On the Value and Lack of Values of Artificial Intelligence[1]

Artificial Intelligence (AI) is being developed by leaps and bounds, and it is accomplishing tasks that were once thought to be uniquely the domain of human intelligence. Some are raising concerns about the implications of this technology not with respect to what it can accomplish but with respect to what it means for understanding humanity. No less a public figure than Henry A. Kissinger has sounded an alarm in *The Atlantic* with "How the Enlightenment Ends: Philosophically, intellectually—in every way—human society is unprepared for the rise of artificial intelligence.[2]" One can respond to this "crisis" by assuming that AI exhausts what constitutes "reason" in order to desperately seek a niche for humanity, or one can question whether what is taken to be "Enlightenment reason" has so truncated the meaning of reason that reason is left impotent in the face of the power of AI.

The irony here is that Kissinger's reflections demonstrate what happens when a reductionist, socially constructed understanding of reason is allowed to dominate the discussion of "intelligence." The result is that the meaning of reason is readily reduced to pragmatic, instrumental reason informed by "analytical" critique of merely empirical data to the exclusion of what makes humanity a "rational" being in the first place – because this reductionist version is the kind of reason that we have come to elevate above everything when it comes to serving our interests. In short, a misanthropic understanding of reason is mistakenly viewed as a threat to human reason.

What follows proposes that we question the assumptions beneath the prevalent, anti-Enlightenment rhetoric concerning rationality not to undermine the clearly productive benefits of AI but to retrieve a far broader and beneficial understanding of reason. This kind of "critique" (not merely analytical criticism) allows the identification of the limits to AI that, at least potentially, can illuminate a pathway through the thick and murky undergrowth of uncertainty and the fear that AI will one-day replace humanity's role in the hierarchy of being.

In other words, in face of the purported collapse of Enlightenment Reason and its impotency over against the developments of AI, one can either join the dirge celebrating the death of reason for its being an arrogant human triumphalism or one can question whether what today is taken to

---

[1] Special and deep thanks to James R. Cochrane for feedback on the initial draft of this piece.
[2] See https://www.theatlantic.com/magazine/archive/2018/06/henry-kissinger-ai-could-mean-the-end-of-human-history/559124/-
For a discussion of the short-sightedness of much of the discussion of "Enlightenment reason" today, see the following posts at https://criticalidealism.org under the category "On Enlightenment:" "Enlightenment: Reflections on Michel Foucault's "Was ist Aufklärung? ["What is Enlightenment"] 7 February 2016" and "What is Enlightenment? A Response to Balcomb's Call for th Retrieval of 'Participation'" originally published in the *Journal of Theology for Southern Africa*, No. 147 (November 2013): 51-73.

be Enlightenment Reason is an adequate grasp of the discussion of reason at the end of the 18th century. Succinctly, what "is" is not necessarily what "ought to be," and the very exercise of investigating why "what is" is not necessarily what "ought to be" demonstrates the power of language both to confound and confuse as well as to illuminate and inform – even empower humanity individually and corporately.

An Eight Point Thesis

"Artificial" intelligence is able to "look" (although it doesn't know that it is "looking") at accumulated data to establish "logical" patterns that it then applies to "new" data.

A first point: human intelligence is "grounded" not exclusively in material phenomena but far more in concealed possibilities that connect all phenomena by means of a common (imperceptible) horizon.   Furthermore, the concealed possibilities of any one object are not isolated from the whole but, rather, inseparable from a shared totality. This horizon of possibilities consists of what "might be," "could be," "is," and "ought" to be so that at least in part, the possibilities that are now "present" have always been there in the "past" and constitute the call of the "future" as they are experienced in the "present." In other words, this transcendental view of time allows one to view time not merely as a progressive sequence moving from the "past," through the "present," to the "future" in the sense of mere linearity. Rather, time consists more of a "future," "past," "present" structure of ambiguous possibilities calling out for actualization and clarity.[3] Human intelligence experiences the world, then, not as a mere collection of things to be calculated, predicted, manipulated, and controlled but as an incessant **projecting out of "the given" ever new possibilities for actualization**. This is a remarkable capacity that involves not merely engagement of the actual but an anticipatory grasp of the possible in self-awareness of one's own sense of possibilities – a capacity one can legitimately question for AI.

