

***Искусственный интеллект, когнитивная наука,
теория систем, философия
Artificial Intelligence, Cognitive Science,
systems theory, philosophy
人工智能、認知科學、系統論、哲學***

Лекторский В.А.

академик РАН, Государственный академический университет
гуманитарных наук, Москва, Россия,
Academician of the Russian Academy of Sciences, State Academic University
Humanities, Moscow, Russia,
俄羅斯科學院院士、國立學術大學
人文學科，莫斯科，俄羅斯，
v.a.lektorski@gmail.com

Аннотация. Обсуждается глубинная связь на первый взгляд достаточно разрозненных направлений, указанных в названии. Акцентируется внимание на обострившейся в настоящее время необходимости понимания места человека в современном мире с его вызовами, в особенности, связанными с проблематикой искусственного интеллекта, необходимости современных ответов на вечные вопросы И.Канта о человеке: «Что я могу знать?», «Что я должен делать?», «На что я могу надеяться?»

Ключевые слова: философия, искусственный интеллект, когнитивные науки, теория систем, нейронные сети, функциональные системы, тектология, познание, сознание

Annotation. The deep connection between the seemingly disparate directions indicated in the title is discussed. Attention is focused on the currently acute need to understand the place of man in the modern world with its challenges, especially those related to the problems of artificial intelligence, the need for modern answers to I. Kant's eternal questions about man: "What can I know?", "What am I?" should I do?", "What can I hope for?"

Keywords: philosophy, artificial intelligence, cognitive sciences, systems theory, neural networks, functional systems, tectology, cognition, consciousness

註解。討論了標題中所示的看似不同的方向之間的深層聯繫。人們的注意力集中在當前迫切需要了解人類在現代世界中的地位及其挑戰，特別是與人工智能問題有關的問題，以及對康德關於人類的永恆問題的現代答案的需要：“我能知道什麼”？、“我是什麼？我應該做什麼？”、“我能期待什麼？”

關鍵詞：哲學、人工智能、認知科學、系統論、神經網絡、功能係統、構造學、認知、意識

Разрешите мне приветствовать участников Конгресса от имени Научного Совета РАН по методологическим проблемам искусственного интеллекта и

когнитивных исследований, председателем которого я являюсь. Совет существует около двух десятилетий. Последние годы мы работаем при Президиуме Российской Академии наук. Наш Совет – один из организаторов Конгресса.

В названии Конгресса объединены проблемы, которые кажутся весьма разными: это и биология, и теория систем, и искусственный интеллект, и математика. В действительности между этими сюжетами существует глубинная связь. При этом я добавлю к обсуждению ещё одну область знания, которая, как я считаю, имеет прямое отношение к обсуждаемой нами тематике: это философия.

Примерно 60 лет тому назад появился термин «искусственный интеллект» (artificial intelligence). На конференции в США ряд специалистов пришел к выводу о том, что можно создавать такого рода системы, которые способны имитировать и усиливать интеллектуальные возможности человека. Первоначально речь шла о логических интеллектуальных системах, потом появились системы, использующие искусственные нейронные сети и возможности глубокого обучения. За 60 лет в развитии систем искусственного интеллекта достигнуты огромные успехи, без них была бы невозможна глобальная цифровизация, которая сегодня во многом определяет нашу жизнь, начиная с экономики и включая оборонную сферу, образование и культуру.

А в начале 70-ых гг. прошлого столетия произошло другое важное событие: возникло новое междисциплинарное движение, которому дали имя когнитивной науки. В неё вошли разработки в области искусственного интеллекта, а также такие дисциплины, как психология, лингвистика, исследования мозга (нейронаука). Вошла в него и философия, прежде всего в той её части, которая изучает познание и сознание (эпистемология и философия сознания). Представители этих разных дисциплин, включившихся в когнитивную науку (когнитивная психология, когнитивная лингвистика, когнитивная нейронаука), разделяли две идеи: во-первых, что когнитивные (познавательные) процессы – не что иное, как переработка информации, получаемой из мира соответствующим устройством (для человека и других живых существ - это мозг), во-вторых, что носителями информации у когнитивных агентов являются особые структуры – ментальные репрезентации. При этом считалось, что ключом к пониманию того, как происходит переработка информации когнитивным агентом, являются модели, предлагаемые разработчиками искусственного интеллекта. Работа головного мозга была понята как работа особого рода компьютера - компьютерная метафора. Философы оказались главными теоретиками когнитивной науки, сформулировав в этой связи особые теории: когнитивного символизма, функционализма, «языка мысли» и др.

