Presented in Prof. Otfried Höffe's
Oberseminar
12 January 2019
Tübingen

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

Abstract

An examination of the claims made on behalf of "artificial intelligence" (AI) that it can and will replace human rationality. Whereas AI has the potential to be beneficial (as the case with any product of autonomous freedom), it also has the potential to be extremely destructive. However, the core thesis of the paper is not simply one of a critique of AI by practical reason but a rejection of the materialistic reductionism on the part of the blind defenders of AI. Already in the 18th Century, Critical Idealism pointed out that there is far more to reason than the "hypothetical" imperatives and "culture" of technical skills. We lose humanity when we overlook the a priori synthetic structures of theoretical and practical reason, the reflecting judgment of "aesthetics," and "pure" religion.

Introduction

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

[1] Special and deep thanks to James R. Cochrane, Emeritus Professor, the University of Cape Town for feedback on the initial draft of this piece. This paper benefited tremendously from the conversation it generated at Prof. Otfried Höffe's Oberseminar on 12 January 2019. The observations with respect to AI have been more carefully formulated, and, especially at the suggestion of Prof. Höffe, a 9th Thesis has been added on the social and historical nature of reason. It was a genuine privilege to have one's work engaged by such a talented group of students.

[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.

whether what is today erroneously 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 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.

Nine Theses

"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.

Thesis 1: Human intelligence experiences the world 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. In other words, 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]

Thesis 2: although the computer's "intelligence" is not natural, 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.

Thesis 3: although artificial "intelligence" makes a kind of "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;[4] 2) a desire (not just fulfilling a logical command) to "make sense" of the phenomena;[5] 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)[6] to the particular phenomena given in perception according to two strategies: determining and

---

[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 can be taken as the claim by Plato that the good is "above being" (ἐπέκεινα τῆς οὐσίας) as the culmination of his similes of the sun and line in the *Republic* (509b). The theme is also found in Aristotle. See Heinz Happ, *Hyle. Studien zum Aristotelischen Materie Begriff* (Berlin: Walter de Gruyter, 1971): 687: "Alle Einzelfälle von ἐνέργεια gründen im Actus purus, von δύναμις in der 'reinen Möglichkeit', die als 'Urgegensatz' Dynamis/Energeia einander gegenüberstehen." Important in the present context is that it appears to have already been engaged by Kant in *Metaphysik Mrongovius* (AA XXIX: 960 ff.). 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 Postumum* (AA XVIII). In *Immanuel Kant. Vernunft und Leben* (Stuttgart: Philipp Reclam, 2007), Volker Gerhardt proposes that the *Opus Postumum* 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" (*Mittelbegriff*) 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, with the notion of "Being" (Sein), Kant never sought 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 Postumum*, 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.

[4] See *Critique of Pure Reason* A 361 f./B 40 f.

[5] This desire consists of both attraction and repulsion (*Lust und Unlust*) to pursue the "lawful" orders of the physical and moral dimensions of experience. See Birgit Recki, *Ästhetik der Sitten. Die Affinität von ästhetischem Gefühl und praktischer Vernunft bei Kant* (Frankfurt a.M.: Vittorio Klostermann, 2001).

[6] See Kant's discussion of the "Table of Categories" in the *Critique of Pure Reason* B 105 f.

reflecting judgment (see below for the distinction); 5) even to be able to make a universal judgment of beauty in nature without a concept;[7] 6) not to speak of the ability to judge the difference between "what is" and "what ought to be."

Thesis 4: humanity is able to generate 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.[8]" 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.

---

[7] In contrast to what is "agreeable" and "good," which are concerned with personal interest, the beautiful "is the faculty for judging an object or a kind of representation through a satisfaction or dissatisfaction without any interest. The object of such a satisfaction is called beautiful." (*Critique of Judgment* AA V: 211) However, Kant adds that beauty is not a "property of the object." (*Ibid*., AA V: 211) Although it is a logical judgment, "[…] this universality cannot originate from concepts." (*Ibid*., AA V: 211; see as well AA V: 215-216)) 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.

