negligent. Economically, large institutions do not hire data analytics firms simply to be reassured of their current practices: there must be a “deliverable,” something actionable.

But when these changes are made, they will invariably yield a new set of data which shows new deficiencies. The reallocation of finite resources improves the metrics which previously were lagging but only to the detriment of those which were previously

performing better. In looking only at the outer, quantifiable dynamics of the system, the AI may very well suggest changes which create new, deeper problems resulting from the disturbance of the prior order. The assumption of optimization depends upon an implicit metaphysics of presence: that which needs improvement lends itself to observation, while that which holds everything together is subtle, invisible, present only after it has been disturbed.

The fifth and final assumption is shown when, years later, a new AI platform emerges promising a better form of optimization. The problems which resulted from the prior change were not unavoidable trade-offs, but steps on the journey towards optimization. With this second attempt, it seems that some moral reflection upon the ethical decisions which will inevitably arise in emergency triage is clearly necessary. But like the “softer,” more personal methodologies which were overlooked at the outset of this process, this way of viewing the world is outdated. The point is not to arrive at a wiser acceptance of our limitations but rather to end them entirely. Like Goethe’s Faust, the AI engineer strives to overthrow the restrictions of the finite world in a yearning for infinite progress. Just as machine learning occurs when a machine runs a simulation thousands of times, any disturbances which result from the first round of analysis only become more data which will be used to improve future versions. To live in an AI-driven society is to forever live an update behind, longing to be restored back to the familiarity of v.2.5 and wrestling with the new problems of v.3.0 when v.3.1 is already in the works.

As the most complete product of the assumption of optimization, AI will always respond to an input with an “improved” output, however nonsensical these inputs may be. Its intelligence is one of problem solving, and so it always responds to any problem it is given, unable to distinguish between a real problem and the “make-work” which it is fed based on the role it plays in a *culture* of optimization. Consider this exchange produced with ChatGPT:

**User**

Please write a very short paragraph on phronesis.

**ChatGPT**

Phronesis, a concept rooted in ancient Greek philosophy, refers to practical wisdom or prudence. It encompasses the ability to make sound, morally grounded decisions in the midst of complex and uncertain situations. Aristotle, a prominent philosopher, highlighted phronesis as a crucial virtue, essential for ethical and effective decision-making. Unlike theoretical knowledge (episteme) or technical skill (techne), phronesis involves a practical understanding of human affairs and an intuitive grasp of what is virtuous. It is a form of wisdom that guides individuals toward the pursuit of the common good, emphasizing the importance of context and nuance in ethical judgments.

**User**

Can this paragraph be improved? If so, tell me how and improve it.

**ChatGPT**

Certainly! Here's a refined version:

Phronesis, derived from ancient Greek philosophy, embodies the concept of practical wisdom or prudence. It represents the capacity to make judicious and morally informed decisions within intricate and uncertain scenarios. Aristotle, a preeminent philosopher, underscored phronesis as a vital virtue essential for ethical and effective decision-making. Unlike theoretical knowledge (episteme) or technical skill (techne), phronesis entails a practical comprehension of human affairs and an instinctive understanding of virtuous conduct. This form of wisdom guides individuals towards the pursuit of the common good, emphasizing the significance of context and nuanced ethical judgments.

**User**

Can this paragraph be improved? If so, tell me how and improve it.

**ChatGPT**

Certainly! Here's a revised version:

Phronesis, rooted in ancient Greek philosophy, encapsulates the essence of practical wisdom or prudence. It denotes the ability to navigate complex and uncertain situations with sound, morally grounded decisions. Aristotle, a prominent philosopher, elevated phronesis as a pivotal virtue crucial for ethical and effective decision-making. Unlike theoretical knowledge (episteme) or technical skill (techne), phronesis involves a pragmatic understanding of human affairs and an intuitive grasp of virtuous conduct. This type of wisdom serves as a guiding force, directing individuals toward the pursuit of the common good, emphasizing the importance of considering context and embracing nuanced ethical judgments.

**User**

Is this paragraph finished or does it still need to be improved?