За 50 лет развития когнитивной науки было сделано немало. Во всяком

случае наладилось плодотворное взаимодействие разных дисциплин, изучающих человека и его познавательные процессы – раньше ничего подобного не было. Вместе с тем сама когнитивная наука в целом претерпела большие изменения. На смену пониманию познания в духе когнитивного символизма и функционализма пришли иные представления: коннекционизм, идея, так называемого, «отелесненного ума», концепция об укоренённости познания в биологических процессах жизнедеятельности, необходимость учёта роли культурных воплощений в познавательной деятельности, понимание человеческого мозга как не просто естественного образования, а как особого биоартефакта, нередуцируемость сознания к когнитивным процессам и др. Познание оказалось более сложным, чем это думалось 60 лет тому назад. Возможности и границы использования моделей искусственного интеллекта для его понимания — это сегодня предмет острых дискуссий.

Ныне ряд теоретиков когнитивной науки связывают её будущее с толкованием когнитивных процессов на основе теории предвосхищающей (предсказывающей) обработки информации – *predictive processing*. Предполагается, что когнитивный агент извлекает информацию из мира, руководствуясь определённой моделью возможных будущих состояний окружения как результата его действий и корректируя эту модель в свете получаемого опыта. Для понимания процессов переработки информации используются некоторые идеи из области искусственного интеллекта. Но ведь эта концепция не что иное, как вариант теории функциональных систем нашего выдающегося учёного Петра Кузьмича Анохина: модель потребного будущего, акцептор результата действия и др.! Так, теория систем оказывается связанной с современными тенденциями в когнитивной науке и с разработками в сфере искусственного интеллекта.

Нужно сказать, что проблематика теории систем и системного анализа опирается на хорошую отечественную традицию. Ещё в начале 20-ых годов прошлого столетия наш философ и учёный А.А. Богданов опубликовал труд под названием «Тектология (всеобщая организационная наука)». В этом труде были развиты определённые системные представления, опередившие своё время. Идеи Богданова не были сразу приняты, а повсеместно оценены гораздо позже, спустя примерно 40 лет, когда под влиянием Л.Берталанфи начали разрабатывать «Общую теорию систем». Эти разработки начались и у нас усилиями философов, логиков, биологов, экономистов. В 80-ые и 90-ые годы в мире и в нашей стране началось исследование сложно-организованных и саморазвивающихся систем, которое продолжается и сейчас. При этом сегодня системные разработки всё более тесно связываются с когнитивными исследованиями. А последние во многом определяют развитие научного знания в целом.

Я хочу затронуть ещё одну проблему, которая в наши дни вызывает

острые дебаты среди специалистов. Это проблема сознания.

Всегда считалось, что именно разум является главным преимуществом человека по отношению ко всем другим живым существам. Недаром наш биологический вид называется *Homo Sapiens* – человек разумный. При этом предполагалось, что познание сопровождается сознанием: нельзя мыслить, не сознавая процесс мышления. Поэтому когда в середине 19 века великий немецкий учёный (физик, физиолог, психолог, философ) Г. Гельмгольц предложил научной общественности понимание восприятия как бессознательного умозаключения, он вызвал волну возмущения: это невозможная идея, заявляли его оппоненты - умозаключение, как и любой другой вид разумной деятельности не может осуществляться без участия сознания! Между тем, сегодня все специалисты в области когнитивных наук принимают положение о существовании «когнитивного бессознательного». Оказывается, что большая часть той огромной когнитивной работы, которую осуществляет мозг (анализ, синтез, обобщение, выдвижение гипотез, разные виды умозаключений и т.д.) происходит без участия сознания. В этой связи возникает вопрос о том, что в таком случае есть сознание и какова его роль в эволюции живого мира и в жизнедеятельности когнитивных агентов.