[8] See Kant's discussion of dreams in *Metaphysik Mrongovius* AA XXIX: 884f, 927. Kant wrote: "The dream is another phenome-non of the imagination. It occurs entirely naturally. Because the imagination is constantly at work and in sleep the effects of understanding have ceased, only the imagination remains and is thereby given free rein. It gives us representations of things [in the dream] rather than understanding … [The] productive imagination is especially manifest in a dream. The dream is a sequence of fabrications that are involuntary. When awake, we are in a shared world; in the dream, though, we are in our own world. – The dream's creativity is similar to that of the waking world but with a difference: in the dream the productive imagination is involuntary, without order and intentionality. In the waking world, in contrast, I can link my fantasy in many ways in all kinds of directions according to an order, and I can always call myself back from my fantasy whenever I wish. In the waking world, fantasy is also involuntary but the creative idea is not so strong as in the dream because in the waking world sense impressions limit us whereas in the dream all of the senses are suspended and only the field [in contrast to territory, where order is possible, and domain where order is necessary] of the productive imagination is active. This is because the dream suspends entirely our consciousness of our circumstance. As a consequence, we have that peculiar experience that we can represent the past without knowledge that it is past. Here a subject of the reproductive imagination is opened up in which we swim in fancies without being conscious of our actual situation." (*Ibid*., 885) ([McGaughey's translation] See as well, Kant's *Prolegomena To Any Future Metaphysics* AA IV: 290-291; and *Kant's Critique of Pure Reason* B 520–21.

Thesis 5: even instrumental "thinking," as AI demonstrates, is more than recognizing patterns, but the human mind possesses the capacity to make determining and reflecting judgments.[9] "Thinking" has to do with "seeing" things that aren't directly and immediately in the phenomena. 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.[10] 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 although one can gain insight into who Achilles was and what he means by considering him to be a lion! The capacity for determining and reflecting judgment is the absolutely incredible, surprising aspect of human intelligence.

Thesis 6: "thinking," then, is far far more than merely applying concepts to data. This is where the notions of "apperception" (self-consciousness),[11] beauty (forming a universal judgment WITHOUT A CONCEPT),[12] and the mathematical and dynamical sublime[13] are crucial to understanding human reason. Neither the mathematical nor dynamical sublime refers to some external phenomena directly, but they are a form of aesthetic judgment[14] 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.[15] All of these notions are saying something about consciousness itself, not the content and patterns that emerge as the content of consciousness.

---

[9] On determining and reflecting judgment, see "Section IV" of the "Introduction" to the *Critique of Judgment* AA V: 179f.

[10] See the *Critique of Judgment* AA V: 187. All determining judgments were "surely" (gewiß) reflecting judgments, originally.

[11] See the *Critique of Pure Reason* B 131 f.

[12] See for example, §8 of the *Critique of Judgment* AA V: 214.

[13] See the *Critique of Judgment* AA V: 244 f.

[14] The term 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."

[15] Already in his so-called "pre-critical" phase, Kant pointed out in his *Vorlesung zur Moralphilosophie* (*Lecture on Moral Philosophy)* (Berlin: Walter de Gruyter, 2004) that humanity's autonomous freedom (in but not reducible to nature's "blind" causal system) in principle has the ability to destroy all of nature (177).

Thesis 7: Critical Idealism acknowledges only two "domains[16]" of experience in which there is clarity and distinctness of perception AND lawfulness because only these two are causal systems capable of "explaining" experience: these two domains are nature and (autonomous) freedom. Nature is capable of being perceived with clarity and distinction, and it conforms to a causal system of "law-"fulness.[17] However, unlike the "blind" lawfulness of nature, humanity (and any other "rational" being) exercises a second causal system – to be sure, not independent from but also not reducible to nature's causal system. This autonomous freedom is what makes it possible for us to assume responsibility for our decisions and agency although we can choose to ignore that responsibility whereas we cannot ignore nature's "law-"fulness. This causal system consists in what Critical Idealism calls autonomous freedom: the ability consciously and intentionally to initiate sequences of events that are never separate from but, also, not reducible to merely physical causality.[18] In short, humanity experiences itself as capable of doing things that nature on its own cannot. Artificial Intelligence itself is an example, as are space probes as well as insights into microbiology and subatomic particles.