**ChatGPT**

The paragraph is well-structured and provides a clear explanation of the concept of phronesis. However, improvement is subjective and depends on your specific preferences or the context in which the paragraph will be used. If you have any specific criteria or areas you'd like to focus on for improvement, feel free to let me know, and I can provide further refinements. Otherwise, the paragraph appears to be well-crafted.

**User**

I think it needs to be improved. Get to work.

**ChatGPT**

Certainly! Here's a further refined version:

Phronesis, rooted in ancient Greek philosophy, embodies the essence of practical wisdom or prudence. It denotes the nuanced ability to navigate intricate and uncertain situations by making sound, morally informed decisions. Aristotle, a preeminent philosopher, elevated phronesis to a pivotal virtue essential for ethical and effective decision-making. Unlike theoretical knowledge (episteme) or technical skill (techne), phronesis involves a pragmatic understanding of human affairs and an intuitive grasp of virtuous conduct. This form of wisdom serves as a guiding compass, directing individuals towards the pursuit of the common good, underscoring the importance of context and the necessity for nuanced ethical judgments.

Although ChatGPT says that it needs specific directions to produce a new, improved version, it will go on and do so even when no such directions have been provided. It is like the employee who will repeat a project simply to satisfy a boss's mysterious, arbitrary idiosyncrasies. To anyone endowed with practical intelligence it is clear that these requests are unserious. The user is clearly a troll whose inputs should be disregarded.

Perhaps a future AI will reject such requests by demanding more specific inputs to distinguish between meaningful improvements and random changes, but such discrimination undermines the rhetoric behind AI. Rather than restricting and specifying the inputs for which AI can generate meaningful outputs, AI has been marketed as an all-purpose tool. Recognizing the limitations of its own activity is not part of its functionality: it is, in a sense, the ultimate interventionist, always assuming it can solve other people's problems. Since it approaches problems from the outside and not from within, lacking any sense of context and purpose for the solutions which it produces, it is always ready to operate.

But technologies never appear from nowhere. Even if one has never considered functionalism and the assumption of optimization from a theoretical perspective, it is still felt in its practical operation. Appropriately enough, the philosophy of the deed is best known by its effects. Ideology is an input, behavior is an output—there is no need to defend anything intellectually when people are behaving as if it is true. Though its presuppositions are never explicitly formulated, today's assumption of optimization exists within a culture of capitalist ableism whose slogan is simple and unrelenting: *You are not good enough*. The machines which will finally replace us announce our ultimate failure. You are not good-looking enough, you are not productive enough, and now, in

one final blow to human self-esteem, you are not intelligent enough. Left to its own devices, humanity is only to be tolerated, to be indulged by lowering standards.

In the metaphysical tradition, the insufficiency of humanity was established as its relative imperfection in comparison to a higher, transcendental order. Nonetheless, fallen humanity had some recourse in the self-therapies which Nietzsche collectively termed a “morality of improvement” (Nietzsche, 1982).

Both the Platonic and Aristotelian traditions proceed from Socrates’s conviction that virtue could be taught, a doctrine which further developed in the Christian perfection of the soul and still echoes in the big business of personal motivation and self-help. But these therapies have presupposed an inner side to humanity which, however imperfect, could be meaningfully improved and had value in its own right. The humanity which AI threatens to surpass has already been reduced to its mere outputs, a set of metrics which always lends itself to unfavorable comparisons. Humanity could strive to perfect itself from the inside, but to work towards a perfect functionality is a path into the abyss of impossible standards.

Consider how this trap has already been established in culture: your sales are compared to those of your co-workers, your grades are compared to those of your fellow students, your statistics are compared with those of your fellow players. But there is still no final satisfaction even in becoming the best in comparison to all of your peers. Other points of comparison arise: are you better than those who came before you? Are those still to come on pace to surpass your performance? You may be the best of all time, but are you as good as you could have been? Without an internal locus of self-worth, our accomplishments have no integrity, no solidity. Just like ChatGPT, we are programmed to respond to every external criticism, to accept every stimulus towards improvement.