---

[3] One might be tempted to attribute this theme to Martin Heidegger's *Sein und Zeit* (Tübingen: Max Niemeyer Verlag, 1979), originally published in 1926. However, it was already developed by Paul Natorp in his *Philosophische Systematic* (Hamburg: Felix Meiner Verlag, 2000) that originated as lectures in 1922/1923 but not published until 1958 (see for example, 276 ff.). Natorp and Heidegger had long conversations on walks while Heidegger taught in Marburg from 1923-1928. Nonetheless, even Natorp is not the origin because the theme has already been engaged by Kant in *Metaphysik Mrongovius* (AA XXIX, 960 ff.), and a case can be at least proposed that this theme provides a coherent framework form what otherwise.is taken to be "wild ramblings" in Kant's *Opus Postmortem* (AA XVIII). Volker Gerhardt proposes that the *Opus Postmortem* confirms that Kant's project failed (*scheitert*) (342) because only in humanity did Kant (342) overcome the famous "gap" (*Kluft*) announced in the *Critique of Judgment* (AA V, 195) between reason and the perceptible, material world (288-289). According to Gerhardt, Kant believed that humanity alone constitutes the "middle concept" (*Mittelbegrif*) between God and the world (342-343) as the being capable of grasping the "living activity" (*lebendige Tätigkeit*) to establish the "unity" an otherwise fragmented totality (342). Gerhardt appropriately says that Kant never sought with the notion of "Being" (*Sein*) a connecting link between reason and the perceptible world (341-342). However, Gerhardt appears to overlook that the "gap" remains only so long as "Being" is taken to be same kind of substance that connects two ontologically distinct dimensions of experience,   However, once one makes the critical turn to conditions of possibility, as Kant appears to do in both the *Metaphysik Mrongovius* and *Opus Postmortem*, rather than objective, substance claims, the "living activity" (*lebendige Tätigkeit*) that constitutes the condition for any and all experience of totality is no longer just a *necessary assumption* (*Critique of Judgment* AA V, 183) of humanity's reflecting judgment but applicable to all phenomena, not just humanity, as the horizon of possibility that is the condition for all dynamism both perceptible and imperceptible.

A second point: although the computer's "intelligence" is not natural (the surprise of it all?!), the computer is also not (!!!) the origin of its "intelligence." It is able to identify logical patterns because a human being has constructed the hardware in ways that nature cannot accomplish on its own and because a programmer has provided it with a sequencing tool on the basis of a humanly constructed, computer language.

A third point: although artificial "intelligence" makes a "judgment" (surely, a metaphor), to make a "judgment" in the full sense of the term requires a consciousness that is different (to a degree that it is for all pragmatic purposes a difference in kind) from the phenomena about which it judges. This is not because we're talking about two kinds of substances (physical and mental) but because judgment requires the addition of things to the phenomena that can't be given by the phenomena themselves: for example, 1) awareness on the part of a "self" that perceives the phenomena; 2) a desire (not just fulfilling a logical command) to "make sense" of the phenomena; 3) which requires a conscious conviction (not a mere automatic recognition) that phenomena capable of being understood conform to a "law-like" order; 4) ability consciously to apply what must be believed to be a system of **universal concepts** (not just a single concept) to the **particular phenomena** given in perception according to two strategies: determining and reflecting judgment (see below for the distinction); 5) even to be able to make a **universal judgment** of beauty in nature **without a concept**;[4] 6) not to speak of the ability to judge the difference between "what is" and "what ought to be."

A fourth point: humanity is able to think up causal explanations for the phenomena that it experiences. Humanity is not limited to instinct, and each individual for her-/himself must generate and learn symbolic systems, which, again, are not given in the phenomena themselves, in order to explain events. What we perceive are the effects of causes. Causes are experienced only indirectly through perceived phenomena. This, in part, is what makes it possible for us to come up with multiple explanations for a set of phenomena. Given that the cause itself is not empirically perceptible (a consequence of the limits to human reason), there is no absolute proof or disproof of a causal explanation. However, what humanity has learned (it does not have it by instinct) is that we understand and explain phenomena best to the degree that we can identify a "law" (for example, a physical law, statistical significance, or algorithm) to the phenomena. This strategy is what allows us to distinguish between nocturnal dreams and the waking state. The former is not, the latter is governed by "laws." Nonetheless, causal explanations involve "good" and "bad" news: the good news is that there is no **absolute** proof that one's explanation is wrong (for example, one can always account for the dismissing judgment as having ignored secondary causes); the bad news is that there is no **absolute** proof that one is right. It helps to have a community that agrees with the one offering and accepting the causal explanation, but, even then, one can be led astray by one's pursuit of status and prestige within the group. In any event, those causal explanations that can be shown to conform to imperceptible "laws" (especially, when capable of mathematical formulation) are the most trustworthy causal explanations.