It is precisely because of autonomous freedom that we are capable of 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 lawfull. 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.

Thesis 8: Critical Idealism suggests that it is important to distinguish between rules that "ought" to govern human behavior that are "particular" (hypothetical imperatives) and rules that "ought" to govern human behavior that are "universal" rules (categorical imperatives).[19] The former govern what is called "consequentialism" or Unitarian ethics, which is an approach to ethical norms as socially constructed either in terms of social laws/rules/conventions taken to be required to successfully negotiate a social world[20] or as the consequence of seeking a self-appropriate "mean" between excess and deficiency with respect to those things in life of which one can have "more and less" (e.g., food, material possessions, sex, etc.). Such particular rules

---

[16] See "On the Domain of Philosophy in General" in the "Introduction" of the *Critique of Judgment* AA V: 174 f.

[17] The rise of the natural sciences is described by Ernst Cassirer as a shift in epistemology from substances to functions. Lawfulness is a symbolic grasp of functional relationality among phenomena. On the "functional" nature of concepts, which are necessary to understand objects (not direct perception of substances), see Kant, *Critique of Pure Reason* B 91-116, especially, Kant's definition of function at B 93; see as well, *Metaphysik Mrongovius* XXIX: 889. For a succinct account of the shift in epistemology from substances to function, see Ernst Cassirer, *Die Begriffsform im mythischen Denken*. Leipzig/Berlin: B.G. Teubner, 1922. For a more detailed account, see *Substance and Function and Einstein's Theory of Relativity* (New York: Dover Publications, 1953) and *Das Erkenntnisproblem in der Philosophie und Wissenschaft der neueren Zeit*. 4 vols. Darmstadt: Wissenschaftliche Buchgesellschaft, 1994.

[18] See the *Critique of Pure Reason* B 473 f.

[19] See the distinction between hypothetical and categorical imperatives in the *Groundwork of the Metaphysics of Morals* IV: 414-422.

[20] See Patricia Churchland's definition of ethics in the "Introduction" to *Braintrust: What Neuroscience Tells Us about Morality*" (Princeton: Princeton University Press, 2011).

are evaluated either in terms of "outcomes assessment" with respect to the successful (or unsuccessful) negotiation of a social world or with respect to one's personal understanding of "excellence.[21]" However, even John Stuart Mills' Utilitarianism, which appeals to past experience to discern ethical norms for the present,[22] fails to be adequate not only because it ignores the novelty in life circumstances but also because it can lead to the violation of human dignity. However, the ability to act on the basis of "universal" principles is concerned with the "archaeology" of moral principles above and prior to all particular circumstances, which is an approach to moral laws that affirms the individual's personal control over her/his ethical life. 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 clever behavior[23] 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 would be able to insist that it is "moral" because of its adherence to the particular, corporate rules that it embraces for governing its behavior if morality was merely a social construct to enable the successful negotiation of a social world. However, universal, moral laws are broad,[24] 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. These "broad" moral principles do not constitute a denial of

---

[21] The ethical "doctrine of the mean" is what Aristotle calls "moral virtues" in contrast to "intellectual virtue" in the *Nicomachean Ethics*. See Books II and X.

[22] See John Stuart Mill, *Utilitarianism* (Indianapolis: Hackett Publishing Company, Inc., 2001): 23-24.

[23] See the *Critique of Judgment* AA V: 431-432.

[24] See "VII [Internal] Ethical Duties are of Wide Obligation, Whereas [External] Duties of Right are of Narrow Obligation" in *Metaphysics of Morals* AA VI, 390-391. Kant distinguished between "narrower" (unrelenting, *unnachlaßlichen*) and "wider" (meritorious, *verdienstlichen*) duty already in the *Groundwork of the Metaphysics of Morals* IV: 424.

sensuousness" or consist of a "moral tyranny[25]" of hypothetical imperatives[26] that are to be taken to govern the "correctness" or "incorrectness" every detail of one's life. 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.

