An International Journal in Philosophy, Religion, Politics, and the Arts

Existenz

Volume 17, No 1, Spring 2022 ISSN 1932-1066

# Artificial Intelligence Against the Backdrop of an Adequately Differentiated Anthropology

Albrecht Kiel
*University of Konstanz, Germany*
a.kiel@uni-konstanz.de

**Abstract:** The achievements and deficits of artificial intelligence systems should be seen against the backdrop of a differentiated anthropology that encompasses not only the rational functions of consciousness, but also those of the subrational-unconscious as well as suprarational functions. These include firstly all forms of fantasies and fictions. Furthermore, they include complex mental achievements such as identity resulting from individuation, imagination, intuition, creativity, and the power of judgement that is capable to separate the relevant from the irrelevant. On the level of semiotics, a distinction must be drawn between images and archetypes and between signs with unambiguous and symbols with ambiguous or connotative-hidden meanings respectively. The quality of a word semantics, sentence semantics, and text semantics also depends on this ability.

**Keywords:** Jaspers, Karl; Jung, Carl Gustav; Furbach, Ulrich; Global Workspace Theory; psychological functions; consciousness; subrational pre-conscious; complex psychological functions; semiotics and semantics.

In his contribution to this issue of *Existenz*, Ulrich Furbach mentions the Global Workspace Theory that was conceived by Bernard Baars and Stan Franklin, it is a modelling theory of consciousness and of higher-order cognition, both of which emerge from competitive and integrated information flows that move via widespread neuronal processes.[1] The theory primarily concerns short-term working memory that is considered to be a spotlight when viewed from a perspective of attention. In it, the human mind (*Geist*) is comprehended as a theater in which conscious processes form matching pairs with unconscious processes. J.W. Dalton criticized it as an "unfinished theatre" in which the role of conscious experience was missing.[2] For him, it is thus at best an account of the cognitive function of consciousness, and he also criticized its methods for decoding unconscious content. The Global Workspace Theory is only one neuroscientific theory of consciousness among other interesting ones. For example, the holonomic brain theory of Karl Pribram and David Bohm, which construes the brain as a holographic memory network in which quantum effects in or between brain cells are occurring; or Antonio Damasio's theory of consciousness, which emphasizes the connection between brain, body, and environment and distinguishes between a proto-self, a core self, and an autobiographical self. Prior

[1] Ulrich Furbach, "On Reasoning, Commonsense Knowledge, and Consciousness," *Existenz* 17/1 (Spring 2022), 1-6. [Henceforth cited as *RCK*]

[2] J. W. Dalton, " The Unfinished Theatre," *Journal of Consciousness Studies* 4/4 (1 April 1997), pp. 319-324.

to these interpretations one adhered to traditional philosophical theories such as functionalism, epiphenomenalism, emergence theory, parallelism, interactionism, and idealism. Traditionally, there are two accounts in which the relation between the brain (body) and psychological functions (mind) is conceived, namely either in a dualistic manner or as an identity of these two functions.

These underlying invisible actors in the Global Workspace Theory that are effective in the background of the blackboard system need therefore be defined more precisely as various psychological functions. Doing this can involve not only unconscious processes, but also areas of the suprarational understood as higher order cognition. Despite all the successes of machine learning, deep learning, and reinforcement learning with neural networks and the Projective Simulation learning model, the question regarding an agent who decides and acts intentionally needs an answer. Likewise, it remains an open question as to whether this agent can put forward purely argumentative reasons or whether this capacity is directly linked to reward systems or mere coincidence. It remains to be seen as to whether a reasoner about reasons posits a self as the basis of that reasoner's free will. In her review of three recent books on the subject matter, Liad Mudrik provides an up-to-date overview of the current state of the debate.[3] According to Mudrik, Daniel Dennet emphasizes the natural selection advantage of consciousness as an organ of control and he would like to dispel the illusion of a core self; rather the self develops out of the experiences of one's life story. In contrast to Dennet, Kevin Mitchell, regards the self as the key to free will, which then can only guarantee top-down control, forward planning, and constancy of the person. However, for Joseph LeDoux consciousness is the highest level of what he considers to be four realms of existence. These readings show the weakness of the debate that is influenced by the analytical theory of mind in the English-speaking world, namely a confusion of basic psychological concepts. Sometimes these self-reflecting capabilities are being referred to as soul or psychological phenomena in general, then again, the label consciousness is being used, and in the same breath the concepts of spirit, mind, or the mental are being used, and subsequently the notion of self is being added.

