Filozofski vestnik
ISSN 0353-4510 | Volume 45 | Number 2 | 2024 | 329–50
 | DOI: 10.3986/FV.45.2.15

Aleš Mendiževec*

The Con-Fusion of Artificial Intelligence[1]

Keywords
artificial intelligence, deep learning, conditions of AI, artificial general intelligence, antihumanism

Abstract
The article discusses various conditions of contemporary artificial intelligence, namely deep learning mechanisms, to emphasize its limitations and argues for an antihumanistic view of contemporary technology. It starts from affirming Turing test and argues that machines can in fact be intelligent but that this intelligence must not be related to a capitalistically hyped idea of artificial general intelligence. Then it outlines various conditions on which deep learning depends in its functioning (brute computing power, capitalist datafication, the world of contingency). These conditions show an epistemological schism in the field of artificial intelligence (between symbolic AI and connectionism) that could be overcome by getting rid of the idea of artificial general intelligence and the competitive relation between humans and machines.

Zmeda pogojev umetne inteligence

Ključne besede
umetna inteligenca, globoko učenje, pogoji UI, splošna umetna inteligenca, antihumanizem

1 This article is a result of the research project Z6-50210 "Nonlinear Causalities: The Problem of (Un)Predictability in Contemporary Technology," which is funded by the Slovenian Research and Innovation Agency.

* ZRC SAZU, Institute of Philosophy, Ljubljana, Slovenia
alesme@zrc-sazu.si

Povzetek
Članek obravnava različne pogoje sodobne umetne inteligence, pri čemer se osredotoča predvsem na globoko učenje. Omejitve umetne inteligence izpostavlja z zagovarjanjem antihumanističnega pogleda na sodobno tehnologijo. Začne z afirmiranjem relevantnosti Turingovega testa in z njegovo pomočjo zagovarja trditev, da so stroji lahko inteligentni, vendar da ta vrsta inteligence ne sme biti povezana s kapitalistično navdahnjeno idejo splošne umetne inteligence (t.i. artificial general intelligence). V nadaljevanju oriše različne pogoje funkcioniranja mehanizmov globokega učenja (surova moč komputacije, kapitalistično upodatkovljenje, svet kontingence). Ti pogoji razkrijejo epistemološko shizmo znotraj področja umetne inteligence (simbolna UI in konekcionizem), ki jo lahko presežemo samo, če se znebimo ideje splošne umetne inteligence in tekmovalnega razmerja med človekom in strojem.

∞

In a way, it is quite simple. If a machine—or anything else for that matter—can imitate intelligent acts, and we, intelligent beings who interact with it, cannot tell if we are interacting with a human being or a machine, we have to affirm that our interlocutor cannot be unintelligent. It is hard to prove intelligence, but it is infinitely easier to demonstrate its possibility. If this artificial entity can imitate intelligence and fool a naturally intelligent being as to its artificiality, then this being has no choice but to affirm the possibility that the fooling entity is also intelligent. At least this is how a popularized version of the Turing test goes.[2]

We might try to complicate it and savvily shift the attention to the human side of the test and object that it presupposes that the human being is intelligent and claim that he is not as intelligent as he thinks he is. That he is himself actually more of a machine playing the game of "as if," imitating himself as a wannabe intelligent being in the act of judging others (un)intelligent. That he is a Pascalian

2 In the original version of the imitation game, as Turing calls it, there are two imitators, one is a machine and the other a human being. The latter can even try to help the interlocutor, while the former has to react in an unartificial manner. If the interlocutor cannot tell who is a human and who a machine, then the machine passes the test. Nonetheless Turing himself mentions a variety of the game that is the same as our version. See Alan Turing, "Computer Machinery and Intelligence," in *The Essential Turing: Seminal Writings in Computing, Logic, Philosophy, Artificial Intelligence, and Artificial Life; Plus The Secrets of Enigma*, ed. Jack Copeland (Oxford: Oxford University Press, 2004), 452.

automaton that made himself practice acts of judging intellect until—in some sort of sci-fi scenario of Althusserian interpellation—he himself started to believe in his own intelligence. That he is only a machine, but with a God complex. If we ascribe to an antihumanist position, this might be the right thesis. But this kind of approach would be a reactionary antihumanism that displaces the problem of artificial intelligence, too busy disgracing the human to even consider the potential intelligence of the machine. We can, of course, legitimately pose different problems from different theoretical positions, but if we consider artificial intelligence as artificial intelligence, then the right problem is only one. Its problem is not the human and not intelligence, but artificiality: not what intelligence is, but how to build it. Not even if we can build it, but how to do so.

We now might want to—in a specifically philosophically arrogant way—critique its unthought presuppositions, namely its idea of intelligence. We would not get far, because we would immediately see that the field of artificial intelligence (AI)[3] is full of reflections on intelligence in trying to define its various characteristics ("What is a rational agent?"; "What are its subfields?"; etc.), that it is full of interdisciplinary attempts at exploring key cognitive mechanisms from different perspectives and so on.[4] The field of AI is in fact orientated towards building an intelligent entity, but that does not mean that it just presupposes a certain notion of intelligence without any analysis, that it neglects different aspects of intelligence. Even if it is practically oriented, it is by no means a naive science. It is even alive, a living science, incorporating its critiques and developing through

3 We will use the shorter version (AI) for the field of computer science and the longer version (artificial intelligence) for the entities that they build.