В наши дни большинство специалистов в области когнитивных наук разделяют мнение о том, что сознание не сводится к тем или иным когнитивным функциям или даже к совокупности таковых. Система искусственного интеллекта может обладать множеством таких функций, но вряд ли ей можно приписать наличие сознания. Что же такое сознание? Есть мнение о том, что это некий самоотчёт когнитивной системы о собственных состояниях. Но разве такой самоотчёт невозможен без сопровождения его сознанием (когда одни уровни мозговой деятельности репрезентируют и контролируют работу других)? Есть иное мнение: сознание есть особого рода переживание, некое феноменальное поле, сопровождающее когнитивную деятельность. Однако в любом случае возникает вопрос: если столь многие познавательные процессы успешно проходят без участия сознания, зачем, вообще, нужно сознание, какова его роль (один философ написал статью под названием: «Почему не все когнитивные процессы идут в темноте?»)? И возможно ли изучать сознание объективным способом, исследуя внешнее поведение когнитивного агента и работу его мозга? Существуют ли, вообще, объективные маркеры сознания или же всегда нужно обращаться к испытуемому и спрашивать его о субъективных переживаниях? Но как тогда быть с изучением сознания животных? Ведь их невозможно спросить об их переживаниях! Если предполагать, что сознание есть не только у человека, но и у других живых существ, то как по внешним признакам узнать о его наличии? Мы обычно считаем, что сознание есть у собак, кошек, лошадей, человекообразных обезьян, других млекопитающих, потому что они в чём-то

сходны с нами. А вот есть ли оно у таракана, червяка? Это непростые вопросы, которые горячо обсуждаются в когнитивных науках. Без ответа на них невозможно думать о проектировании искусственного сознания - а такие попытки сегодня есть.

Использование систем искусственного интеллекта в жизни и в научном познании, а также интенсивное развитие когнитивных наук создают ситуацию антропологической революции и антропологического вызова.

Машинная обработка больших данных порождает такие предсказания, которые не может делать человек. Но обычно предсказание (на основе простых наблюдений или же с помощью научной теории) предполагает возможность его объяснения и понимания. Машинное предсказание не сопровождается объяснением и пониманием того, как оно было получено (человек этого не понимает, машина тоже, но последней оно и не нужно). Мы имеем в этом случае дело с чуждым нам разумом. Это рациональность, но не та, которая была до сих пор привычной и делала нас «Человеком разумным». К тому же машина может в некоторых случаях выдавать неверные прогнозы - слепое доверие к её рекомендациям рискованно. Использование машинной обработки больших данных в научных исследованиях привело некоторых специалистов к идее о том, что научная теория (на основе которой можно понять и объяснить изучаемые явления) больше не нужна, как не нужны также научные гипотезы и эксперименты. Если всерьёз отнестись к этой идее, то привычная нам наука не имеет будущего.

Использование искусственного интеллекта для управления техническими системами (в частности, в беспилотном транспорте), а также в праве (т.н. «машиночитаемое право»), в других областях жизни остро ставят ряд этических вопросов: сохранение за человеком в этих новых условиях роли автономного морального агента, отвечающего за последствия своих действий, создание «этических» интеллектуальных систем, не способных навредить человеку, критерии доверия к рекомендациям искусственного интеллекта. А это наряду с вопросами о познании, рациональности, сознании, структуре научной теории - фундаментальные философские проблемы, которые имеют не только огромный исследовательский смысл, но сегодня оказываются жгуче практическими.

В своё время великий философ Кант написал о том, что для ответа на вопрос «Что такое человек?» нужно ответить на три других: «Что я могу знать?», «Что я должен делать?», «На что я могу надеяться?». Сегодня мы должны заново отвечать на все эти вопросы. От этого ответа зависит будущее человека.

Allow me to welcome the participants of the Congress on behalf of the Scientific Council of the Russian Academy of Sciences on methodological problems of artificial intelligence and cognitive research, of which I am the chairman. The

council has existed for about two decades. In recent years we have been working under the Presidium of the Russian Academy of Sciences. Our Council is one of the organizers of the Congress.

The name of the Congress combines problems that seem very different: biology, systems theory, artificial intelligence, and mathematics. In fact, there is a deep connection between these stories. At the same time, I will add to the discussion another area of knowledge, which, as I believe, is directly related to the topic we are discussing: philosophy.

About 60 years ago the term “artificial intelligence” appeared. At a conference in the USA, a number of specialists came to the conclusion that it is possible to create systems of this kind that can imitate and enhance human intellectual capabilities. Initially, we were talking about logical intelligent systems, then systems using artificial neural networks and deep learning capabilities appeared. Over the past 60 years, enormous strides have been made in the development of artificial intelligence systems; without them, global digitalization, which today largely determines our lives, starting with the economy and including the defense sector, education and culture, would not be possible.

