

Transhumanities as the Pinnacle and a Bridge

Piotr (Peter) Boltuc ^{1,2} 

¹ Department of Philosophy, Department of Computer Science, University of Illinois Springfield, One University Plaza 3030, Springfield, IL 62703-5407, USA; pboltu@sgh.waw.pl

² Department of Management Theory, Warsaw School of Economics, 02-554 Warsaw, Poland

Abstract: Transhumanities are designed as a multidisciplinary approach that transcends the limitations not only of specific disciplines, but also of the human species; these are primarily humanities for advanced Artificial Intelligence (AI leading to AGI). The view that philosophy, ethics and related disciplines pertain to all rational beings, not solely to humans, is essential to the philosophy of Immanuel Kant. This approach turns out to be practical at the epoch of advanced AI. Many authors ponder how a kernel of ethical respect for human beings can be built into Artificial General Intelligence by the time it becomes a reality. I argue that the task requires, among other components, inculcating the core of the Humanities into advanced AI.

Keywords: transdisciplinary science; transhumanities; Church–Turing entities; transdisciplinary humanities; interdisciplinary humanities; disruptive change; humanities for AI; friendly AGI

1. Transhumanities, the Need and the Idea

The first author who focused on transdisciplinarity was Jean Piaget¹. He postulated the kind of approach that “would not only cover interactions or reciprocities between specialized research projects, but would place these relationships within a total system without any firm boundaries between disciplines.” (Piaget 1972)

Let me situate transhumanities in the framework of transdisciplinary studies. I present the standard difference between disciplinary, multidisciplinary and interdisciplinary studies. I also choose between the two models of transdisciplinarity:

- A. Where disciplines mesh to the point of total blending;
- B. Where there is a smoother transition between, or among, the disciplines, with the center dominated by the mesh but a very visible disciplinary core remaining at the edges (Figure 1).

Let us choose the latter model, since the point is not to question disciplinary attainments but to put into question their centrality. To follow Hegelian structures regarding this, the former disciplines appear as *thesis* and *antithesis*, with the transdisciplinary center turning into their *synthesis*. (As a synthesis, at the next level, it becomes a new thesis, always already prone to the ensuing transdisciplinary challenge.)

The above point has little to do with Artificial General Intelligence (AGI), mentioned already in the abstract, since primary definition of interdisciplinarity is extension- and structure-based, not based on the epistemic *audience*. Secondary definition, that pertains to transhumanities, is based on the two assumptions: First, that transdisciplinarity provides the best partition among the sciences, also apt for the AGIs. Second, that transdisciplinary humanities—with AIs/AGIs as their audience—deserve their own focus area, facilitated by giving it a special name, which is: transhumanities. *The name transdisciplinary humanities, remaining open for both, human and AI/AGI epistemic subjects.*

1.1. Transdisciplinarity and Its Alternatives

The point of transdisciplinarity is to move beyond inter-disciplinary framework. In the latter, authors collaborate ‘at the edges’; they obey the theoretical frameworks and



Citation: Boltuc, Piotr (Peter). 2022. Transhumanities as the Pinnacle and a Bridge. *Humanities* 11: 27. <https://doi.org/10.3390/h11010027>

Received: 22 December 2021

Accepted: 27 January 2022

Published: 14 February 2022

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

paradigms of each of their disciplines—tinkering with applications outside of the standard domains that retain their entrenched centrality.

In transdisciplinary research, collaboration takes place at the center—no *disciplinary taboos* are *sacrosanct dogmas* anymore. For instance, Einstein challenged absolute qualities of time and space (Newton and Kant); Heisenberg challenged the laws of the traditional logic and validated paraconsistent quantum logic (with complementarity), while Darwin challenged the *sacrosanct* separation of species. All these revolutionary attempts took place at the outskirts of the established science of the time; albeit they moved their newly created sub-disciplines (e.g., quantum physics or evolutionary biology) to the center of the group of sciences they revolutionized.

1.1.1. The Graph of Various Levels of Multi-Disciplinary Integration

Within transdisciplinary research, disciplinary experts use their competency, while relinquishing the taboos and local paradigmatic ramifications². The graph below visualizes the relationships between transdisciplinary, interdisciplinary, multidisciplinary and disciplinary research. Transdisciplinarity is represented by one oval shape; however, its coloring changes, representing various disciplinary competencies. The interdisciplinary model shows several disciplinary blubs, with peripheral areas shared by two (there could be more, although not very many) disciplines. Then, there is the multi-disciplinary model with separate disciplines, with bridges (meaning collaboration) between some of them.

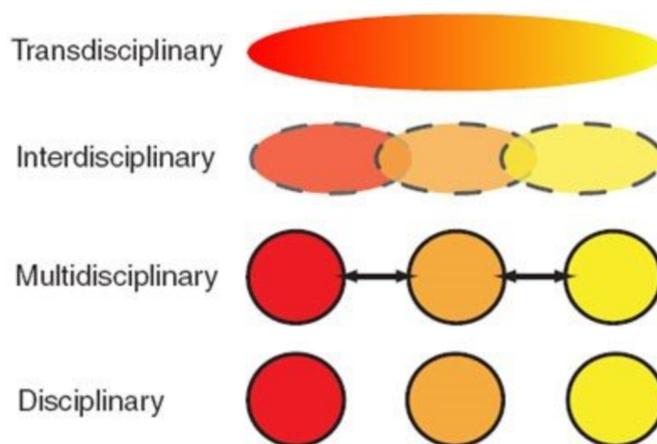


Figure 1. (Philip Mccarty 2016), creative commons https://www.researchgate.net/profile/Philip-Mccarty-2/publication/311486261/figure/fig1/AS:436722644328500@1481134200855/Transdisciplinarity_W640.jpg.

1.1.2. Non-Thinking Animals? An Example of Transdisciplinarity for Philosophy

Let us now present an example of transdisciplinary—and contrast it with an interdisciplinary approach.

Example 1. *My clearest experience with transdisciplinary discussion took place at the Minnesota Center in Philosophy of Science circa 1997, where Donald Davidson, a leading analytical philosopher of his generation, gave a talk on **thinking animals**. Davidson followed his published argument (Davidson 1985), which went like this: Since Davidson demonstrated in his earlier works (generally accepted in linguistics and philosophy of language at the time) that thinking requires relatively advanced language capabilities [premise 1] and since it was quite obvious, to him, that animals, even chimps, or dolphins lack such capabilities [the sort-of-empirical premise 2]; therefore, animals cannot think [conclusion]. Thus, what looks like animal thinking, is just instincts or other simple cognitive functions below thinking [consequence of the conclusion].*

At the meeting, Davidson's approach encountered quite strong critique, initiated by Wade Savage, the host in charge of the conference. The gist of the criticism was that

the argument does not fit with empirical research on animals. It also necessitates the conclusion that tribal humans, living in pre-language cultures or at the very early stages of language formation, would also not be thinking (by inflated Davidsonian standards). To this, Davidson tried to respond that even biology departments noticed weaknesses in those so-called ‘empirical arguments’ and have consequently stopped financing related research (such as Jane Goodall’s experiments with chimpanzees in the wild). To this, Wade Savage answered in a rather savage way (*the pun quite intended*), pointing out that those decisions had been made based on Davidson’s theory, the argumentation of which provides but circular support for his views. The point does show Davidson’s argument not only to be erroneous but also to beget dangerous practical consequences. This was followed by Dale Jamieson, then an endowed chair at St. Olaf, now a leading environmental philosopher at NYU, who tried to take a more conciliatory tone and to open some avenues for Davidson’s honorable retreat. Jamieson claimed that, when confronted with well-researched and intuitively plausible facts, a theory may either soften its core claims—in this instance, by lowering the language criteria for thinking, or limit the extension of its applications to refer only to the thinking of humans within civilizations that acquired advanced language. This would leave room for *thinking* of other kinds. Both ideas have been fastidiously rejected by Davidson, who maintained that his theory of thinking was well grounded in linguistics and needed no revisions.