4 On this we have to disagree with Franchi and Güzeldere, who criticize the lack of an interdisciplinary approach in the field of AI. They confuse arrogance for hermeticism (and in this aspect, cyberneticians were no better, contrary to what Franchi and Güzeldere claim). Computer scientists might be antimetaphysic positivists, who think that mathematics and quantification are the only way to formalize a field of knowledge, as Phillip Agre (himself a former computer scientist) explains it, but even so they are themselves (at least nowadays) fully engaged in studying philosophy, cognitive psychology, physiology, and brain neurology, even interested in sociological and ethological themes, etc. Regarding the former, see Stefano Franchi and Güven Güzeldere, "Machinations of the Mind: Cybernetics and Artificial Intelligence from Automata to Cyborgs," in *Mechanical Bodies, Computational Minds: Artificial Intelligence from Automata to Cyborgs*, ed. Stefano Franchi and Güven Güzeldere (Cambridge: MIT Press, 2005), 15–149; regarding the latter, see Philip Agre, "The Soul Gained and Lost: Artificial Intelligence as a Philosophical Project," in Franchi and Güzeldere, *Mechanical Bodies*, 153–73.

them. Computer scientists who are designing artificial intelligence are very much aware of its shortcomings and of the critiques that are being addressed to them. We probably would not be very far off if we imagined Rodney Brooks, the famous robotics engineer, as a reader of Jean-Pierre Dupuy and his classic study *On the Origins of Cognitive Science* (subtitled *The Mechanization of the Mind*), trying to overcome the disembodiment issue and embody mechanisms of artificial intelligence so that there might somehow, as some would claim, emerge a lived experience and with it a truly purposive self-organized artificial entity, this ultimate cybernetic dream.

It is of no use to judge these attempts as possible or impossible. The usual objections that artificial intelligence does not understand anything and just follows rules, that it is disembodied and has no lived experience, that it has no common sense, etc., in one word, that it is not humanly intelligent and that it never will be, all have a modern humanist basis. Humanists, too busy protecting the human to even consider that the machine might have its own specific intelligence. This story of humanist attacks on AI reads as a constant downplaying of the latter's achievements. But it is getting harder and harder to downplay or ignore the achievements and capabilities of contemporary technology, which are forcing humanists to endlessly "update" (practical examples in) their critiques.

Today it is in no way naive to think that an intelligent machine could be built or even that it has already been built: after Deep Blue beat grandmaster Garry Kasparov in a chess match in 1997, after Watson won *Jeopardy!* in 2011, after AlphaGo beat the Go champion Lee Sedol in 2016 and was itself beaten a year later by its self-taught successor AlphaGo Zero (with a score of 100–0), after the deep-learning boom starting with AlexNet winning the ImageNet competition in 2012 and the subsequent emergence of generative AI and the ChatGPT explosion in 2022, which this time shook the whole world, renewing debates about so-called artificial general intelligence or AGI—artificial intelligence that would reach the level of human cognition.[5] Man-made machines have beaten and surpassed humans in almost every game that had an aura of that kind of intelligence that is accessible only to human geniuses—the now stolen crowning jewels of

5 Regarding the magnitude of the impact of ChatGPT in trends of technological developments and economical investments, see Eliza Strickland, "15 Graphs That Explain the State of AI in 2024," *IEEE Spectrum*, April 15, 2024, https://spectrum.ieee.org/ai-index-2024.

human intelligence. From AGI to technological singularity,[6] the beginning of the 21st century, with its unforeseen technological breakthroughs, not only revived the old dreams of (let us say with a bit of spiteful provocation) "more-than-human" intelligence, but even repainted the picture (if you are a fan of Dall-E) or rewrote the narrative (if you are a fan of ChatGPT) of what an infallible human or even God-like intelligence would look like and how it would function, making it increasingly more tangible.

So, for a second, let us hold off on the usual humanistic objections; let us nonetheless be antihumanist on this point—but let us also put on hold the savvy displacement and stick to machines. This does not mean that we have to outright reject humanist critiques, which are by themselves actually quite legitimate. What this does mean is that we have to suspend them at this point and change our theoretical position and with it the problematic so that the correct question that pertains to it could be posed: not to pose the question of whether a machine is intelligent, not even to pose the question of whether a human being is intelligent, but rather to posit that if a human as an intelligent being is itself some sort of a machine, then a machine might as well be intelligent in its own (artificial) way.

Humans as some sort of machine? Yes, some sort; for we are not claiming that a human being is an information-processing machine, as the vulgate of cognitive scientists goes. What we are claiming is that a human being is itself a machine in the sense of an automaton, that he is not a creature capable of fully-transparent self-consciousness and is in fact governed by and is the effect of automatic processes that he himself cannot control. Not every automation implies digitalization, we have to be wary of that.[7] So, to elaborate our thesis differently, again and

6 Technological singularity is a point where artificial intelligence becomes capable of improving itself and "takes off" on its own, leaving humans behind in the process of the development of intelligence. Or in the words of its most notorious proponent, Ray Kurzweil: "There are actually two schools of thought on the singularity: there's a hard take off school and a soft take off school. I'm actually in the soft take off school that says we will continue to progress exponentially, which is daunting enough. The idea of an intelligence explosion is that there is a magic moment where a computer can access its own design and modify it and create a smarter version of itself, and that it keeps doing that in a very fast iterative loop and just explodes in its intelligence." Ray Kurzweil, interview by Martin Ford, in *Architects of Intelligence: The Truth about AI from the People Building It* (Birmingham: Packt, 2018), 238.

7 We have to have in mind that automation does not necessarily mean digitalization and with it information-processing for the following reason: Jean-Pierre Dupuy, in his great

for the final time: if a human being is itself some sort of machine, then instead of claiming that intelligence is exclusively allocated, we have to claim that it is inclusively dispersed, in some way present also in contemporary machines, which are just built differently—and this would define the specificity of its intelligence.

Let us be bluntly direct: we are not humanistic technophobes; we suppose that humanity has in fact succeeded in building some sort of intelligence. But this does not amount to advocating AGI. We have to insist on that, because we are also not antihumanistic (procapitalist) technophiles.[8] The less savvy antihumanists also love to disgrace the human, but by emphasizing the superior functioning of machine intelligence. That it is just like human intelligence but infinitely better (faster is what this really means). But AGI has turned into a capitalist trap: Is not the desired "generality" of artificial intelligence more than anything the flexible universality of an automated tool, useful for anyone and applicable to everything and at any time and any place, the ultimate capitalist product capable of satisfying the needs of each and every singular consumer? Instead of advocating AGI, we want to open up the possibility for the specificity of artificial intelligence. That artificial intelligence is specific means precisely that it is different than human intelligence and as such does not—or that at least "by itself" would not—necessarily strive towards human intelligence's "generality" and to become superhuman.