And in the early 70s. In the last century, another important event occurred: a new interdisciplinary movement arose, which was given the name cognitive science. It includes developments in the field of artificial intelligence, as well as such disciplines as psychology, linguistics, and brain research (neuroscience). Philosophy was also included in it, primarily in that part that studies cognition and consciousness (epistemology and philosophy of consciousness). Representatives of these different disciplines included in cognitive science (cognitive psychology, cognitive linguistics, cognitive neuroscience) shared two ideas: firstly, that cognitive (cognitive) processes are nothing more than the processing of information received from the world by a corresponding device (for of humans and other living beings - this is the brain), secondly, that the carriers of information in cognitive agents are special structures - mental representations. At the same time, it was believed that the key to understanding how information is processed by a cognitive agent is the models proposed by artificial intelligence developers. The work of the brain has been understood as the work of a special kind of computer - a computer metaphor. Philosophers turned out to be the main theorists of cognitive science, formulating in this regard special theories: cognitive symbolism, functionalism, “language of thought”, etc.

Over the 50 years of development of cognitive science, a lot has been done. In any case, fruitful interaction between different disciplines studying man and his cognitive processes has been established - there was nothing like this before. At the same time, cognitive science itself has undergone great changes as a whole. The understanding of cognition in the spirit of cognitive symbolism and functionalism has been replaced by other ideas: connectionism, the idea of the so-called “embodied

mind”, the concept of the rootedness of cognition in the biological processes of life, the need to take into account the role of cultural embodiments in cognitive activity, the understanding of the human brain as not just natural formation, but as a special bio-artifact, the irreducibility of consciousness to cognitive processes, etc. Cognition turned out to be more complex than it was thought 60 years ago. The possibilities and limits of using artificial intelligence models to understand it are the subject of heated debate today.

Nowadays, a number of theorists of cognitive science associate its future with the interpretation of cognitive processes based on the theory of anticipatory (predictive) information processing - predictive processing. It is assumed that a cognitive agent extracts information from the world, guided by a certain model of possible future states of the environment as a result of its actions and adjusting this model in the light of the experience gained. To understand information processing processes, some ideas from the field of artificial intelligence are used. But this concept is nothing more than a version of the theory of functional systems of our outstanding scientist Pyotr Kuzmich Anokhin: a model of the required future, an acceptor of the result of an action, etc.! Thus, systems theory turns out to be connected with modern trends in cognitive science and with developments in the field of artificial intelligence.

It must be said that the problems of systems theory and systems analysis are based on a good domestic tradition. Back in the early 20s of the last century, our philosopher and scientist A.A. Bogdanov published a work entitled “Tektology (general organizational science).” In this work, certain systemic ideas were developed that were ahead of their time. Bogdanov’s ideas were not immediately accepted, but were widely appreciated much later, about 40 years later, when, under the influence of L. Bertalanffy, they began to develop the “General Theory of Systems”. These developments began in our country through the efforts of philosophers, logicians, biologists, and economists. In the 80s and 90s, research into complexly organized and self-developing systems began in the world and in our country, which continues to this day. At the same time, today system developments are increasingly closely associated with cognitive research. And the latter largely determine the development of scientific knowledge as a whole.

I want to touch on another issue that is causing heated debate among specialists these days. This is a problem of consciousness.

It has always been believed that it is the mind that is the main advantage of a person in relation to all other living beings. It is not for nothing that our biological species is called Homo Sapiens - reasonable man. It was assumed that cognition is accompanied by consciousness: one cannot think without being aware of the thinking process. Therefore, when in the mid-19th century the great German scientist (physicist, physiologist, psychologist, philosopher) G. Helmholtz proposed to the scientific community an understanding of perception as an unconscious inference, he

caused a wave of indignation: this is an impossible idea, his opponents said - inference, like any other type intelligent activity cannot be carried out without the participation of consciousness! Meanwhile, today all specialists in the field of cognitive sciences accept the existence of the “cognitive unconscious”. It turns out that most of the enormous cognitive work that the brain carries out (analysis, synthesis, generalization, hypotheses, various types of inferences, etc.) occurs without the participation of consciousness. In this regard, the question arises about what consciousness is in this case and what its role is in the evolution of the living world and in the life of cognitive agents.