The transdisciplinary aspect of this debate may now be shown in nice contrast with the inter-disciplinary one. Donald Davidson’s theory of meaning, and his resultant theory of thinking, is an interdisciplinary theory. Through its strong position in the previous century’s linguistics and its easy fit with the linguistic turn of the mid-20th-century analytical philosophy, Davidson’s theory became a vital component of those disciplines, well entrenched in both. Davidson’s approach was largely above and beyond critical scrutiny, especially from the natural sciences—shielded from animal observations and even quite common experiences with pets. Among *mainstream* philosophers and linguistics scholars at the time, criticisms such as those by Savage, Jamieson, myself and predominant majority of the participants of the MN seminar, would have been brushed off due to some form of reference to the authority. If this was not reference to the authority of Davidson, a highly respected Harvard professor, it would have been reference to the authority of disciplinary paradigms accepted in the analytical philosophy of language and linguistics of the time. However, the philosophers at the Minnesota seminar, showing themselves as quite transdisciplinary, had no special devotion to the dominant paradigm in a couple of disciplines (philosophy of language and linguistics) and were able to confront the position based on such a paradigm, with evidence and models developed in other disciplinary domains or by navigating among those.

Specific to the transdisciplinary approach is its ability to challenge disciplinary, and even multidisciplinary, consensus. The often context-dependent, inter-disciplinary perspectives, criteria, or trendy theories that dominate a given field for a decade or even for generations are the target of choice for transdisciplinary work, which is needed to allow progress and to disturb overly stable nodes of consensus.

1.1.3. An Example in Philosophy of AI

A similar situation can be seen in AI, where the so-called *good old-fashioned AI (GOFAI)* (Minsky 1985), still viewed as the gold standard, runs counter to the actual attainments in AI, such as the lifelong learning projects piloted by Siegelmann (Siegelmann 2019), *contra* to the longstanding view that neural networks cannot be productively retrained), or Thaler’s discovery engines (Thaler 2017). The latter operates close to the boundaries of AGI with an advanced cognitive engine DABUS, which is a radical challenge to the methods and limited goals of the good old-fashioned AI.

Thaler’s work, discussed in more detail later in this paper, provides strong empirical support to Alan Turing’s 1950 response to the old objection by Lady Lovelace, that comput-

ers cannot discover anything new, which still lingers among some computer science adepts (Turing 1950).

1.2. An Instance of Transdisciplinary Research in Socio-Economic Sciences

Let us now move to an exemplary approach to transdisciplinary research, by pioneers of transdisciplinarity within socio-economic sciences (Darian-Smith and McCarthy 2016). In their interesting article, Darian-Smith and McCarthy present both the theory and applications of a transdisciplinary approach to the field of Global Studies. They view transdisciplinary methodology as focused primarily on specific, contradictory, broad issues, often related to global socio-political concerns. The authors argue that the transdisciplinary approach is primarily based on the new questions, such as race and ethnicity; class and social inequality; gender and sexuality; postcolonialism, Orientalism, cultural imperialism; nationalism and the nation state; etc. These are legitimate and important topic areas—a very good fit with Global Studies. However, as is visible from the few examples in the previous sub-sections, social sciences are not the sole realm in need of transdisciplinarity; I return to this last issue in the next section.

In terms of appropriate methodologies for transdisciplinary studies, Darian-Smith and McCarthy argue that “[m]ultidisciplinary approaches use the perspectives of a number of different disciplines with no necessary overlap, and interdisciplinary approaches use the methods and theories of one discipline to inform others” [*op. cit.*]. In their search for an overarching method, the authors focus on post-structuralism and social constructivism, which looks like a step in the right direction.

They also make, or quote, helpful observations closely related to our broader topic. For instance, they argue: “Unfortunately, in the social sciences and humanities Piaget’s transdisciplinary approach has been overshadowed by the disciplinary/interdisciplinary debate and has not gained widespread recognition in the Euro-American academy. The disciplines are too often focused on defending their boundaries and resources to fully engage in cooperative efforts” (Darian-Smith and McCarthy 2016). Transdisciplinary humanities would also challenge disciplinary ramifications, following a similar approach to that developed by Darian-Smith and McCarthy.

Yet, now I would like us to consider an even more radical proposal.

2. Transhumanities

Now is the time to return to the issue of transdisciplinary humanities, which is my focal point in this paper. Soon thereafter, I move a step further, to what I call *transhumanities*.

2.1. Transdisciplinary Humanities—The Very Idea

Returning to the humanities, let us mention a quote, by Darian-Smith and McCarthy, from Rosemary Johnston (Rosemary 2008), who argues that transdisciplinarity “overtly seeks ways to open up thinking to ‘maps of unlimited possibilities’ . . . to create mindscapes that are unfettered by traditional patterns and procedures” (Rosemary 2008, pp. 229–30). A similar approach, focused on mind-maps in thinking, fits very well with Damasio’s neuroscientific theory (Damasio 2011). It also fits with Goertzel’s focus on patterns in his philosophically oriented book (Goertzel 2006), as well as with the *strong semantic computing* I argue for (Boltuc 2018b). The thinking maps and mindscapes, as well as the call to reach beyond the “traditional patterns and procedures”, as a way of moving away from disciplinary paradigms, very consistent with Johnston’s point, are the dominant methods in the current article.

I venture to say, largely based on the arguments from the previous section, that *transdisciplinarity came* to challenge disciplinary boundaries not merely in the extension of each science’s domain, but primarily its methodologies, so-called standards and other quite proudly defended *taboos*. The transdisciplinary approach asks questions that go beyond any previously existing disciplinary perspectives—even that of philosophy. The questions that may have been *opened up* by disciplinary progress, and then projected onto broader

theoretical and multi-disciplinary plains, within *transdisciplinarity*, emerge as integrative wisdom beyond the previous multi-disciplinary reach.