These questions should be answered against the backdrop of an adequately differentiated anthropology (based on the ones by Karl Jaspers and Carl Gustav Jung, among others),[4] which can differentiate psychological functions in a comprehensive manner on different levels of reality (higher orders) with corresponding levels of communication.

Pre-conscious thinking in images and archetypes of landscapes, animals, people (and their animistic projections) was followed by using language and abstractions of images, such as pictograms and ideography or hieroglyphics. It was only on this basis that the rational object consciousness could form that is linked to the fourth memory bank of objective world knowledge, which is used to argue, calculate, plan, and, organize (reason, intellect, *nous poietikos*).

However, according to the current state of brain research it is interwoven with three other memory systems and two other consciousness systems: After priming and procedural memory (which is significant for robotics) comes autobiographical memory (that is anchored in historically evolved collective traditions), which is linked to the individuation process of self-consciousness. And concurrently, it is associated with the most primal form of consciousness, conscious experience as the origin of dance, song, music, and language; it is closely connected to an increased phenomenological intelligence of perception.

For Jaspers, the dialectic of inner action is not only determined by these three types of consciousness. What needs to be added to this is the dialectic of knowing and not knowing as well as the dialectic of dynamic movement and the attempts to securing it, whereby certainties and forms of fulfilment (faith, fantasy, love, orientation, trust) are being sought.[5] Jaspers is thus the first scholar to concretize the abstractions of a discussion held since antiquity regarding the contrast between outer and inner speech as *dianoia*, *oratio mentalis*, as thinking, soliloquy, *langage intérieur*. Also relevant in this context is Jerry Fodor's language of

[3] Liad Mudrik, "The Grand Challenge of Consciousness," *Nature* 623/7985 (2 November 2023), 25-26.

[4] A detailed account of this topic can be found in Albrecht Kiel, "The Logic of Karl Jaspers as an Intercultural Basic Knowledge," *Existenz* 13/2 (Fall 2018), 8-18; and Albrecht Kiel, "Sixteen Perspectives on Karl Jaspers' Orientative Knowledge," *Existenz* 13/2 (Fall 2018), 19-24.

[5] Karl Jaspers, *Philosophy, Volume 2*, transl. E. B. Ashton, Chicago, IL: University of Chicago Press 1970, pp. 281-6. Karl Jaspers, *Von der Wahrheit*, München, DE: R. Piper & Co. Verlag 1958, pp. 180f., 352, 393f., 540f.

thought as central code, central computing language, and medium for the computations' underlying cognitive processes.[6]

In this dialectic, the complex mental functions come into play: dream fantasies and fantasy on various levels as waking dream, imagination, fiction, and hypothesis; creativity (novel synthetic re-combinations that have new functions) and intuition; power of judgment regarding the "relevance" of "background knowledge" that Furbach mentions at the end of his text (*RCK* xyz); ethos and ethics; and specifically: the themes of scientific ethos; the nexus of individuation with identity. Suprarational, communicative examples include: Indirect communication; metaphor and allegory; pathos and enthusiasm versus sober clarity; sarcasm, parody, and satire; humor and psychological resilience.