The Turing test would most definitely validate our thesis, precisely because of its simplicity. Today, most computer scientist do reject the Turing test, but not because it is wrong, rather they reject it because it was confirmed and became useless to them—too broad for a conceptual differentiation of intelligence and

study *On the Origins of Cognitive Science: The Mechanization of the Mind* (trans. D. B. Debevoise [Cambridge: MIT Press, 2009]), accuses some antihumanist structuralists (Lacan, Althusser), on which we are relying here, of adopting cybernetic metaphors of mechanization when in fact they merely discussed the human as an automaton in the sense that he is not—if, for the sake of the comparison, we use cybernetic terminology—self-regulated by his fully-transparent consciousness. This difference is precisely why today, pace Dupuy, an antihumanist (structuralist) approach might prove far more fruitful than the old phenomenological or humanist approach that is stuck on the critique of the lack of meaning and sense due to the disembodiment of the digital or virtual world of artificial intelligence.

8 Indeed, it is hard to be an antihumanist today. We constantly hear Althusser's voice asking: *Est-il simple d'être antihumaniste en philosophie (de technologie)?*

too narrow for practical experimentation.[9] But in its outdated provocativeness it is just simple enough for a correct theoretical positioning in our contemporary conjuncture. The Turing test today falls nicely in between technophilia and technophobia, demonstrating the possibility of a different kind of intelligence without attaching it to human "general" intelligence.

We will therefore admit the success of contemporary technological development and we will suppose (with Turing) that engineers have in fact built some sort of intelligence; but we will do that exactly to avoid the hype around the latest advancements with the deep learning "revolution" that media sensationalism and the capitalist drive for financial investment and future profits produced. Because we will not dwell on the AGI dilemma, which is the name of this hype game. As we have said, the correct problem is not intelligence but artificiality. But even this would be too big of an issue for this occasion; our attempt is actually much humbler. We will not ask what exactly this specific intelligence of machines is, in what way artificial intelligence as intelligence is specific; instead of defining machine intelligence in its difference to human intelligence, we will merely try to outline its different sources to see what it is made of, we will try to present various conditions that enabled this artificiality to be built in order to open up the possibility of its specificity.

And here, on the side of artificiality, the story gets complicated, even to the point of total confusion. Here the practical success of AI itself proves to be theoretically problematic, infusing confusion into its conceptual core. The deep learning hype is in fact blowing up promises and making speculations that have no substantial basis.[10] Deep learning methods have proved to be so powerful that it has be-

9 For example, Yann LeCun, one of the key architects of the deep learning "revolution," states the following: "The Turing test is not actually an interesting test. In fact, I don't think a lot of people in the AI field at the moment consider the Turing test to be a good test. It's too easy to trick it, and to some extent, the Turing test has already been and gone. [. . .] There is a whole component of intelligence that has nothing to do with language, and we are ignoring this if we reduce AI to just satisfying the Turing test." Yann LeCun, interview by Martin Ford, in *Architects of Intelligence*, 129.

10 Stuart Russell, one of the authors of the standard textbook on artificial intelligence, for example, claims: "What I see in a lot of the discussions and presentations from people talking about this is that there's probably an over-estimate of what current AI technologies are able to do and also, the difficulty of integrating what we know how to do into the existing extremely complex functionality of corporations and governments, and so on." Stuart

come hard to define their limits, especially since the hype is not produced by the enthusiasm of engineers, who have, at least some of them, outlived one or even more of the so-called AI winters, but mostly by the capitalist overflow of investment in future applications turned consumer-oriented personalized assistants.

Let us take, for example, Gary Marcus, a cognitive scientist and a hard-line critic of the deep learning "revolution." In his article *Deep Learning: A Critical Appraisal* from 2018, just a couple of years before ChatGPT entered the public stage, he mentions that "chatbots in general have not lived up to the hype they received a couple years ago."[11] This statement of course did not age well. It probably comes as no surprise that he later doubled-down on his critique when considering ChatGPT. Even if his arguments are sound, they are nonetheless pretty much one-sided (for example, considering so-called hallucinations as a deadly chatbot sin), which is probably also the reason why he gets such harsh treatment from the deep learning architects, such as from its "godfather" Geoffrey Hinton. The lesson here is not about who is right and who is wrong, nor which side will prevail; the lesson is that, despite sound theoretical criticism, practical success nonetheless arrives and this makes it increasingly tougher to draw the theoretical limits of deep learning and with it the demarcation lines that define the field of AI.

What, therefore, is at stake here is the field of AI itself and what is causing this problem is the obscured difference between deep learning and artificial intelligence as such. Geoffrey Hinton does not hide his convictions that the future of artificial intelligence is deep learning with its neural nets, and that conventional or symbolic artificial intelligence with its knowledge representation and problem-solving mechanisms is obsolete, because it is "just wrong," and that the idea of hybrid systems which would combine deep learning and neural nets with the symbolic processing of conventional artificial intelligence is "just an attempt to hang on to the view they already have, without really comprehending that they're being swept away."[12] Deep learning will take over the whole field of AI; it will become artificial intelligence itself, this is Hinton's bet.

Russell, interview by Martin Ford, in *Architects of Intelligence*, 55. Some even fear that the hype will bring about another so-called AI winter. See Gary Marcus, "Deep Learning: A Critical Appraisal," *Arxiv* (January 2018), https://doi.org/10.48550/arXiv.1801.00631.