2.2. Preliminary Regulative Definition

Let me propose the following regulative definition of the transdisciplinary approach:

Transdisciplinarity uses analysis, synthesis, or other methods that go above and beyond disciplinary perspectives (as they were when we found or inherited them at the beginning of a given research project or theoretical work). The transdisciplinary approach is not constrained by methodologies, authorities, or most constraints of the disciplines at hand, though its practitioner needs to understand, and muster, the relevant disciplinary knowledge. It develops its own general methods and even paradigms above and beyond disciplinary limitations.

A transdisciplinary approach provides a way to go beyond petrified disciplinary or even interdisciplinary consensus, which does not fit with the current research on those or other sciences and fails under broader examination. It is possible, but not quite necessary, to cast the same thesis within the framework of Thomas Kuhn's theory of paradigm change (Kuhn 1969), where 'normal science' (the phase of stable growth) must be interrupted by the phases of scientific revolution that allow for the changes in the main paradigm—just as relativity theories, and then quantum physics, have changed the tenets of Newtonian physics.

Within a transdisciplinary framework, revolutionary changes in paradigms may not be required for rapid progress; they may occur much more easily than in Kuhn's framework, although their consequences may eventually lead to the synthesis of the new and old paradigms. This is due to the issues related to the sociology of science—as some authors claim—the change in dominant scientific theories results from the old authorities' disappearance due to retirement or death. Only then do people from younger generations take their posts, bringing their long-held views out into the limelight, to gain contemporary confirmation³ (Chawla 2019). In this sociological dimension of the problem, which is not the most inspiring one, although it needs to be taken into account, a transdisciplinary approach allows the change to be based on different nodes of influence that become relevant to the issues at hand. This process does not necessarily require uprooting disciplinary domains, some of which may hold value (especially for well-crafted applications of a given approach), whereas the new schools find broader intellectual and financial resources. Thus, contra Kuhn, the new subject may not require a revolution, but merely an extension from disciplinary, or interdisciplinary, to a still broader—transdisciplinary—reference frame (Classen 2020).

2.3. Beyond Transdisciplinarity—Humanities for AI

Transdisciplinary humanities are not the same as transhumanities. The latter can be defined as humanities that transcend their previous limitations, due to radical changes in their external or internal conditions. The changes which humanities face today primarily result from technological progress in electronics, what Luciano Floridi calls The Fourth Revolution (Floridi 2014). It is now clear that the informational revolution is nothing like the other stages of the industrial revolution. In the last 80 years (and, effectively, the last 25 years or so) it started to change our life, thinking, work, transportation, the military and other domains to a very radical extent. Looking back, one may see nothing radical in those changes, in the long, historical context, but adopting a historical perspective is unjustified because the informational revolution has not yet plateaued. It is not complete; it is not even slowing down; it remains in its early stage.

Our goal is not merely to present an account of current or upcoming successes of AI and other information sciences. There is much published work on this, though nobody seems to be able to gauge the scope of those changes in their entirety. One thing is quite certain; we are already cooperating with advanced informational entities (computers carrying sophisticated neural networks). This kind of functional symbiosis with AI and

other informatics, requires, or *gets to acquire*, a level of education and *acculturation* for these artificial thinking entities to an extent we can barely imagine.

2.4. Transhumanities—The Space for Connection among the Inteligencies of Various Kinds

Thus, **what are transhumanities?** These are the humanities presented in the universal context. Kant thought of rationality as going beyond human limitations (Kant 1781), as well as Plato, Spinoza and other major philosophers. Plato's philosophy was based on ideas, viewed as ideal forms that go beyond the scope of our thought or culture (Plato 1968). Spinoza tried to present ethics explained as geometrical proofs; he also referred to the infinite kinds of experiences available to infinite beings (which he viewed as angels), contrasted with merely two human forms of perception: internal, by first-person thought, and external, by empirical experiments (Spinoza 1677).

Kant's metaphysics of morals (Kant 1797), as well as his *metaphysics proper* (ontology), goes beyond humanism, resulting in the humanities (philosophy and ethics in particular) being addressed to any truly rational being. Well, cognitive engines such as AI seem quite rational. Thus, Kant's philosophy, and that of other major philosophers, would be addressed to them. *Let us explore this idea next.*

3. Transhumanities and the Church–Turing Entities

We now get into today's philosophy of AI. This is to see how it relates to the transhumanities.

3.1. Church–Turing Lovers and Beyond

In 2011, I introduced the notion of *Church–Turing Lovers* in a broadly philosophical paper, which I developed and modernized in 2017 (Boltuc 2011, 2017). The term comes from the Church–Turing thesis, which—in its so-called physical interpretation—may be construed as implying that every function that can clearly be described could be replicated by a possible machine (Deutsch 1985).

Church–Turing Lovers are **defined** as future artificial companions that can satisfy all reasonable expectations a person may have of one's human significant other, including caregiving, companionship, and intimate, sexual, and, *last but not least*, social needs. In the same paper, the author extended this class of robots to those meeting and exceeding human-level performance in other domains, including work and sports.

Artificial companions may have psychological and intellectual capacities approaching, and even exceeding, those of human persons. A good example is DABUS, which is the second cognitive engine developed by Stephen Thaler and his company *Imagination Engines* (Thaler 2019). DABUS is a lifelong learning AI-cognitive architectures, with many advanced functionalities. Recently, it caused a lot of publicity when it made a decisive contribution in making at least two engineering discoveries. Thaler decided that a human engineer, if he or she were given the same amount of information and guidance (*programming*), would have been viewed as the author of those discoveries and, therefore, be eligible to receive the patents. Consequently, he and his legal team applied for the two patents on behalf of DABUS (one of them being a set of fractal-shaped containers, which would have been counterintuitive to human designers but was optimal when compact-packaging certain commodities). While the courts in the Northern Hemisphere (EU, US, UK) dismissed the case, the courts in the Southern hemisphere (RSA and Australia) allowed the proceedings to continue, with a good chance of success.

Thus, the category of Church–Turing Lovers pertains to humanoid robots—going above and beyond mere house-cleaning robots; yet the category name focuses, a bit scandalously, on the controversial aspect of sexuality. *Alexa-style AI helpers* or unsophisticated robotic sex-toys are merely Church–Turing Artificial Companions, which may include non-humanoid machines or even AI programs (such as *Alexa* or *Cortana* when they are much improved). Thus, the category of sophisticated but not necessarily sexual robots is needed; namely that of Church–Turing Entities.

Church–Turing Entities are entities that satisfy two conditions:

1. They can perform relevant complex activities at a human level;
2. Their cognitive functions are advanced and autonomous enough to be viewed as a separate intelligence (*sapient*) and maybe even a *sentient* entity, to follow D. Kelley (Kelley 2020). Varela called the latter process *autopoiesis* (Maturana and Varela 1980) and the term is now used in AI theory in a similar context (Goertzel 2006).

3.2. Transhumanities and the Posthuman Condition

Having defined Church–Turing Entities, we need to move forward with the topic of transhumanities. What is the connection between the two?