Thesis 9: Human reason is by no means merely subjective and limited to self-selected goals and achievements but, rather, is unequivocally social and historical.[27] Given that reason is profoundly limited, Kant's philosophy of history is not driven by an absolute goal (e.g., Christian salvation or Hegel's meta-Idea of the One). Kant's philosophy of history is governed by a, again non-Hegelian, "cunning of reason" that he labels reason's "unsocial sociality.[28]" Given that history begins, for Kant, with the conscious emergence of humanity's transcendental capacities of autonomous freedom as the ground of its theoretical and practical reason, history is viewed here as an open-ended project in which humanity seeks to become human (i.e., to *properly* exercise both its theoretical and practical reason).

Individual and groups can and will act *contrary to their self-interests* in the name of higher moral principles. Yet, even when humanity acts exclusively on the basis of self-interest, its ability to do so always includes the capacity of practical reason to act morally. Hence, the unsocial sociality of humanity consists in the possibilities of humanity's very practical reason. This by no means constitutes an embracing of dystopia because humanity's hope is not dependent upon its achievements or failures but on its originating capacities that can never be eradicated as long as there is such a rational species.

---

[25] Kant rejects the charge that morality is a form of tyranny in the *Metaphysics of Morals* AA VI: 409. Kant also rejects the claim that morality involves suppression of the "flesh." The "flesh" (included in "animality") is affirmed as the most fundamental, material basis of any and all experience (see *Religion within the Boundaries of Mere Reason* AA VI: 26-27). The criterion for sexuality is found in the second form of the categorial imperatives that is anchored in the recognition of human dignity: "So act that you use humanity, in your own person as well as in the person of the other, always at the same time as an end, never merely as a means." (*Groundwork* AA IV: 429) Furthermore, our "animality" is necessarily presupposed for the two "higher" capacities (*Anlagen*) achievable by a rational being: "humanity" as status and prestige in the eyes of others and "personality" as respect for the moral law as sufficient incentive for governing one's moral responsibility (see the *Metaphysics of Morals* AA VI: 27-28). In short, Kant denies that morality requires the denial of "sensuousness." See *ibid.*, AA VI: 408 but also 384, 390, 394, and 405. Kant also rejects "ethical asceticism." See „Ethical Ascetics §53" of The Doctrine of Virtue in *ibid.* AA VI: 484-485: "[…] monkish ascetics, which from superstitious fear or hypocritical loathing of oneself goes to work with self-torture and mortification of the flesh, is not directed to virtue but rather to fantastically purging oneself of sin by imposing punishments on oneself […] [I]t cannot produce the cheerfulness that accompanies virtue, but rather brings with it secret hatred for virtue's command." (*Ibid.*, AA VI: 485)

[26] On the crucial different between hypothetical and categorical imperatives, see Kant's *Groundwork to the Metaphysics of Morals* AA IV: 413-415. The former are imperatives that govern particular "skills" (either technical or pragmatic) whereas the latter are moral imperatives that are universal.

[27] See „Fünfter Teil: Geschichte" in Otfried Höffe's *Kants Kritik der praktischen Vernunft. Eine Philosophie der Freiheit* (Munich: C.H. Beck, 2012): 273-337).

[28] See the „Vierter Satz („Fourth Thesis") of Kant's *Idea for a Universal History with a Cosmopolitan Aim* AA VIII: 20-22.

In other words, the very rational capacities that constitute our species marker lead us not merely to exercise our individual, creative capacity but also to democratic social orders grounded in representative government with a constitutionally guaranteed division of powers (legislative, administrative, and judicial) as well as to international cosmopolitanism and the negotiation of national interests under the auspices of a league of nations but not a world government.[29]

The very "nature" of humanity's "unnatural" capacities – all of which, of course, are not of our own or any other human being's creation – ground reason in a social order, as fragile and precarious as the human condition is.[30]

Conclusion

There is little in these nine theses 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 but make 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.

---

[29] For an account of Kant's influence on Woodrow Wilson's vision for the League of Nations, see Gerhard Beestermöller, "Die Umsetzung der Völkerbundsphilosophie in politische Wirklichkeit durch Woodrow Wilson." In *Die Völkerbundsidee. Leistungsfähigkeit und Grenzen der Kriegsächtung durch Staatensolidarität* (Stuttgart: Verlag W. Kohlhammer, 1995): 94-142.