A fifth point: "thinking" is more than recognizing patterns but the conscious ability to make determining and reflecting judgments. Cryptically, "thinking" has to do with "seeing" things

---

[4] Any- and everyone who has not had the capacity beaten or trained out of her-/himself would find the scene of a sunrise over Three Finger Jack above Lower Burley Lake in Central Oregon beautiful although there is no concept "beauty" that unites the phenomena sunrise, mountain, lake, much less, water fall, orchid, eagle, coast line, etc.

that aren't there in the phenomena (directly or indirectly!). For example: no matter how many times one sees a table, the understanding that one is seeing/has seen a table requires more than the data. It requires grasping a concept (within a system of concepts) as well as applying that (system of) concept(s) to data in one of two ways: 1) when one already "knows" the appropriate concept for the data, one makes a **determining judgment** by applying the known concept to the data; or 2) when one does not already have an appropriate concept (e.g., "I don't have the faintest idea what the professor is saying"), one has the conscious option of choosing (or not choosing) to "go find the appropriate concept(s)" for the phenomena for oneself in order to understand by means of **reflecting judgment**. Neither form of judgment is the mere application of a pre-programed, symbol system (e.g., computer language that the computer itself hasn't generated), and no computer (as far as we will ever be able to determine) is able to take pleasure in the discovery of the "right" judgment. Furthermore, all determining judgments were once reflecting judgments – for a rational consciousness. Finally, a determining or reflecting judgment can be made even in ways not governed by the data (e.g., as in the case of figurative language). For example, Achilles is NOT a lion! The capacity for determining and reflecting judgment is the absolutely incredible surprising aspect of human intelligence.

A sixth point: "thinking," then, is far far more than merely applying concepts (not just past patterns) to data. This is where the notions of "apperception" (self-consciousness), beauty (forming a universal judgment WITHOUT A CONCEPT), and the mathematical and dynamical sublime are crucial. Neither the mathematical nor dynamical sublime refers to some external phenomena directly, but they are a form of aesthetic[5] judgment about consciousness. There are no beginnings nor ends to concepts in what is the mathematical sublime, and consciousness' ability to initiate sequences of events that nature cannot initiate exclusively on its own means, in principle, that consciousness can destroy all of nature in what is the dynamical sublime. All of these notions are saying something about consciousness itself, not the content and patterns that emerge as the content of consciousness.

A seventh point that draws on the other points, thus far: Critical Idealism acknowledges only two "domains" of experience in which there is clarity and distinctness of perception **AND** lawfulness because only these two are causal systems capable of "explaining" experience: nature and freedom. Nature is capable of being perceived with clarity and distinction either directly or indirectly, and it conforms to a causal system of "law-"fulness. However, unlike the "blind" lawfulness of nature, there is a second causal system that humanity (and any other "rational" being) is capable of experiencing – as well as, capable of ignoring. This causal system consists in what Critical Idealism calls **autonomous freedom**: the ability to initiate sequences of events that are never separate from but, also, never reducible to merely physical causality. In short, humanity experiences itself as capable of doing things that nature on its own cannot. Artificial Intelligence is an example, as are space probes and insights into microbiology and subatomic particles. It is precisely because of autonomous freedom that we are capable of

---

[5] Yes, the terms aesthetic comes from the Greek αἴσθησις, which means "perception by the senses." However, Critical Idealism's "critique" of sense perception involves the Copernican Turn to the "transcendental" (i.e., conscious) conditions that make sense perception possible, not just a "critical" analysis of sense data themselves. An aesthetic of the sublime, then, is ultimately concerned with the transcendental conditions that make a certain kind of sense perception (an unlimited universe and an overwhelming natural power capable of destroying the individual) in which the individual is reduced to "meaninglessness."

holding ourselves responsible for our understanding, decisions, and actions to the extent that they **ARE NOT** determined by physical causality. Given that autonomous freedom is a causal system, it is "law-"full. However, in this case the "law" is not the physical law, which would be blindly determining of human creativity, but the moral law. As with all causal explanations, we cannot prove that there are universal, moral laws applicable to all times and all places, but, as with our conviction with respect to privileging lawful explanations to natural phenomena rather than to fold our hands and say, "life is just a dream!," we are better off as individuals and a species when we do assume lawful responsibility for our understanding, decisions, and actions.