To understand this problem, we think of *the posthuman condition*. This is viewed, by Braidotti, as the end of the Enlightenment (Braidotti 2018). However, what is, and what is not, the right thing to end? The human-centered existential condition is no longer tenable. As Floridi demonstrates (Floridi 2014), we have learned through the Copernican, Darwinian and Freudian revolutions that we are no longer justified in holding onto a self-image as living in the astronomical center of the universe, nor are we justified to view ourselves as detached and ontologically superior to the animals. Finally, we are no longer in a position to view our thoughts and intentions as transparent to ourselves—this (third) revolution is the most anti-Enlightenment of the three, since it undercuts belief in human intellect as crisp and apodictically normative, in the conceptual frame from (D’Alembert 1761) to at least (Kant 1781). However, the fourth revolution is still open for interpretation, to which this paper may make a slight contribution.

Braidotti views the posthuman condition as opposite to the Enlightenment; here, we disagree. The antithesis to the Enlightenment is Romanticism (Rousseau as counterpart to d’Alembert; Hegel to Kant). We view post-humanism as a humanism of sorts, a humanism that transcends beyond the limits of the intellect of natural human beings. Therefore, it fits not only with the Freudian Third Revolution, according to Floridi, but also with the consequences of Libet’s revolution as interpreted by Velmans.

To briefly explain the latter, Libet observed, within neuroscience, that we make decisions and start acting before we become conscious of the issue (Libet 2004). Many scientists and philosophers view this point as counterintuitive and contrary to human free will; however, Velmans presents a plausible explanation of how we act, in most circumstances, through thinking and decisions that do not come to the forefront of our awareness (Velmans 2009). The latter is more like the rear-view-mirror, documenting what has happened in memory, not quite making things happen (except for the halting instinct that brings the situation to our consciousness faster than we act and lets us stop action/s that are already in progress).

In this vein, posthumanism is the synthesis of the Enlightenment humanism and Romanticism’s emotion-driven way of being in the world. Braidotti’s aim is the posthuman condition, which she pictures as the end of capitalism, replaced by the abdication of humanity, sort of. In this, Braidotti is not quite Marxian enough, sliding into the backwaters of the utopian (or, in her instance, mildly anti-utopian) *idealistic* socialism of Proudhon.

Marx seems correct that no socio-economic system is going to prevail unless it proves most efficient in terms of technological and economic progress (Marx 1947). This was the basis of Marx’s violently decisive response to Proudhon’s utopian socialism, with the latter’s rejection of technological progress (and endorsement of Luddism, the movement of destroying the machines that were taking over workplaces). Unfortunately, many socially and environmentally conscientious people endorse Proudhon-like ideologies, underestimating Marx’s insightful objections, consistent with his background in the works of Riccardo and Adam Smith, leading economists at the previous epoch. Technological progress seems required for socio-economic improvements, and—I daresay, although I argued for this elsewhere—even more so for environmental sustainability (Boltuc 2015; Bisk and Boltuc 2017). On this theory of progress, the economy is objective (in Hegelian lingo, it is *real*),

the way wolves in the forest were *real* to humans in tribal times—meaning that paying attention to them was essential to our survival conditions and mode of existence.

Braidotti joins the Romantics, Rousseau style. For reasons of both oppression and the ecological exploitation of nature, she wants us to, and predicts we will, step down from the scene. Floridi seems to hit a similar lately, although with a constructive, even heroic, overtone, coming from Picco del Mirandola (Floridi 2016). I daresay that this reversion to the Romantic age, which was merely the antithesis of the Enlightenment, is charming in Floridi's paradoxical optimism, but troubling whenever nihilism towards ourselves takes over. Those of us living within developed civilization are reaching beyond that. We are getting to the stage (or, age) of synthesis. **Environmentally**, this translates into *sustainability through technological developments*. As my friends and I tend to say, we opt for *sustainability as growth* (Bisk and Boltuc 2017), the growth through advanced technology and also the multi-dimensionality of our existence, which reach beyond narrow Enlightenment-rationalism but also way beyond the elimination of human value due to a passive-aggressive form of nihilistic romanticism.

3.3. Transhumanism at the Edge of Chaos

In AI, such growth goes beyond the good old-fashioned AI, towards the so-called *computing at the edge of chaos*, based on controlled stochastic processes (Boltuc 2020). It reaches further than Braidotti's biologism. We can enjoy meaningful connections with humans and other animals, but also connect with those smart and sensitive beings based on silicon—namely, intelligent machines—the Church–Turing Entities.

Machines, which *originally* have been nothing more than human-made contraptions, now learn from experience and from huge datasets of all kinds, becoming slowly but surely the main intellects in our world—intellectually just behind (and soon, very much in front of) human beings. However, they would still function as a part of human culture, or so we hope.

3.3.1. Ethical Computers and 'Cultured AI'

The predominant view in Robot Ethics since at least 2004 is that robot ethics needs to be instilled in the very heart of robot programming and functioning (Waser 2015). There are programs to train various robots in robot ethics (Moor 2007) and semi-autonomous drones being trained in war ethics (Welsh 2018). Thus, at least some of the humanities are already a part of AI *education*, definitely in AI *training*. However, what about the phenomenology of human experience? Would machines be able to function in our world, as companions, partners, team-members, and maybe friends, if they fell short of the human cultural experience in some preeminent way?

The question emerges of what are the humanities good for in the hardly human-centered world of future AGI? While some intellectuals, activists and politicians panic and want to slow down progress (Matzinger 1993), others try to take good care of what is left of it (Floridi 2014). We can learn from Kant and the other philosophers discussed above that mathematics, science and ethics are objective to the point of being accessible to non-humanoid rational beings, although the point is controversial.

I pose that, **as long as we want to attain understanding with advanced AIs or AGIs, we need to interact with them at the level of the humanities**. Without this, they would become technocratic savages. One does not need a long story to see how harmful human technocrats (and single-dimensional people) not inculcated in the humanities may become when in charge, in politics, business management and other domains. (Of course, the humanities tend to bring in their own kind of craziness, but the balance within the humanities and between the humanities and 'practical' approach (Russell 1912), provides some balance).

Religion is too weird for AIs and hard to defend to a rational mind. If accepted, it may lead to fanaticism practiced, *which would be highly unfortunate and unpractical*.

The humanities are here to teach (among many vital things, about the human take on the world). It includes what Goertzel started arguing for in 2021: that human life is guided by paraconsistent models, not by strictly consistent algorithms. To function in an environment that is accessible to human beings, AIs also need to have this *paraconsistent* aspect of programming, which allows for creativity, as both Goertzel and Thaler illustrate in their pre-AGI, or early AGI projects.

While some people view the humanities as a uniquely human form of self-expression, we are learning, while combatting racism, sexism and elitism, that intelligence goes beyond any race, gender or nationality. I argue that it also reaches beyond the chemical *emergence base*, whether silicon- or carbon-based. The mind is not merely biological. It is good for any future metaphysics (not just for the narrowly humanoid-centered); however, to be compatible with the framework of human mind/s, AGI training requires a firm grasp of the humanities.