Only once such differentiations are being made, could the boundaries of the industrial revolution determined by artificial intelligence be defined more precisely. Via this clarification one could also counter objectively utopian hopes and dystopian fears (for instance, human-AI symbioses). In the subdivisions of artificial intelligence there are different expansion possibilities and limitations, for instance, supervised, reinforcing, or stochastic learning through artificial neural networks; knowledge discovery and data mining (resulting in an expanded knowledge space that can only be achieved with complex computing systems, especially in systems biology); pattern recognition (for instance in exemplars, language, and in handwriting); knowledge engineering and knowledge modelling; expert systems that differ with regard to whether the recommendations for action are derived from a knowledge base that relates to legal, medical, or psychiatric problems (such as suicide risk); general game playing; computer vision and robotics.

Artificial intelligence regarded in terms of an industrial revolution creates competitive advantages if communication and ideas can be formulated in a manner that is accessible to computational language. Building on this, the challenge consists then in advancing into the spheres of intellectually demanding creativity. Using artificial intelligence systems in this way, risk decisions can be prepared better, ultimately, however, they can only be made based on value judgments (executed by humans).

For artificial intelligence research, it would be conducive if a semiotic distinction were at least made between images and unconsciously effective archetypes. Furthermore, signs with an unambiguous meaning ought to be distinguished from symbols with an ambiguous, connotative-hidden meaning (*RCK* xyz7). In this context reference should be made to Paul Ricœur's dialectic of three symbol levels, yet above all to C. G. Jung's research into symbols and archetypes. According to them, symbols have an unconsciously effective sense and background for a cultural community. The quality of semantics also depends on the differentiation of semiotics. The semiotic triangle (of sign system, concept, and object) is the subject of the cited pages of Jaspers' *Von der Wahrheit*. Only it allows for the emergence of the various linguistic types of word, sentence, and text semantics. In this context it needs to be clarified whether they capture the unconscious-preconscious, the rationally conscious, the aforementioned complex-suprarational (spiritual) as well as the existential functions of the self or self-consciousness respectively.

As Aeneas Rooch elaborates in his book *Die Entdeckung der Unendlichkeit*, the development of mathematics from Georg Cantor to David Hilbert and from Kurt Gödel to Paul Cohen also evidences that psychological functions belong to the objective truths that cannot be completely discerned in the formal framework of mathematics. I agree with Rooch that philosophical anthropology needs to ask whether and how complicated intellectual functions such as inventiveness, intuition, or reason (in contradistinction to rationality) can be represented by a sum total of simple functions and which functions are suitable for summing up.[7]

Jaspers' further treatment of semiotics of the signa of existential decisions, which determine the direction of individuation, and the personally experienceable ciphers of transcendence (contained in the progression of Jaspers' work spanning from *Philosophy, Volume III* to his lectures given in the summer semester of 1961) most

[6] Stephan Meier-Oeser, "Inneres Wort; Innere Rede," in *Historisches Wörterbuch der Philosophie, Band 12, W-Z*, eds. Joachim Ritter, Karlfried Gründer, Gottfried Gabriel, Basel, CH: Schwabe Verlag 1971, see the entry "Wort, Inneres; Rede, Innere" that also contains further references.

[7] Aeneas Rooch, *Die Entdeckung der Unendlichkeit: Das Jahrhundert, in dem die Mathematik sich neu erfand*, 1870–1970, München, DE: Wilhelm Heyne Verlag 2022, pp. 51, 340, 355, 379, 384.

likely would be difficult to grasp by artificial intelligence systems at current levels of development; if anything, his dialectic of the types of the encompassing (of objects and horizons) on the various levels of reality might be grasped by them.

And perhaps the connecting functions of individual-basal reason (*Verstand*) as an approximate completeness of psychological functions, and higher reason (*Vernunft*)—in respect to the entire immanent and transcendent reality (*Wirklichkeit*) and one's own shadow of dysfunctions—could also be represented in a universally valid manner. For instance, this could be done with the aim of transcending (sublating) the incompatible theistic notions of God in a panentheistic synopsis. After the eclipsing of ideologies, theocracies in overpopulated regions seem to be the main cause of war today—currently this goes as far as once again there being a danger of nuclear war.[8]

[8] This contribution, originally submitted in German language, has been translated by the *Existenz* editors.