11 Marcus, 17.

12 See Geoffrey Hinton, interview by Martin Ford, in *Architects of Intelligence*, 78, 85.

There is another reason for this confusion of AI. It is completely practical, namely, that machine learning and deep learning mechanisms have now become integrated into most basic computer programs; they are present almost everywhere and we probably use them daily even if we do not realize it. A lot of programs and applications that we use and that do not seem to learn or even be intelligent in their most basic functions in fact do include artificial intelligence or even machine learning (Microsoft Word, for example, which I am using now, is helping me with my phrasing; or Google Translate, which is helping me with my poor English vocabulary), as these mechanisms were progressively built in. This is also the reason why computer scientists are constantly emphasizing (in their interviews) the difference between deep learning, machine learning, and artificial intelligence, that not everything is deep learning, that deep learning is a specific kind of machine learning, which is a specific kind of artificial intelligence, besides the more classical symbolic mechanisms such as general-purpose search mechanisms, expert systems (based on a more domain-specific knowledge), probabilistic reasoning, etc.[13]

What this gradual and almost imperceptible integration of deep learning mechanisms into our technologically mediated lives and the instantly explosive success of their concepts, reshuffling the field of their origin, have resulted in is the eclipsing of their conditions. It is well known that deep learning mechanisms rely on the brute force of today's computing power. But deep learning itself is not a new technology; some of its core mechanisms can be traced back to the 80s when different computer scientists developed the idea of a backpropagation algorithm in relation to neural nets, among them Geoffrey Hinton.[14] And with gradual improvements in its architecture, such as Yann LeCun's convolutional neural network in the mid-90s (which was later used in the ImageNet competition, where deep learning first proved to be by far the superior AI tool), and the constant adding of extra neural layers—deep learning is deep precisely and only in the sense that it uses plural hidden layers of neurons between the input and

13 Here, we are relying on Stuart Russell and Peter Norvig, *Artificial Intelligence: A Modern Approach*, 4th ed. (Harlow: Pearson Education, 2022).

14 Even though he is sometimes credited with the invention of the backpropagation algorithm, he himself emphasizes that this is false. There were some unrelated attempts at defining this algorithm in 70s and 80s, but its main figure is David Rumelhart. Hinton credits himself with the invention of learning distributed representations via a backpropagation algorithm. See Geoffrey Hinton, interview, 73–75.

the output layers—the engineering framework of contemporary deep learning was built.[15] But it became efficiently functional and superior (to other AI tools) only when computing power improved enough to be capable of calculating the value of all the millions of its parameters and adjusting all the weights of the neural connections through backpropagation.

This brute force of computing power is essential for this mechanism to self-adjust, to effectively learn. As Thompson, Greenewald, Lee, and Manso explain: "Fortunately for such artificial neural networks [. . .] decades of Moore's law [i.e. that the number of transistors in an integrated circuit doubles about every two years] and other improvements in computer hardware yielded a roughly 10-million-fold increase in the number of computations that a computer could do in a second. So when researchers returned to deep learning in the late 2000s, they wielded tools equal to the challenge."[16] So, on one hand, deep learning is conditioned by the increase in computing power, which is approaching a limit because there is less and less physical space for the miniaturization of the transistors that are powering computers;[17] and on the other, there are problematic material effects and even ecological consequences of this technological improvement which are attracting ever increasing attention from critics.[18] The forgotten materiality is slowly turning its back on deep learning and contemporary technology

[15] See Charles C. Tappert, "Who Is the Father of Deep Learning?," *International Conference on Computational Science and Computational Intelligence (CSCI)* (2019): 343–48, https://doi.org/10.1109/CSCI49370.2019.00067.

[16] See Neil C. Thompson et al., "Deep Learning's Diminishing Returns," *IEEE Spectrum*, September 24, 2021, https://spectrum.ieee.org/deep-learning-computational-cost. For a bit more technical discussion, see Neil C. Thompson et al., "The Computational Limits of Deep Learning," *Arxiv* (July 2022), https://doi.org/10.48550/arXiv.2007.05558.

[17] They call it the end of Moore's law. Some are here wagering on the future role of quantum computing, 3D stacking, etc.; but because of the uncertainty of these nascent technologies, others are emphasizing that instead of focusing on this approach "from the Bottom," we should focus on approaches "from the Top," such as the performance-engineering of software, the development of algorithms, and hardware streamlining—but then again the gains from these approaches would be "opportunistic, uneven, sporadic, and subject to diminishing returns." See Charles Leiserson et al., "There's Plenty of Room at the Top: What Will Drive Computer Performance After Moore's Law?," *Science* 368, no. 6495 (June 2020), https://doi.org/10.1126/science.aam9744.

[18] "Extrapolating the gains of recent years might suggest that by 2025 the error level in the best deep-learning systems designed for recognizing objects in the ImageNet data set should be reduced to just 5 percent [. . .]. But the computing resources and energy required to train such a future system would be enormous, leading to the emission of as much car-

as such, revealing its problematic conditions, which were eclipsed by the engineering success and capitalistic technophilic dreams of the emergence of AGI.

There is another condition of deep learning mechanisms which is also pretty much known. Deep learning mechanisms would not learn much if there was nothing to learn from—all the endless data that is today available, mainly due to the internet. There are even disputes about large language models, such as ChatGPT, that have been trained on the "general intellect" of the internet—because these models are also shaped by the free labour of millions and millions of individuals—which are being used and exploited for capitalistic profits. It would be hard to deny that this exploitation is unjust; but instead of focusing on (problematic) effects, we have to, in our case at least, emphasize the conditions of its functioning. What is most important for us here is the sheer size of the data that is necessary to train deep learning mechanisms so that they become at least minimally efficient. There is no deep learning without big data, which means that it requires not only a huge amount of data but also diverse data so that datasets are not biased.[19] The greater the size and the greater the variety, the better the machine learning mechanism that was trained on them functions.

This practically begs for a massive data extraction apparatus, even greater than the one of the State established in the middle of the nineteenth century with the development of modern state institutions such as statistical bureaus. After the "avalanche of printed numbers," as the great historian of statistics Ian Hacking calls it,[20] came total "datafication," as Couldry and Mejias call it—data extraction became so widespread in the twenty-first century, they claim, that social relations themselves can now be defined as "data relations."[21] The reason for

bon dioxide as New York City generates in one month." Thompson et al., "Deep Learning's Diminishing Returns."

19 As Barbara Grosz explains: "The computer system can 'read all the papers' (more than a person could) and do certain kinds of information retrieval from them and extract results, and then do statistical analyses. But if most of the papers are on scientific work that was done only on male mice, or only on male humans, then the conclusions the system is coming to are limited." Barbara Grosz, interview by Martin Ford, in *Architects of Intelligence*, 346.