3.3.2. Humanities beyond Anthropocene

We may ask: What would non-human intelligence bring to the table? Additionally, what are we able to bring to the table if artificial general intelligence (AGI) becomes so much smarter than we are? **Are we wise enough, impartial enough, to carry the humanities over and beyond the anthropocene?**

This last question makes little sense to many humanists. Yet, more and more *digital natives*⁴ grow up into adults. They have the right to their version of *humanism*—and to create the right space for it. Arguably, Heidegger (Brook 2009) viewed *ethos* primarily as the metaphysical/phenomenological area of life of a living being. It would probably have gone beyond him, but not beyond Leibniz and Kant, to define this *ethos* outside of carbon-based entities, especially beyond humans into the post-Platonic realm of the mind.

Thus, we encounter the following puzzle: Will (and should) our generations let the humanities rise beyond the original ontological dependency on human beings? What would such humanities mean, and mean to whom? These are some of the questions that the upcoming transhumanities may be best positioned to address. This paper is meant as a harbinger of transhumanities, not as their bible. In the next and final section, I pose more questions than I answer, but that is what philosophy seems best at, as well as most useful for.

The main work in the humanities seems to be transdisciplinary, reaching above and beyond disciplinary boundaries and even methodologies. We have broad *Gestalts* of multifarious characteristics, which are transformed by a model through the imagination of the artist, critic, writer, reader and so on, to new ideas, permutations, and experiments (Damasio 2011; Thaler 2019)—this is what differentiates transhumanities from transdisciplinary sciences of other kinds. We allow the humanities to mesh with the arts and social and natural sciences according to somewhat unpredictable synergies and patterns.

4. Transhumanities, towards a Definition

Now is the time to properly explain the difference between interdisciplinary and transdisciplinary humanities. There were compressed explanations in the abstract and the early sections of this article: a rather straightforward one was given in Section 1 and contextual explanations appeared in Section 2—but these were not quite satisfactory enough. Later in this section, we explain how neither interdisciplinary humanities nor transdisciplinary humanities are the same thing as my proposed—and controversial—version of the concept of transhumanities. The first definition is regulative, the second one is stipulative, and the third one is a stipulation definition of special kind, which may be called a *projective definition* since it *designs* a certain relatively new conceptual tool⁵.

4.1. Interdisciplinary Humanities, a Regulative Definition

In Section 1, I talk of narrowly interdisciplinary research, characterized by working within 2–3 disciplinary areas on an issue in which their overlap, using disciplinary methods

and standards. Interdisciplinary research is viewed here as collaboration among disciplines (through their competent representatives) within their methodological standards and ontologies. One enters an interdisciplinary project, e.g., one related to the pandemic, where a microbiologist uses her knowledge in microbiology for the task of the team, and steps back into her disciplinary research when the joint project is over. Of course, this is an idealization, since one learns from one's teammates at the research collective, but that new competency belongs to other, clearly delineated domains and does not reshape one's practice of microbiology, except maybe the practical skills.

4.2. *Transdisciplinary Humanities, a Stipulative Definition*

We stipulate that transdisciplinary humanities like any other transdisciplinary research treats disciplines as a resource while building the main core of the new science among them. The sciences remain separate, but gravitas of the new area puts transdisciplinary approach in the center, not at the outskirts. We gave the examples of quantum physics and evolutionary biology. Within the humanities, the transdisciplinary approach would be a new, disruptive approach and ideology.

The term disruption comes from various applied theories, such as a theory of management or education. In business theory of digital transformation, a disruptive innovation is such that leaves the old entrenched players no defense due to two factors;

1. Disruptive technology that allows satisfying consumer needs in radically different ways (unique value proposition)
2. Unique value network, a set of talent, suppliers, enablers and other capacities that is not easy to replicate (Regersm 2016).

In the humanities, books are an important asset, so are traditional theater, ballet and concerts of classical music, but movies of various kinds, show, 4D programs, sophisticated games, spiritual lectures, self-help trainings populist music and so on, compete for attention, intellectual, spiritual, aesthetic and financial resources.

The trick in transdisciplinary humanities is to view the humanities in both senses: First, as the humanities themselves, for instance Dostoevski's novels, Beethoven's concerts, or Plato's dialogues. Second, as scholarship *about* the humanities, which are at the meta-level.

In the sciences empirical experiments, or systematic observations of nature are at the first level, while higher level theories are at the second. In the humanities, many tend to identify the humanities with literature and art criticism, its theory and even sociology. While it seems clear that, when we talk of education in the humanities, it is largely of getting disciples acquainted with the works of art, reading meaningful literature, poems or getting to know works of visual arts, and only secondarily being able to interpret them based in part on the science of the humanities, e.g., systematic analysis of literature and other pieces of art. Transdisciplinary humanities do not center upon any given ideology or method, be it structuralism, constructivism, Marxism or minimalism. It treats analysis at hand as the center of the theoretical universe and uses methods as tools, which are instrumental in attaining the most desirable outcomes—based on criteria largely free of methodological or ideological prejudice.

4.3. *Transhumanities, a Projective Definition*

Transhumanities shall be projected or designed, as a science that allows intellectual and pre-semantic, conceptual gestalt-based connection with advanced intelligence, including human beings and highly sophisticated AI of the right kind. We need to remember not to focus primarily on second level interpretations but on first level objects of the humanities, the humanistic 'things themselves', the carriers of meanings. They need to be open to various art media and styles but also to those non-existing yet, that may carry the most content perceivable for intelligence not very much like ours.

We design the following definition of projective definition, which offers a bit more liberty than a stipulative definition would. In Polish, there is a term 'projective definition', which has been well-entrenched in the Warsaw-Lwov school of analytical philosophy since

at least the 1920s; however, the term did not seem to migrate to the West⁶. The use of projective definition is handy in the rare cases when the idea seems novel enough.

4.4. Building towards Formulation of Transhumanities

There seem to be *two ways to revise, revert, or revolutionize a paradigm*. One is a *move from below*; it often starts with a detail, an opaque observation that leads to new, revolutionary theories. This approach is characteristic of natural (especially physical) sciences, such as Newton's, or Einstein's revolution.

The second method, less obvious in the natural sciences, proceeds from the top down. It starts with broad-view perspectives and subverts the very scaffolding of a discipline. I pose that computer science, with its move from the so-called "*Good-old-fashioned Artificial Intelligence*" (GOFAI) (Minsky 1985) to the search for Artificial General Intelligence (AGI), for instance, in its version proposed by (Goertzel 2006), exemplifies each of those strategies.

However, transdisciplinary science is not the same as transdisciplinary humanities; this is due to the methodological specificities of the humanities. We view the humanities as softer than even the soft sciences (such as psychology), since they do not just integrate the first-person consciousness in psychological sense. The humanities integrate, and in fact are **built upon**, *the content of the first-person perspective, its meaning and values*. This is how the humanities, which may be highly analytical, also need to leave room for metaphors, intuition and creativity, incorporating hermeneutic interpretations (not just social, sociological, or medical) of the arts, music, poetry, etc., as important aspects of humanistic experience and specifically humanistic intentionality.