An eighth point: Critical Idealism suggests that it is important to distinguish between rules that "ought" to govern human behavior that are "particular" and rules that "ought" to govern human behavior that are "universal" rules. The former govern what is called "consequentialism" or Unitarian ethics, which is an approach to ethical norms to govern events **over which one has no control**. Even John Stuart Mills' appeal to past experience to discern ethical norms for the present fails to be adequate because it ignores the novelty in life circumstances. The latter is concerned with the "archaeology" of moral principles above and prior to all particular circumstances, which is an approach to moral laws **over which the individual has control**. I alone can give myself "permission" to do something. No one else can determine for me what moral law I should invoke.

No one can be forced to embrace a moral law because a moral law can only be self-legislated. The rules that govern technical skills and pragmatic behavior are "particular" rules in light of **merely personal self-interest**. Furthermore, the rules that govern the "successful negotiation" of a social world (e.g., a particular culture, civic laws, corporate ethical climate, etc.) are "particular" rules tied to a particular, social world. A good example that illustrates the danger of confusing particular, social rules for universal, moral laws are the two formulations of the so-called "Ten Commandments" in the First Testament (Exodus 20 and 34): the former is a set of social rules that were needed to successfully negotiate a "nomadic culture;" whereas the latter is a set of social rules that were needed to successfully negotiate a "sedentary culture" with private property and domestic animals. However, both systems of social rules require that their citizenry be "moral," that is, that they seek to live according to universal, moral laws, not slavish adherence to particular technical skills or social conventions. One can do everything properly according to a given set of technical standards or social rules and still be immoral. Every Mafia Clan and Drug Cartel insists that it is "moral" because of its adherence to the particular, corporate rules that it embraces for governing its behavior. However, universal, moral laws are broad, and they include, for example, that one should not lie, one should not break one's promises, one should not exploit the ignorance of others to one's personal advantage, one should develop one's talents, one should respond to the suffering of others, etc. Above all, the moral law requires the affirmation of the dignity of all "rational" beings because only rational beings possess the capacity of intentional, autonomous freedom and the self-legislating of moral laws to govern the exercising of that freedom. In short, then, dignity is not a particular, social rule, but a universal, human capacity: it is not a "human right" that is owed to someone; it is a "human capacity" that each and every individual always and already possesses so long as s/he is alive.

Conclusion

There is little in these eight points that applies to Artificial Intelligence so that one might legitimately propose that there is little "intelligence" to AI. Judgment involves more than the blind application of algorithms, and morality involves more than being satisfied with "what is." Only a rational being capable of exercising autonomous freedom is capable of even raising the question, much less of making an effort to answer, what "ought to be."

AI offers rational beings a powerful tool for the pursuit of technical ends, but it borders on misanthropy to claim that it is capable of replacing humanity in the hierarchy of beings. However, Critical Idealism reminds us that we misunderstand our position in the hierarchy of beings when we assume that our privileged position gives us absolute and uncontrolled sovereignty over nature (and others) because of our "superiority." Our superiority consists in a set of non-instinctual capacities that we can choose to exercise or ignore as well as a set of capacities that we can choose to exercise or ignore makes it possible for us to determine what "ought to be."

We live in a computer age that particularly valorizes youth. We sell our youth short and abnegate our responsibilities to the extent that we reduce "rationality" to instrumental reason that stops short of asking what "ought to be" – and leaves open to commercializing self-interest the decision to do or not to do something. We all, but particularly the young, need education. The heart of rationality is education because humanity's instincts are so poor. Education requires acknowledgement that we "don't know" and "can't know" simply by birth and given our rational limits. It also requires that we engage most rigorously those symbol systems that humanity has come to learn to be in-sightful. Among these symbol systems are the natural sciences with its commitment to lawful understanding of phenomena. However, it is time that we began to learn from Immanuel Kant's Critical Idealism that there is a second lawful system that makes it possible for a rational species to discern what "ought to be" -- rather than to read Kant superficially to identify those grounds for dismissing him.