20 See Ian Hacking, *The Taming of Chance* (Cambridge: Cambridge University Press, 1990), 27–35.

21 See Nick Couldy and Ulises Mejias, *The Costs of Connection: How Data is Colonising Human Life and Appropriating It for Capitalism* (Stanford: Stanford University Press, 2019).

this is the change in the nature of processing and analysing data. As Couldry and Mejias explain: “Whereas the earlier social knowledge was built from specific calculations performed on so-called structured data (for example, entries in statistical tables and databases), today’s social knowledge can be built from unstructured data, drawn directly from the traces left in the flow of everyday life.”[22] The more machine learning mechanisms are integrated directly into life, or more precisely, directly into unstructured data traces of the flow of everyday life, the more (patterns) they can figure out, the more “knowledge” they can produce. Because they are capable of penetrating directly into unstructured data and predict their patterns without our “meddling,” there are no more limits to data extraction, every single datum might prove useful for machine learning algorithms to find new patterns, to produce new “knowledge” (and to use it for competitive advantage and future profits).

Computer scientists might have succumbed to the naturalization of “data relations”; they might think that data just happens to be here, available to them, but what made it so are two centuries of the development of data extraction apparatuses, starting with the modern State and ending in contemporary platform capitalism. But let us not get it the other way around: first the State and then capitalism developed this apparatus because of their own reasons; the technology of machine learning (and deep learning) did not invent it, it merely fell back on this apparatus as its contingent condition and then started to merge with it and enhance it, to fuel and transform it to the point of almost completely integrating data extraction with the flow of life itself. The political (State control, political manipulation, etc.), legal (the issue of unknown consent, copyright issues, etc.), and ethical (the automatic incorporation of racial and other social prejudices, etc.) problematic effects are well known, but what we have to emphasize here is the intertwinement of the capitalist datafication and technology of deep learning, where the latter at a certain point accepted the former as its own contingent condition for development.

Strongly connected to this is another condition of deep learning mechanisms. We have already mentioned an important difference within AI: between deep learning with its neural nets or connectionism, as it was historically called, and symbolic or conventional AI or even Good Old-Fashioned AI (GOFAI), as it is

[22] Couldy and Mejias, 125.

sometimes called. Some computer engineers today consider this difference to be a technical difference of approach: between bottom-up and top-down. Connectionism starts building intelligence from lower levels of cognition (such as perception and learning), hoping that the highest level (logical reasoning) will emerge from it, while symbolic AI starts from the highest level of cognition in the conviction that the lower levels are inessential and can be added or integrated later. This difference might be only technical if we aim to build a hybrid system (which is now, for some, the future of AI), but epistemologically this difference is not a mere difference, it is a schism (as we have seen with Hinton's vision of the future of AI). This is the schism of different epistemologies or two different logics. As Matteo Pasquinelli succinctly explains, symbolic AI "professes that intelligence is a representation of the world (knowing-that) which can be formalized into propositions and, therefore, mechanized following deductive logic," connectionism, on the other hand, "argues that intelligence is experience of the world (knowing-how) which can be implemented into approximate models constructed according to inductive logic."[23]

The paradigmatic case for symbolic AI is the game of chess.[24] It is a perfect game for symbolic AI due to its closed environment, clear goal, and strictly defined rules. It is perfect for rule-based symbol manipulation. The only real issue here is how to find the best way towards the end goal using predetermined rules and knowledge. This is what is basically called problem-solving through search mechanisms: "Choosing which sequence of actions to adopt was a matter of search."[25] Of course, there are certain problems such as combinatorial explosion, but it could be and in fact is side-stepped by not searching for the best sequence of action or the direct final solution, but rather for a "good enough" choice.[26] The computer uses its predetermined knowledge to scan and evaluate the current state of affairs (the position of the chess figures on the board in our

23 See Matteo Pasquinelli, *The Eye of the Master: A Social History of Artificial Intelligence* (London: Verso, 2023), 21–22.

24 See Franchi and Güzeldere, "Machinations of the Mind," 46–56.

25 Agre, "Soul Gained and Lost," 161.

26 "The [search] tree can be created one step (one node) at the time by the recursive application of the rules of the game. [. . .] The complete tree is of course impossible to create, but the search procedure does not have to rely on a complete tree if it settles for less-than-optimal results. It may create just a few positions per turn, deciding to explore only one or two moves beyond the current one." Franchi and Güzeldere, "Machinations of the Mind," 52.

case) and calculates the best possible or the "good enough" move by exploring different possible outcomes considering "only" a couple moves in advance. "Only" in the sense that even in a game such as chess, with its strictly determined and rather small environment, a computer cannot calculate all the possible variations because there are too many of them;[27] but this "only" is still good enough to beat any human.

There is nothing left to chance in symbolic AI, everything happens according to rules and deductive reasoning—the reasoning that does not take any risk.[28] "Chess is the intellectual game par excellence [. . .]; without a chance device to obscure the contest, it pits two intellects against each other in a situation so complex that neither can hope to understand it completely, but sufficiently amenable to analysis that each can hope to outthink his opponent [. . .]," as Allen Newell and Herbert Simon, who advanced and determined the development of (symbolic) AI for decades, somewhat poetically put it.[29] Without a chance device, they say. And for a good reason. Every search mechanism in symbolic AI has to somehow tame the complexity of the game. Any unforeseen event that was not included in the *a priori* rules and knowledge would "obscure" the game, effectively ending it, exploding the machine. And not only adding something new and leading it towards combinatorial explosion, even complicating the existing order a little bit by infusing ambiguity in it, would ruin it because

27 As Claude Shannon, who was the first to focus on the problem-solving aspect of a chess game, explains it with basic math: "In a typical chess position there will be about 32 possible moves with 32 possible replies—already this creates 1024 possibilities. Most chess games last 40 moves or more for each side. So the total number of possible variations in an average game is about 10^{120}. A machine calculating one variation each millionth of a second would require over 10^{95} years to decide on its first move!" Claude Shannon, "A Chess-Playing Machine," *Scientific American* 182 (February 1950): 48–51, quoted in Franchi and Güzeldere, 51.