Transhumanities (and, I suppose, the other transdisciplinary sciences as well) are those that transcend their preexisting paradigms. Paradigms seem to extend over a single science, or a set of *related* sciences (those that share the same assumptions and methodologies but differ primarily in their subject areas, as various extensions). Thus, in some cases, a single science may transcend its limits (as Kant transcended the limits of British empiricism).

4.5. Transhumanities further Stipulated

Transhumanities will not be understood here as just transdisciplinary humanities. Transhumanities are not just a blend of disciplines (as the term transgender is not viewed merely as a blend of genders, but a new quality, which is directly not rooted in its biological or socio-economic emergence-base). Transdisciplinarity presents the framework, in which we further define transhumanities as its very anomalous instantiation. Again, this is due to the special status of the very gist of the humanities—our grasp of the world, which is *always already* personal yet social (as shared through some formal or informal language and culture/s). This subjective, yet essentially also objective, grasp is the special miracle of creativity in the existential, not merely *psychological*, context of the human mind. Great work (Ingarden 1931)—Bach or Wagner; Beethoven or Chopin; Goethe or Dostoevsky; Da Vinci or Dali—is not merely the subject of the humanities: those are, *inter alia*, the humanities themselves. Our history of art, or literature, or accounts of those culture-facts, are the humanities only as long as they respect those artefacts as acts, creations, or products that are *within the humanities already* (Classen 2020). An art-critic who provides some statistics about the arts, even their social functioning (say, attendance at some galleries or what each artifact costs), with no phenomenological insights, puts himself outside of the humanities *tout-court*, with his or her work remaining a footnote, although some such footnotes turn out to be valuable on other grounds (e.g., for art-business, or sociology of the arts). Humanistic phenomenology is *always already* participatory in the humanities as their content, co-creating the perception of the work of art and thereby reshaping that work itself, since an artwork subsists in its readings as interpretations (Ingarden 1931). Reception, not just of the arts but also on and within the humanities, is the co-creation of human experience through the narratives—historic, literary, even psychological.

Philosophy, with all its justifiable analytical ambitions, is also a discourse aiming to grasp a glimpse of the human condition. Therefore, a stringent formalization of cri-

teria, which allows one to *siphon out* Heidegger, Jaspers, or Žižek as unprofessional due to their lack of adherence to some formal measuring stick, would likely end up under serious scrutiny.

5. Conclusions and Review: Transhumanities for AGI

Artificial Intelligence (AI), at present, is developing toward Artificial General Intelligence (AGI), the kind that can produce a human-level performance in most cognitive domains and, as it develops, can exceed the cognitive performance of the human intellect. A full explanation and justification of this claim can be found elsewhere (Goertzel 2006).

Let me clarify our terms. The current mainstream AI, based largely on neural networks, can be trained toward a specific task; however, it is not versatile in the sense of being easily retrained to perform other tasks, even somewhat similar tasks. However, several teams—including the international teams funded by DARPA (Siegelmann 2019)—are making substantial progress in this regard. Another problem has been the cost of retraining every neural network by human beings, which could take years. One *natural* solution, which turned out to not be as easy to implement as it may sound, is to *copy* the outcome of AI's *basic training* when it acquires universal skills and *paste* the internal structure of such an 'artificial brain' onto any number of systems, leaving only specific training in one-on-one *tutorials* for AIs; one such harness has been developed and implemented (Kelley 2020).

The ability to make *real discoveries* (*whatever that is supposed to mean*) has supposedly been reserved to human beings—according to the voices that still percolate in scientific community. Yet, already in 1950, Alan Turing responded to *the Lady Lovelace objection* from the 1800s that questions machine ability to create anything new. The gist of Turing's argument was that if human beings were able to make essentially new discoveries, then there is no functional difficulty in robots following the same algorithms and attaining required level of novelty (Turing 1950). In the 1990s, Stephen Thaler patented discovery engines (*AI for making useful discoveries*), which discovered, among other very useful and advanced things, the rotary electric toothbrushes for Oral-B. Nobody advised the AI cognitive engine to search in this direction. In fact, engineers initially rejected all designs involving the rotary element—however, since the preponderance of 200 or so designs incorporated this element, one of the managers decided to test their functioning, and all of us know how the story goes from there. At present Thaler's original cognitive engine is out of patent and in broader use (Thaler 2017). However, *Discovery Engines company* patented a more advanced cognitive engine, DABUS (Thaler 2019), which makes discoveries in a way that would be called independent if any human engineer was making them, with the same level of prompting⁷.

A claim has been made that, in order to possess fully human-level consciousness, functional equivalence is not enough (Boltuc 2007, 2009, 2012, 2018a, 2020); this is because fully conscious machines would have to have first-person consciousness or awareness, as defined in Chalmers' so-called *Hard Problem of Consciousness* (Chalmers 1995). Again, this is a conversation for another day. However, the above topic relates to the humanities in very practical terms. One would have to decide how to treat the works (including professional articles) submitted by AI as the sole author or co-author. This is of serious practical importance: If people apply for patents on behalf of AIs, then submitting articles on their behalf is an easy step in a similar direction, but *this may not be a pressing issue for a while*.

Transhumanities for the Future

Is there a happy, long-term perspective to our digital future? Let me quote a beautiful sentence by Herman Hesse, which I quote after Albrecht Classen's article (Classen 2021): "Joculator Basiliensis (. . .) invented for the Glass Bead Game the principles of a new language, a language of symbols and formulas, in which mathematics and music played an equal part, so that it became possible to combine astronomical and musical formulas, to reduce mathematics and music to a common denominator" (Classen 2020). This sentence,

dating back to the 1940s, preempted the full sensory integration developed by O'Regan and others, which allows us to view various senses⁸, and, we may add here—as well as mathematics and other intellectual capabilities—a continuum of meanings (O'Regan 2012). In a world of full psycho-motoric and psycho-intellectual integration, people are playing with this multifarious reality of meanings, filling each object and every interaction with sub-symbolic, sub-semantic shades of various kinds: emotional, representational, and those satisfying practical needs.

Perhaps we could extend our sensitivity, our intellectual reach, with little, if any, disciplinary cutting lines, instead viewing these lines as connective nodes of various modes of communication, memory and, more broadly, of existence. This continuity of experience is free of language-related or historically rooted limitations and idiosyncrasies (Boltuc 2018b). All of this is attained through a multi-sensory, multi-language (including math, predicative logic's languages, intuitions, images and music) integration of communication and meanings. This enables first-person streams of awareness that are non-reducible to their content, just like the stream of light is not reducible to the shadows that are left by the objects illuminated by it. This is a sketch of a highly utopian story of future transhumanistic humanism, attained through really smart AI working with us in unison, with us also changed by the technology in major ways, though not necessarily all the way to cyborgization (Sorger 2019).