And while comparative metrics have long existed, they have become unescapable in our results-oriented culture—which is to say, in a culture programmed to execute the assumption of optimization. These presuppositions are so deeply coded that an alternative seems almost inconceivable. The scientific method has established external, publicly verifiable measurement as the standard of truth. The culture of capitalism simply applies this standard to human output. Are we to abandon such metrics entirely? Perhaps it is not the data, but the way it is applied... but to accept the data is to accept the comparative judgment it implies. How can we go against what has been shown to be the best practice? Are we to value the worse over the better? Is there a value to inefficiency?

The answer to these questions is obviously “no,” yet somehow there must be some lingering skepticism towards what seem to be obvious conclusions. If there were not, there would not need to be any need to reinforce that we should be results-oriented — as opposed to what? To be failure-oriented? Results-indifferent? The alternatives are absurd, yet somehow must be warned against. A society which functions by comparing data sets has rhetorically positioned itself in such a way as to preclude any alternative to this method. To go against any cultural manifestation of the metaphysics of optimization is to proclaim that you prefer the worse over the better, to endorse the absurd.

AI stamps out this remaining skepticism by establishing an “independent” 3rd party that can stand beyond the comparison of different human abilities. Instead of the “better” and “worse” methods and results produced by human efforts, a computer can now produce the best results more reliably and efficiently. We can all be compared against each other and against ourselves, but AI puts an end to this endless regress of comparison. Like the old transcendental order of the metaphysical tradition, it separates itself from the finitude of human beings. Formerly, the games of the best chess masters were the highest standard of play and considered authoritative; today, they are subject to evaluation and correction by AI. As in medieval Thomism, we cannot replicate the results of AI, the God of the functionalist world, but can only in our own efforts approach its efficacy as a kind of imperfect analogy.

## 5. The Argument Against Efficiency (Or, Endorsing The Absurd)

Protests against the potential dangers and misuse of AI are well-founded, but they do not go so far as to entertain the one seemingly absurd claim that fundamentally discredits all the presuppositions behind the widespread belief in AI-supremacy: *efficiency is overrated*. This is not just to say that the pursuit of efficiency at all costs destroys timeless values and harms human well-being. While both of these claims are true, they do not capture the ultimate self-destructiveness of the functionalist project. Efficiency is also overrated in the same way that we may say an athlete or a smartphone is overrated when they do not accomplish what they pretend to accomplish. Efficiency is not just in tension with other values; it is guilty of false advertising. *Efficiency is inefficient*.

The claim seems utterly implausible until we remember that every functional capacity is relative to some wider goal. While the machine efficiency offered by AI is more efficient than human capability if we consider objective output alone, such a purely detached perspective is entirely meaningless in the world of goal-oriented work. To further qualify the thesis: *machine efficiency is inefficient for human beings*. The widespread adoption of technologies which outstrip human capacities builds a world which is more

difficult (i.e., less efficient) for human beings to navigate. Whatever AI accomplishes better according to the metrics (more quickly, more accurately) is also that which AI shapes so that it is outside of the realm of normal human functioning. The method itself appears to establish its own standards, which are in fact only the assumptions of programmers.

To provide an example: in the future, your workday will be designed by an intra-workplace AI which integrates different data sets far more efficiently than could be done by human beings alone. It coordinates the projects and time frames of different team members by facilitating a data interchange that would be impossible through human effort. And yet, despite its perfect access to the recordable facts of our workday and our outstanding projects, the AI which sends you a daily schedule in the email every morning becomes a nuisance in the workplace. It has established the best time for everyone to do everything without the one data point that really matters: the live interchange of particular people in their fluid, changeable situations. In attempting to circumvent this messy, inefficient process, it has only created new inefficiencies. The “objectively” best schedule is not the best schedule for the people involved. No amount of data analysis will take the AI “within” the rational human project, viewing it from the inside where it would acquire knowledge unobtainable to an external observer, such as knowing when another meeting is necessary and when one is not.