[30] On humanity's precarious position, see *Groundwork of the Metaphysics of Morals* AA IV: 425-426.

WORKS CITED

Kant's works are cited according to the Akademie Ausgabe (AA) [Academy Edition] of the Königlich Preußischen Akademie der Wissenschaft [Royal Prussian Academy of Sciences]. The exception is the Critique of Pure Reason, which is cited according to the pagination of its two editions: the 1781 First Edition (A) and the 1787 Second Edition (B). An example: Kant, *The Critique of Judgment* AA V: 431-432 refers to the Academy Edition Volume V: page #s 431-432. Translations into English of Kant's works often have the AA pagination in the columns. Unless otherwise indicated, translations of Kant's works come from the Cambridge University Press (CUP) edition.

Aristotle. *The Nicomachean Ethics*. Translated by David Ross. Oxford: Oxford University Press, 2009.

Beestermöller, Gerhard, "Die Umsetzung der Völkerbundsphilosophie in politische Wirklichkeit durch Woodrow Wilson." In *Die Völkerbundsidee. Leistungsfähigkeit und Grenzen der Kriegsächtung durch Staatensolidarität*. Stuttgart: Verlag W. Kohlhammer, 1995.

Cassirer, Ernst. *Die Begriffsform im mythischen Denken*. Leipzig/Berlin: B.G. Teubner, 1922.

_____. *Das Erkenntnisproblem in der Philosophie und Wissenschaft der neueren Zeit*. 4 vols. Darmstadt: Wissenschaftliche Buchgesellschaft, 1994.

———. *Substance and Function and Einstein's Theory of Relativity*. New York: Dover Publications, 1953.

Churchland, Patricia. *Braintrust: What Neuroscience Tells Us about Morality*. Princeton: Princeton University Press, 2011.

Gerhardt, Volker. *Immanuel Kant. Vernunft und Leben* (Stuttgart: Philipp Reclam, 2007.

Happ, Heinz. *Hyle. Studien zum Aristotelischen Materie Begriff*. Berlin: Walter de Gruyter, 1971.

Heidegger, Martin. *Sein und Zeit* [*Being and Time*]. Tübingen: Max Niemeyer Verlag, 1979. 1926.

Höffe, Otfried. *Kants Kritik der praktischen Vernunft. Eine Philosophie der Freiheit*. Munich: C.H. Beck, 2012-

Kant, Immanuel. *Critique of Judgment*. AA V: 165-485. 1790.

———. *Critique of Practical Reason.* AA V: 1-164. 1788.

_____. *Critique of Pure Reason*. 1781/1787.

_____. *Groundwork of the Metaphysics of Morals*. AA IV: 387-463. 1785.

———. *Idea for a Universal History with a Cosmopolitan Aim*. AA VIII: 17-31. 1784.

———. *The Metaphysics of Morals in Two Parts: The Doctrine of Right and The Doctrine of Virtue*. AA VI: 204-355, 376-43. 1797.

———. *Metaphysik Mrongovius*. Akademie der Wissenschaften der DDR Band XXIX: 747-1047. 1783.

_____. *Opus Postumum* AA XVIII.

———. *Prolegomena to Any Future Metaphysics.* AA IV: 255-383. 1783.

———. *Religion within the Boundaries of mere Reason.* AA VI: 1-202. 1793.

———. *Vorlesung zur Moralphilosophie* . Edited by Werner Stark and Manfred Kühn. Berlin: Walter de Gruyter, 2004. 1774/1775.

McGaughey, Douglas. :" "Enlightenment: Reflections on Michel Foucault's "Was ist Aufklärung? ["What is Enlightenment"] at https://criticalidealism.org.

_____. "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.

Mill, John Stuart. *Utilitarianism*. Indianapolis: Hackett Publishing Company, Inc., 2001.

Natorp, Paul. *Philosophische Systematic*. Hamburg: Felix Meiner Verlag, 2000. 1922/1923, Pub. 1958.

Recki, Birgit. *Ästhetik der Sitten. Die Affinität von ästhetischem Gefühl und praktischer Vernunft bei Kant*. Frankfurt a.M.: Vittorio Klostermann, 2001.