28 As Ian Hacking defines inductive logic in difference to deductive logic: "Valid arguments are risk-free. Inductive logic studies risky arguments. A risky argument can be a very good one, and yet its conclusion can be false, even when the premises are true." Ian Hacking, *An Introduction to Probability and Inductive Logic* (Cambridge: Cambridge University Press, 2001), 11.

29 Allen Newell, J. C. Shaw, and Herbert Simon, "Chess-Playing Programs and the Problem of Complexity," *IBM Journal of Research and Development* 2 (October 1958): 320–35, quoted in Franchi and Güzeldere, "Machinations of the Mind," 54.

it requires a "deterministic environment" and an unambiguously defined "state space," as computer scientists would say.[30]

Deep learning, with its neural nets, functions in a completely different way. Deep learning has no problem with chances and taking risks, because it deals with probabilities and uses inductive logic. A deep learning mechanism does not search for the best or good enough choice on the basis of its predetermined knowledge; it learns to predict its supposed outcome; it makes a statistical evaluation of the output. In a sentence, for example, a deep learning mechanism predicts the next word on the basis of what is has learned during its training; it does not determine it on the basis of a predetermined rule of syntax (like noun → verb). That is why training the algorithm is so important for deep learning—as well as all the (various) data that it is trained on. It is trained on such so that it can learn to make out patterns which have different probabilities as outputs. If you show a deep learning mechanism a picture of a cat, it will infer from its previous "experiences" the probability that it is a cat (and the probability that it is a dog, for example); it will take its chances and try to guess, to put it a bit simplistically.

While symbolic AI focused on the game of chess, connectionism focused on visual pattern recognition and speech recognition (and later expanded to the field that was "reserved" for symbolic AI, namely intellectual games such as chess and go). It would be hard not to see that these two "approaches" are in fact in opposition: the one counts on ahistorical logical rules, the other bets on changing them through learning; the one does not take any chances, the other operates through them; the one starts with the core of human rationality, the other with lower-level cognitive mechanisms; the one is fascinated with developed minds, the other with the learning of children, etc. These two approaches technically might come together, but "ideologically" they are miles apart—as far apart as necessity and contingency are. To put it a bit more philosophically, they are supported by two radically different ontologies, one betting on order, the other on chaos. The more the world is complex and unpredictable, the more connectionism thrives and calculates the contingencies into probabilities and the more symbolic AI loses its the necessitarian grounds. Where the one gains a world, the other loses one. It then goes hand in hand that when the world seems

30 See Russell and Norvig, *Artificial Intelligence*, 63, 83.

more complex, more absorbed into contingency, the more worldly connectionism seems. And the world—both culture and nature—in the last couple of decades has definitely seemed increasingly complex and contingent.

We can see this schism between symbolic AI and connectionism even more clearly in the history of AI. The problem of the limits of deep learning, this present battle internal to the field of AI, does not concern only its future, but it is also affecting its past. Its past is constantly being rewritten, first by the then victorious symbolic AI and then by the now victorious connectionism.

The field of AI was, as is now widely agreed, established with the Dartmouth Conference (The Dartmouth Summer Research Project on Artificial Intelligence) in 1956. Its initiators were John McCarthy and Claude Shannon. But the main actor was in fact John McCarthy, who was a bit disappointed with their shared project of editing *Automata Studies* and for that reason advanced the idea of a conference.[31] This is an important fact because Shannon was already by then a hugely influential intellectual figure (as "the father" of information theory) and was an active member of cybernetic conferences, which ended a few years before, and was also in touch with Alan Turing and the British Ratio Club. We could perhaps say that he was the connecting point for all major intellectual figures, groups, or movements that were dealing with machine intelligence. On the other hand, the young John McCarthy was a lot more hostile towards cyberneticians and a lot more fascinated by the work of Allen Newell and Herbert Simon. This fact was reflected in the guests and consequently the topics of the conference. And in the name of the field that is now known as artificial intelligence, which he coined: "As for myself, one of the reasons for inventing the term 'artificial intelligence' was to escape association with 'cybernetics.' Its association with analogue feedback seemed misguided, and I wished to avoid having either to accept [. . .] Wiener as a guru or having to argue with him."[32]

[31] See Ronald R. Kline, "Cybernetics, Automata Studies, and the Dartmouth Conference on Artificial Intelligence," *IEEE Annals of the History of Computing* 33, no. 4 (October–December 2011): 5–16, https://doi.org/10.1109/mahc.2010.44.

[32] John McCarthy, review of *The Question of Artificial Intelligence: Philosophical and Sociological Perspectives*, ed. Brian P. Bloomfield, *Annals of the History of Computing* 10, no. 3 (July–September 1988): 227, quoted in Kline, 13.

If we could say that Shannon was the connecting point, then McCarthy was most certainly the disconnecting point. He did not have an issue only with Norbert Wiener and his notoriety, but what was really bugging him was the basic idea of cyberneticians, namely the circular causality of the feedback mechanism (or adaptive self-regulation) and their obsession with electronic circuitry and mechanization—he believed that this was a dead end.[33] He dismissed the idea of brain modelling; he did not consider that neural nets as electronic circuits and the idea that they could have the ability to learn and adapt was the future of AI—or that such even could be a part of it. His sole focus was on symbolic AI, and his main partners were Newell and Simon. So, when he proposed a new name, i.e. "artificial intelligence," he effectively tried to banish cybernetics from the field of studying and building machine intelligence. The story repeated itself in the late 60s when Marvin Minsky (one of the guests at the Dartmouth Conference) and Seymour Papert attacked the work of Frank Rosenblatt, the main figure of connectionism (as the successor of cybernetics), who built the first neural network mechanism called Perceptron, and with this (and also due to other factors) buried connectionism until its small revival in the 80s, with its full resurrection and eventual success occurring in the last decade.