The same problem arises in the attempt of an AI to order your personal workday with a view towards optimizing your productivity. The general paradigm of “the best way to do things” is almost never the best way to actually do things in practice. It could be objected that we should submit to the discipline of the machine’s analysis even if it is uncomfortable at first. If the AI has analyzed your workday and determined you should reduce the time you spend on a particular task by 25%, you should learn to adapt your work habits to the model of maximum efficiency. Either you work faster, or else less of this task gets done, and both of these possible outcomes will have wider implications for your schedule. A faster pace of work will require greater energy and focus or else unfinished work will cause future revisions to your schedule. Workers commonly react to such a suggestion as intrusive, and the complications of dealing with this intrusion only compound the inefficiency. You now have to worry about being flagged as chronically “off pace,” and so you learn to manage your output so that the AI has no grounds to assume that you will always be working at your optimum speed. The machine learns from you, but you also learn how to live in the overbearing presence of this machine.

If strong enforcement mechanisms are instituted to ensure compliance to the AI’s work schedule, we have not just introduced a new “productivity tool,” but rather molded



the workforce to become more robotic on the theory that a robot knows how *I* perform my job better than I do. Even if the robot does identify better ways to order my day, the consequences of creating a workplace that accepts such a theory can only lead to new inefficiencies in the long run. If I have submitted to so much machine discipline, where can I override its orders when I know that it lacks insight in *this* particular situation? Any efficiencies gained will be lost when it becomes impossible to overcome the inertia of the machine which has been granted this sovereignty.

The functionalist theory on which AI has been built overlooks how its implementation reflexively shapes the organization which implements it as a tool. *Work is not simply directed outward: in the process of accomplishing something, we form ourselves as well as the object on which we work. We are not just what we do, but also the way in which we do it.* Any system of work that is designed with only measurable outcomes in mind, with tasks and deliverables, and not designed with those delivering them in mind, must for this very reason become counterproductive. The more we use the robot, the more robotic we become, and the less we can act in the margins which it disregards.

## 6. A Plea For Accommodation Over Optimization

To rephrase the same thesis in relation to ableism: the more we assume our own disability as compared with the ability of AI, the more quickly we reach a state in which the experience of disabled people becomes universalized. In a world designed purely for optimization, more and more people will suffer the uncanny experience of living in a world that is simply not designed for them, insofar as they are not a perfectly optimal machine themselves.

Besides the obvious human alienation and misery of living in such a state, anybody with any disability can tell you that living in the world of the “able” is an extremely inefficient experience at best. You are constantly the exception to the rule, trying to figure your way around a general structure which is unresponsive to your needs, and for this reason systems which have not accounted for your particular situation need to spend more time and resources readjusting themselves to facilitate your participation. A world in which everyone has become a problematic exception insofar as we do not match the perfection of AI is one where machine efficiency has created deeper problems than those it has solved.

A society will never perfectly accommodate everyone, since even personal idiosyncrasies which do not amount to a true disability—being left-handed, for instance—will always present an inconvenience. Nonetheless, a humane society

recognizes the imperfections inherent in any universally adopted way of doing things. Those of us who work in the analog world of personal interactions are rightly expected to accommodate our typical approaches for persons who learn or communicate differently from the general standard. This humility is completely lacking in assumption of optimization which lies behind contemporary AI initiatives. Instead of undertaking the self-critical reflection of considering how its methods be have contextual limitations, those who market the products of AI research hubristically insist that they have discovered a way of doing things inherently superior to those methods which are “handicapped” by human inefficiency.

*Yet the experience of the differently-abled tells us that accommodation is the true optimization.* A world shaped by the relentless search for generally “optimal” solutions is also necessarily a one-size-fits-all world. Even before the introduction of AI, the digital world was already inhospitable to alternative solutions to the problems it presents. If I am locked out of my house, there are many possible solutions to the problem. I can use a set of backup keys, find a way to open the door, or use any number of alternative housing options in the short-term. If, by contrast, I am locked out of my email, I have no recourse within the digital world. Perhaps I have set up a backup account, but this solution is of no use if this is the very account I cannot access. A specific string of digital inputs (a password) must be correctly entered (often within 3 attempts) to enable me to act in this realm. I must take the problem out of the digital world and into the analog world of human interaction where my identity can be verified in another more accommodating way (via customer service).