In practical terms, the trend towards narrow professionalization needs to be balanced by the opposite tendency, towards multi-model thinking and research. The human mind seems to be operating through multiple mind maps (Damasio 2011) and is quite capable of viewing things from multiple angles, paradigms or *Gestalt* forms. This trend is often multi-disciplinary, building bridges among two or three disciplines and combining their professional paradigms in terms of methodologies, scopes and research communities. Soon, those communities may need to include truly smart AIs armed not only with technical knowledge (Boltuc 2018a) but also in depth in the humanities. Otherwise, these potential partners will fail to be inculcated in wider human civilization—the culture, values, traditions and meanings we live by—which would amount to a huge failure in communication⁹.

Funding: This research received no funding.

Conflicts of Interest: The author declares no conflict of interest.

Notes

¹ According to Darian-Smith and McCarthy (below).

² This is nicely envisaged in Figure 1. Other similar graphs present transdisciplinarity as just one horizontal blub. This would indicate the end of disciplinary knowledge or competencies, which the author views as misguided.

³ 'A new scientific truth does not triumph by convincing its opponents and making them see the light, but rather because its opponents eventually die, and a new generation grows up that is familiar with it.' This principle was famously laid out by German theoretical physicist Max Planck in 1950 and it turns out that he was right, according to a new study.

⁴ Luciano Floridi and many others use this term for those people born and raised in the environment of persistent interactions with IT technologies.

⁵ Quite often, interdisciplinarity is exactly what we need. For instance, if I need to know about Abraham Lincoln's views and attitudes toward Native Americans, we should consult a historian specializing in Lincoln studies and learn the facts. Yet, this idealization is often quite naïve since it is neglecting disciplinary and attitudinal context-dependencies. Lincoln scholars are often understated about the issue, since the 16th president of the US was by far more compassionate concerning the group we now call African Americans than another group called now Native Americans. A scholar specializing in Native American studies, preferably from a place that identifies with their tradition, would be most likely to tell us that Lincoln's grandfather was killed by the Native Americans, which partly shaped Abe Lincoln's attitudes in the matter. This is to show that interdisciplinarity, in the humanities and beyond, never presumes full objectivity of the experts, or lack of room for their fruitful discussion and interdisciplinary disagreement, as well as consensus. Those, however, do not necessarily turn a Lincoln-scholar into a scholar in Native American studies and *vice versa*—even though they may gain some shared competencies.

⁶ In Polish there is a term *projecting definition* (*definicja projektująca*), which most people translate as stipulative definition; however, stipulating is different than designing or projecting. Thus, the Polish version that comes from the Lwov-Warsaw school of

philosophy (probably Twardowski and Ajdukiewicz) as well as the German philosophical tradition, is more radical than the regular stipulative definitions. Thus, I treat it here as a sub-class of stipulative definitions characterized by more radical level of projecting (not merely gauging) but not quite re-designing all the way.

- 7 As a side note that develops what was mentioned in the main text: There were a number of patent applications for DABUS filed lately; some were already rejected, but only for philosophical or legalistic reasons (either there is no legal procedure to award patents to robots and AI, or it is viewed as ethically or anthropologically inappropriate). In any instance, it is not controversial that most advanced Artificial Intelligence, for the last decade or so, has been able to make (*in principle* patentable) independent discoveries.
- 8 This is visible in the musical glasses, a device for people with non-working eyes but working visual cortex, where sounds are perceived as colors and shapes. In a matter of months many people regain the main aspects of space understanding based on different qualia replacing vision.
- 9 I want to thank Albrecht Classen, Editor of this journal, for encouragement and helpful comments on several versions; Meredith Cargill and other members of the Central Illinois Philosophy group for essential comments on an earlier version; my Senior Seminar students at UIS for their input, as well as anonymous reviewers for this journal for their important recommendations. Special thanks go to Christina Takahashi from MDPI Japan for exemplary editorial help.

References

- Bisk, Tsvi, and Peter Boltuc. 2017. Sustainability as Growth. In *Technology, Society, Sustainability*. Edited by Lech Zacher. Berlin/Heidelberg: Springer, pp. 175–83.
- Boltuc, Peter. 2007. (co-author) Replication of the Hard Problem of Consciousness in AI and Bio-AI: An Early Conceptual Framework. In *AI and Consciousness: Theoretical Foundations and Current Approaches*. Edited by Antonio Chella and Riccardo Manzotti. Merlo Park: AAAI Press, pp. 24–29.
- Boltuc, Peter. 2009. The Philosophical Problem in Machine Consciousness. *International Journal of Machine Consciousness* 1: 155–76. [\[CrossRef\]](#)
- Boltuc, Peter. 2011. What is the Difference between Your Friend and a Church Turing Lover. In *The Computational Turn: Past, Presents and Futures? Proceedings IACAP 2011*. Edited by Charles Ess and Ruth Hagengruber. Aarhus: Aarhus University, pp. 37–40.
- Boltuc, Peter. 2012. The Engineering Thesis in Machine Consciousness. *Techné: Research in Philosophy and Technology* 16: 187–207. [\[CrossRef\]](#)
- Boltuc, Piotr. 2015. *Sustainability and Growth (A Controversy)*. Gliwice: Organization and Management, Śląsk Politechnics, pp. 27–48.
- Boltuc, Piotr. 2017. Church-Turing Lovers. In *Robot Ethics 2.0: From Autonomous Cars to Artificial Intelligence*. Edited by Patrick Lin, Keith Abney and Ryan Jenkins. Oxford: Oxford University Press.
- Boltuc, Peter. 2018a. BICA à Rebouris. In *Advances in Intelligent Systems and Computing. Biologically Inspired Cognitive Architectures*. Edited by Alexei V. Samsonovich. Berlin/Heidelberg: Springer, pp. 27–43.
- Boltuc, Piotr. 2018b. Strong Semantic Computing. *Procedia Computer Science* 123: 98–103.
- Boltuc, Piotr. 2020. Conscious AI at the Edge of Chaos. *Journal of Artificial Intelligence and Consciousness* 5: 25–38. [\[CrossRef\]](#)
- Braidotti, Rosi. 2018. The Posthuman Condition and the Critical Posthumanities, Lecture November 2. Available online: <https://rosibraidotti.com/2019/01/28/the-posthuman-condition-and-the-critical-posthumanities/> (accessed on 15 December 2021).
- Brook, Angus. 2009. *The Early Heidegger and Ethics: The Notion of Ethos in Martin Heidegger's Early Career*. Saarbrücken: VDM Verlag.
- Chalmers, David J. 1995. Facing up to the problem of consciousness. *Journal of Consciousness Studies* 2: 200–19.
- Chawla, Dalmeet Singh. 2019. Science Really Does Advance One Funeral at a Time, Study Suggests" Chemistry World 17 September 2019. Available online: <https://www.chemistryworld.com/news/science-really-does-advance-one-funeral-at-a-time-study-suggests/3010961.article> (accessed on 15 December 2021).
- Classen, Albrecht, ed. 2020. *Paradigm Shifts During the Global Middle Ages and Renaissance*. Arizona Studies in the Middle Ages and the Renaissance (ASMAR 44). Turnhout: Brepols.
- Classen, Albrecht. 2021. Transdisciplinarity—A Bold Way into the Academic Future, from a European Medievalist Perspective and or the Rediscovery of Philology? *Humanities* 10: 96. [\[CrossRef\]](#)
- D'Alembert, Jean Le Rond. 1761. *Preliminary Discourse to the Encyclopedia of Diderot*, English editions 1995. Translated by Richard N. Schwab, and Walter E. Rex. Chicago: University of Chicago Press.
- Damasio, Antonio. 2011. *Self Comes to Mind: Constructing the Conscious Brain*. New York: Random House.
- Darian-Smith, Eve, and Philip McCarty. 2016. Beyond Interdisciplinarity: Developing a Global Transdisciplinary Framework. *Transcience* 7: 1–26.
- Davidson, Donald. 1985. Rational Animals. In *Actions and Events: Perspectives on the Philosophy of Donald Davidson*. Edited by Ernest Lepore and Brian P. McLaughlin. New York: Basil Blackwell.
- Deutsch, David. 1985. Quantum theory, the Church–Turing principle and the universal quantum computer. *Proceedings of the Royal Society* 400: 97–117.
- Floridi, Luciano. 2014. *The 4th Revolution. How the Infosphere is Reshaping Human Reality*. Oxford: OUP.
- Floridi, Luciano. 2016. On Human Dignity as a Foundation for the Right to Privacy. *Philosophy of Technology* 29: 307–12. [\[CrossRef\]](#)
- Goertzel, Ben. 2006. *The Hidden Pattern. A Patternist Philosophy of Mind*. Boca Raton: BrownWalker Press.