And today, after this "final" success of deep learning methods and the connectionist side, researchers are trying to figure out who the "father" of deep learning is. The answer is usually Frank Rosenblatt with his Perceptron.[34] This was the first neural network mechanism capable of learning (by adjusting the weights of connection between neural nodes), but it had only one layer of neurons between the input and the output layers, so its learning was not deep in this sense. This was also the issue that Minsky and Papert brought up as its impossible limit, but this is precisely the issue that was resolved in the 80s and in the 90s with the backpropagation algorithm and convolutional neural nets, etc. Conceptually, connectionism was not wrong; what seemed impossible were in fact just "technicalities," which were solved by technological development (as we have seen). That is why today researchers are "excavating" Frank Rosenblatt and naming him the father of deep learning. Some go even further and say that the "emergence of neural networks" is "a key idea for AI" because the first to propose the idea that biological neurons could function the same way as elec-

---

33 Kline, 7–8.

34 See Tappert, "Who is the Father of Deep Learning?"

tronic circuits (this means the Boolean logic of AND, OR, and NOT) were Warren McCulloch and Walter Pitts.[35] So McCulloch and Pitts, the two main figures of the cybernetic Macy conferences, should then be seen as the beginning of AI—even the standard textbook on AI confirms this.[36]

We have no intention to raise issue with this, but we would like to emphasize how the history of AI changed: first, the field of AI was established by denegating cybernetics; with the development of symbolic AI came the assertion of its dominance and further denegation, this time of connectionism as the successor to cybernetics; and then after deep learning won the engineering battle, Frank Rosenblatt became one of the key figures of AI and cybernetics became its origin. A total reversal. We could say that there are always various opinions as to what the historical influences and predecessors are, and this is of course true; we could even say that there are always different (maybe even conflicting) interpretations within a theoretical field, but this story of the origins and conditions of AI shows a far deeper confusion: its own formal structures are shaken up; the conceptual field itself is being torn apart.

We are, of course, in no position to judge and predict the future of AI; that is not our aim. We are only attempting to show the conceptual confusion that is currently determining the field of AI and to outline the con-fusion of the various conditions and limits of its most developed and hyped area—deep learning. Its mechanisms were constructed from various "parts"; some were foreign in their nature and domesticated for a heavy price (capitalistic datafication and the dreams of AGI), some pertain to its necessary but forgotten physical substrate, which now calls for our attention (the material and environmental limits), and some reveal the contingency of its worldly context (the world of necessity versus the world of contingency). This is the artificiality of today's AI, with deep learning in its foreground. With it coming to the foreground, it eclipsed its conditions and limitations and succumbed to the capitalist hype and convinced

35 Pasquinelli, *Eye of the Master*, 128. Let us, just for the sake of intrigue, mention that in opposition to this discrete logic of neuronal functioning the renowned Gestalt psychologist Wolfgang Köhler proposed a continuous logic. See Steve J. Heims, *The Cybernetics Group 1946–1953: Constructing a Social Science for Postwar America* (Boston: MIT Press, 1991), 239.

36 "The first work that is now generally considered as AI was done by Warren McCulloch and Walter Pitts (1943)." Russel and Norvig, *Artificial Intelligence*, 35.

itself that it is the sole future of AI as the best potential candidate for AGI. Even worse, it became intertwined with it; even if capitalism did not invent it, it did develop it with all the possible data and infused it with the desire for AGI. But on the other hand, it has pushed it to its materialistic limits and has shown its contingent worldly context, which might prove to be in fact guidelines for the development of the AI field itself. In any case, one thing is certain: it will have to get rid of the desire for AGI, which is fuelling the combativeness of the field and attaching machine intelligence to human intelligence and establishing a relation of competition so that there can emerge something like a paradoxically submissive superior entity (a tool more capable of knowing what to do than its pretended user).

We need to lose this desire for AGI and get rid of the transhumanist fantasy of technological singularity.[37] Not in order to denounce machine intelligence, as humanists have been doing for decades, but in order to affirm it as a specific intelligence. They would, of course, disagree and say that it is not intelligent, that this intelligence is merely a statistical output from a human-trained mechanism. Well, okay, but from such a reductionist point of view even logic is not logic but just some rules. And it is not just some semi-random output, but a precise result of inference, because it uses logic, inductive logic, the same that scientists have been using for the last two centuries. And which human is not himself "human-trained"? Then some would say that a machine does not understand the meaning of its output, that it is not sensible. Of course, it is not, but even a human can lose his own purpose and become lost when the whole of humanity is showing signs of senselessness in today's posthumanistic and anthropocenic world. And then some would say that machine intelligence is not embodied and alive. Yes, it might be seen that way, but in fact it is embodied and alive, and its limits are pressing us to see it as such; and if being alive means gaining experiences and self-adapting through them, then this is precisely what it has begun to do.[38]

37 Regarding transhumanist fantasy and its relation to contemporary capitalism, see Aleš Mendiževec, "No future, končno! Plastičnost med destrukcijo, ustvarjalnostjo in simulacijo" [No Future, Finally! Plasticity between Destruction, Creation and Simulation], afterword to *Ontologija naključja: Esej o destrukcijski plastičnosti*, by Catherine Malabou (Ljubljana: Maska, 2023), 87–116.

38 We refer here mainly to artificial intelligence as connectionism and we in fact posit it as the dominant technique of AI, but we refuse to buy into the schism that does away with sym-

These arguments will convince no one, of that I am sure. And in fact, I agree: machine intelligence does not have its own (general) purpose in the world, because it has no inner experience to relate to and cannot have and express emotions (even if it can imitate them); and even if a human uses scientific methods of statistical inference, it is itself not a statistical tool, the human brain is not merely a one-dimensional scientist, it has multidimensional common sense (as they call it); yes, machine intelligence is data greedy and human intelligence can make a whole lot from just a small amount of data, as they say, because the latter excels at knowledge transfer, while the former lacks this capability, etc. As we have said, humanist critiques are valid, but with this comes a certain *méconnaissance*, as the French would put it. Humanists exclusively allocate intelligence to humans instead of inclusively dispersing it, as an antihumanist might try to do; because they see in machine intelligence the danger (and others the opportunity) of AGI and this is why they transform the machine-human relation from a comparative one (with inclusive differences) into a competitive one (with excluding differences). The goal of our arguments is not to convince anyone by providing proof of machine intelligence, but to de-competify this difference by showing the differentiational proximity of human and machine intelligence in order to open the possibility of considering machine intelligence as intelligent in its own way. If only we get rid of AGI.