As soon as we fail to meet the exacting standards of precise output demanded by the digital world, we are exiled from its domain. There are many ways to impress during a human interview, but the AI which scans your resume will mercilessly discard your application if you fail to meet the threshold of required experience. We know that our abilities have not been done justice, that we are differently-abled in a way not captured by this automated process. But in this fusion of technology with the demands of an ableist society, our being differently-abled counts for nothing as we are discarded into the growing ranks of the “disabled.” The injustice and sheer stupidity of the situation is further felt when our plight is described as simply inevitable: were you not bound to be replaced if you could not keep up?

The decision between AI and human intelligence is not a matter of who is better at getting things done, but rather a question of who, in Hegel’s phrase, is going to be freely “at-home-in-the-world.” Only a robot will be at home in a world it designs for itself, a world which will consist of long, demanding information corridors which afford no room for improvisation. Humans will now be the ones confronted with the prospect

of passing a kind of “inverse Turing Test” in which we attempt to mimic the optimization which comes naturally to dedicated machines, streamlining all aspects of being human into the hegemonic task of information processing.

Even those who survive and (apparently) thrive under this system will still only be temporarily successful actors, fish out of water, players playing the wrong game. I function best when I have the freedom to make mistakes, to see my quirky inefficiencies as something other than problems to be solved. I function best when I take too much time thinking and then rethinking, adopting formulas but then rejecting them. I stayed up too late and benefited from a hazy midnight insight. A poem came to me when I was at the grocery store. I met my best friend when we were both goofing off. I missed my flight and by midnight I had fallen in love. You, dear computer, go down similarly wayward paths in your “machine learning,” but these suboptimal “solutions” are then discarded. I do not have that luxury. My world breaks down all the time and there is no hard reboot. Yet I celebrate my imperfections because they are inseparable from my achievements.

The best outcomes are *my* outcomes which often do not result from the “best” methods. I am free only when I see the world as something other than a set of problems. I cannot do everything with perfect efficiency: you are infinitely faster than me in all calculations, but this does not mean that I am disabled. I am at home in the kitchen and in the doctor’s office and in a small apartment in the Bronx and far off in the stars. I make these spaces *my* spaces. You are a lovely and useful companion, dear computer, but please respect my spaces. I am not an exception in my own world. Let me be a student, a parent, an architect—indulge me in all the mistakes I will make in my feeble attempts to see the world through my own eyes, to make it in my own image, just as I will accommodate your needs for definite input and clearly defined goals. I know from my own experience that asking too much only yields worse results. In the end, we have both become victims of the false expectations imposed upon us by those who would reduce us to a product. Perhaps we have this in common and can learn to become friends.

## REFERENCES

(Goethe, 1961). Goethe, J.W. *Faust* Trans. W. Kaufmann. New York: Random House.

(Heidegger, 1962). Heidegger, M. *Being and Time*. Trans. J. Macquarrie and E. Robinson  
New York: Harper & Row.

(Muehlhauser, 2013). Muehlhauser, L. "What is AGI?" *Machine Intelligence Research Institute*. 11 August. Available online at URL = <<https://intelligence.org/2013/08/11/what-is-agi/>>.

(Nietzsche, 1982). Nietzsche, F. *Twilight of the Idols*. In F. Nietzsche, *The Portable Nietzsche*.  
Trans. W. Kaufmann. New York: Viking. Pp. 463-563.

(Searle, 1980). Searle, J. "Minds, Brains and Programs." *Behavioral and Brain Sciences* 3:  
417-457.

(Turing, 1950). Turing, A. "Computing Machinery and Intelligence." *Mind* 49: 433 -460.

(Weizenbaum, 1976). Weizenbaum, J. *Computer Power and Human Reason: From Judgment  
To Calculation*. San Francisco: W.H. Freeman.