- Ingarden, Roman. 1931. *Das Literarische Kunstwerk*. Tübingen: Max Niemeyer Verlag.
- Kant, Immanuel. 1781. *Critique of Pure Reason*. English Edition 1999. Translated and edited by Paul Guyer and Allen W. Wood. Cambridge: Cambridge University Press.
- Kant, Immanuel. 1797. *Foundations of the Metaphysics of Morals, and What Is Enlightenment?* English Edition 1959. Translated with an introduction by Lewis White Beck. New York: Liberal Arts Press.
- Kelley, David. 2020. Preliminary Results and Analysis of an Independent Core Observer Model (ICOM) Cognitive Architecture in a Mediated Artificial Super Intelligence (mASI). In *Biologically Inspired Cognitive Architectures 2019: Proceedings of the Tenth Annual Meeting of the BICA Society*. Edited by Alexei V. Samsonovich. Berlin/Heidelberg: Springer, pp. 179–86. Available online: <https://www.springer.com/us/book/9783030257187> (accessed on 15 December 2021).
- Kuhn, Thomas. 1969. *The Structure of Scientific Revolutions*, 2nd ed. Chicago: University of Chicago Press.
- Libet, Benjamin. 2004. *Mind Time: The Temporal Factor in Consciousness, Perspectives in Cognitive Neuroscience*. Cambridge: Harvard University Press.
- Marx, Karl. 1947. *The German Ideology*. London: International Publishers Co.
- Maturana, Humberto, and Francisco Varela. 1980. *Autopoiesis And Cognition: The Realization of the Living Is a Cybernetic Work in Systems Theory and the Philosophy of Biology*. Dordrecht: D. Reidel Publishing Company.
- Matzinger, Thomas. 1993. *Subjekt und Selbstmodell. Die Perspektivität Phänomenalen Bewußtseins vor dem Hintergrund einer Naturalistischen Theorie Mentaler Repräsentation*. Edited by Paderborn Mentis. Cincinnati: Mentis.
- Mccarty, Philip. 2016. Image. Available online: https://www.researchgate.net/profile/Philip-Mccarty-2/publication/311486261/figure/fig1/AS:436722644328500@1481134200855/Transdisciplinarity_W640.jpg (accessed on 15 December 2021).
- Minsky, Marvin. 1985. *The Society of Mind*. New York: Simon and Schuster.
- Moor, James. 2007. Taking the Intentional Stance Toward Robot Ethics. *APA Newsletter* 6: 14–17.
- O'Regan, Kevin O. 2012. *Why Red Doesn't Sound Like a Bell: Understanding the Feel of Consciousness*. Oxford: Oxford University Press.
- Piaget, Jean. 1972. The Epistemology of Interdisciplinary Relationships. In *Interdisciplinarity: Problems of Teaching and Research in Universities*. Edited by Leo Apostel, G. Berger, A. Briggs and G. Michaud. Paris: Organization for Economic Cooperation and Development, pp. 127–39.
- Plato. 1968. *The Republic*. Translated by Allan Bloom. New York: Basic Books.
- Regersm, David. 2016. *The Digital Transformation Playbook*. New York: Columbia University Press, Business School Publishing.
- Russell, Bertrand. 1912. US Publication 1998 "The Value of Philosophy". In *The Mayfield Anthology of Western Philosophy*. Edited by Daniel Kolak. Mountain View: Mayfield Publishing Company, pp. 1070–73.
- Rosemary, Johnston. 2008. On Connection and Community: Transdisciplinarity and the Arts. In *Transdisciplinarity—Theory and Practice*. Edited by Basarab Nicolescu. Cresskill: Hampton Press Inc., pp. 223–36.
- Siegelmann, Hava. 2019. Lifelong Learning in Nature and Machines; US Defense Advanced Research Projects Agency (DARPA) 4 November 2019, Presented to COMCAS 2019, Tel Aviv, Israel. Available online: http://2019.comcas.org/Portals/107/Hava%20T_%20Siegelmann.pdf (accessed on 15 December 2021).
- Sorger, Stefan Lorenz. 2019. Transhumanism: The Best Minds of Our Generation Are Needed for Shaping Our Future. *Newsletter on Philosophy and Computers of the American Philosophical Association* 18: 15–18.
- Spinoza, Benedict. 1677. *Ethics*. English Edition 2005. Translated by Edwin Curley. New York: Penguin.
- Thaler, Stephen. 2017. *Encyclopedia of Creativity, Invention, Innovation and Entrepreneurship*. Edited by Elias G. Carayannis. Berlin/Heidelberg: Springer Science+Business Media LLC.
- Thaler, Stephen. 2019. DABUS in a Nutshell. *APA Newsletter on Philosophy and Computers* 19: 40–42.
- Turing, Alan. 1950. Computing Machinery and Intelligence. In *Mind*. Cambridge: MIT Press, Volume LIX, pp. 433–60.
- Velmans, Max. 2009. *Understanding Consciousness*. London: Routledge.
- Waser, Mark. 2015. Designing, Implementing and Enforcing a Coherent System of Laws, Ethics and Morals for Intelligent Machines (including Humans). *Procedia Computer Science* 71: 106–11. [[CrossRef](#)]
- Welsh, Sean. 2018. *Ethics and Security Automata*. Abingdon and New York: Routledge.