References

Agre, Philip. "The Soul Gained and Lost: Artificial Intelligence as a Philosophical Project." In *Mechanical Bodies, Computational Minds: Artificial Intelligence from Automata to Cyborgs*, edited by Stefano Franchi and Güven Güzeldere, 153–73. Cambridge: MIT Press, 2005.

Couldy, Nick, and Ulises Mejias. *The Costs of Connection: How Data is Colonising the Human Life and Appropriating it for Capitalism*. Stanford: Stanford University Press, 2019.

Dupuy, Jean-Pierre. *On the Origins of Cognitive Science: The Mechanization of the Mind*. Translated by D. B. Debevoise. Cambridge: MIT Press, 2009.

Franchi, Stefano and Güzeldere, Güven. "Machinations of the Mind: Cybernetics and Artificial Intelligence from Automata to Cyborgs." In *Mechanical Bodies, Computational*

bolic AI and other segments of AI precisely because we reject the desire for AGI. Hybridity just might lead to the specificity of machine intelligence which would be disentangled from capitalistic desires.

Minds: Artificial Intelligence from Automata to Cyborgs, edited by Stefano Franchi and Güven Güzeldere, 15–149. Cambridge: MIT Press, 2005.

Grosz, Barbara. Interview by Martin Ford. In *Architects of Intelligence: The Truth about AI from the People Building It*, edited by Martin Ford, 333–54. Birmingham: Packt, 2018.

Hacking, Ian. *An Introduction to Probability and Inductive Logic*. Cambridge: Cambridge University Press, 2001.

——. *The Taming of Chance*. Cambridge: Cambridge University Press, 1990.

Heims, Steve J. *The Cybernetics Group 1946–1953: Constructing a Social Science for Postwar America*. Boston: MIT Press, 1991.

Hinton, Geoffrey. Interview by Martin Ford. In *Architects of Intelligence: The Truth about AI from the People Building It*, edited by Martin Ford, 71–94. Birmingham: Packt, 2018.

Kline, Ronald R. "Cybernetics, Automata Studies, and the Dartmouth Conference on Artificial Intelligence." *IEEE Annals of the History of Computing* 33, no. 4 (April 2011): 5–16. https://doi.org/10.1109/mahc.2010.44.

Kurzweil, Ray. Interview by Martin Ford. In *Architects of Intelligence: The Truth about AI from the People Building It*, edited by Martin Ford, 227–50. Birmingham: Packt, 2018.

LeCun, Yann. Interview by Martin Ford. In *Architects of Intelligence: The Truth about AI from the People Building It*, edited by Martin Ford, 119–42. Birmingham: Packt, 2018.

Leiserson, Charles, Neil Thompson, Joel Emer, Bradley Kuszmaul, Butler Lampson, Daniel Sanchez, Tao Schardl. "There's Plenty of Room at the Top: What Will Drive Computer Performance After Moore's Law?" *Science* 368, no. 6495 (June 2020). https://doi.org/10.1126/science.aam9744.

Marcus, Gary. "Deep Learning: A Critical Appraisal." *Arxiv* (January 2018). https://doi.org/10.48550/arXiv.1801.00631.

McCarthy, John. Review of *The Question of Artificial Intelligence: Philosophical and Sociological Perspectives*, edited by Brian P. Bloomfield. *Annals of the History of Computing* 10, no. 3 (July–September 1988)

Mendiževec, Aleš. "No future, končno! Plastičnost med destrukcijo, ustvarjalnostjo in simulacijo." Afterword to *Ontologija naključja: Esej o destrukcijski plastičnosti*, by Catherine Malabou, translated by Iztok Ilc, 87–116. Ljubljana: Maska, 2023.

Newell, Allen, J. C. Shaw, and Herbert Simon. "Chess-Playing Programs and the Problem of Complexity." *IBM Journal of Research and Development* 2 (October 1958): 320–35.

Pasquinelli, Matteo. *The Eye of the Master: A Social History of Artificial Intelligence*. London: Verso, 2023.

Russell, Stuart. Interview by Martin Ford. In *Architects of Intelligence: The Truth about AI from the People Building It*, edited by Martin Ford, 39–69. Birmingham: Packt, 2018.

——, and Peter Norvig. *Artificial Intelligence: A Modern Approach*. Fourth Edition, Global Edition. Harlow: Pearson Education, 2022.

Shannon, Claude. "A Chess-Playing Machine." *Scientific American* 182 (February 1950): 48–51.

Strickland, Eliza. “15 Graphs That Explain the State of AI in 2024.” *IEEE Spectrum*, April 15, 2024. https://spectrum.ieee.org/ai-index-2024.

Tappert, Charles C. “Who Is the Father of Deep Learning?” *International Conference on Computational Science and Computational Intelligence (CSCI)* (2019): 343–48. https://doi.org/10.1109/CSCI49370.2019.00067.

Thompson, Neil C., Kristjan Greenewald, Keeheon Lee, Gabriel F. Manso. “Deep Learning’s Diminishing Returns.” *IEEE Spectrum*, September 24, 2021. https://spectrum.ieee.org/deep-learning-computational-cost.

———. “The Computational Limits of Deep Learning.” *Arxiv* (July 2022). https://doi.org/10.48550/arXiv.2007.05558.

Turing, Alan. “Computer Machinery and Intelligence.” In *The Essential Turing: Seminal Writings in Computing, Logic, Philosophy, Artificial Intelligence, and Artificial Life: Plus The Secrets of Enigma*, edited by Jack Copeland, 433–64. Oxford: Oxford University Press, 2004.