Nowadays, most specialists in the field of cognitive science share the view that consciousness is not reducible to certain cognitive functions or even to a set of them. An artificial intelligence system may have many such functions, but it can hardly be attributed to the presence of consciousness. What is consciousness? There is an opinion that this is a kind of self-report of the cognitive system about its own states. But is such self-report impossible without accompanying it with consciousness (when some levels of brain activity represent and control the work of others)? There is another opinion: consciousness is a special kind of experience, a kind of phenomenal field that accompanies cognitive activity. However, in any case, the question arises: if so many cognitive processes successfully take place without the participation of consciousness, why is consciousness needed at all, what is its role (one philosopher wrote an article entitled: “Why don’t all cognitive processes take place in the dark?”)? And is it possible to study consciousness in an objective way by studying the external behavior of a cognitive agent and the functioning of its brain? Are there, in general, objective markers of consciousness, or is it always necessary to turn to the subject and ask him about subjective experiences? But what then to do with the study of animal consciousness? After all, it is impossible to ask them about their experiences! If we assume that not only humans, but also other living beings have consciousness, then how can we recognize its presence by external signs? We usually believe that dogs, cats, horses, apes, and other mammals have consciousness because they are similar to us in some way. But does a cockroach or a worm have it? These are thorny questions that are hotly debated in cognitive science. Without an answer to them, it is impossible to think about designing artificial consciousness - and there are such attempts today.

The use of artificial intelligence systems in life and in scientific knowledge, as well as the intensive development of cognitive sciences, creates a situation of an anthropological revolution and an anthropological challenge.

Machine processing of big data generates predictions that humans cannot make. But usually a prediction (based on simple observations or with the help of a scientific theory) presupposes the possibility of its explanation and understanding. A machine prediction is not accompanied by an explanation and understanding of how it was obtained (a person does not understand this, nor does a machine, but the latter

does not need it). In this case we are dealing with a mind alien to us. This is rationality, but not the one that was hitherto familiar and made us “Homo sapiens.” In addition, the machine can, in some cases, produce incorrect forecasts - blind trust in its recommendations is risky. The use of machine processing of big data in scientific research has led some experts to the idea that scientific theory (on the basis of which the phenomena under study can be understood and explained) is no longer needed, just as scientific hypotheses and experiments are no longer needed. If we take this idea seriously, then the science we are accustomed to has no future.

The use of artificial intelligence to control technical systems (in particular, in unmanned vehicles), as well as in law (the so-called “machine-readable law”), in other areas of life, acutely raises a number of ethical questions: maintaining the role of an autonomous person in these new conditions a moral agent responsible for the consequences of his actions, the creation of “ethical” intelligent systems that cannot harm a person, criteria for trusting the recommendations of artificial intelligence. And this, along with questions about knowledge, rationality, consciousness, and the structure of scientific theory, are fundamental philosophical problems that not only have enormous research meaning, but today turn out to be extremely practical.

At one time, the great philosopher Kant wrote that to answer the question “What is a person?” three others must be answered: “What can I know?”, “What should I do?”, “What can I hope for?” Today we must answer all these questions anew. The future of a person depends on this answer.

請允許我代表俄羅斯科學院科學委員會歡迎與會人員參加關於人工智能和認知研究方法論問題的大會，我是該委員會的主席。該委員會已經存在了大約二十年。近年來我們一直在俄羅斯科學院主席團領導下工作。我們的理事會是大會的組織者之一。

大會的名稱結合了看似截然不同的問題：生物學、系統論、人工智能和數學。事實上，這些故事之間有著深刻的聯繫。同時，我將在討論中添加另一個知識領域，我認為它與我們正在討論的主題直接相關：哲學。

大約 60 年前，“人工智能”一詞出現。在美國的一次會議上，許多專家得出的結論是，創建這種可以模仿和增強人類智力的系統是可能的。最初，我們談論的是邏輯智能系統，然後出現了使用人工神經網絡和深度學習功能的系統。過去 60 年來，人工智能系統的發展取得了巨大進步；沒有它們，全球數字化就不可能實現，全球數字化在很大程度上決定了我們的生活，從經濟開始，包括國防部門、教育和文化。

70 年代初。上個世紀，發生了另一件重要事件：一場新的跨學科運動興起，被命名為認知科學。它包括人工智能領域的發展，以及心理學、語言學和大腦研究（神經科學）等學科。哲學也包含在其中，主要是研究認知和意

識的部分（認識論和意識哲學）。認知科學（認知心理學、認知語言學、認知神經科學）中不同學科的代表們有兩個共同的觀點：首先，認知（認知）過程只不過是通過相應的設備（例如，認知）從世界接收到的信息的處理過程。人類和其他生物——這就是大腦），其次，認知主體中信息的載體是特殊的結構——心理表徵。與此同時，人們認為理解認知代理如何處理信息的關鍵是人工智能開發人員提出的模型。大腦的工作被理解為一種特殊計算機的工作——計算機的隱喻。哲學家成為認知科學的主要理論家，在這方面提出了特殊的理論：認知象徵主義、功能主義、“思想語言”等。認知科學 50 年來的發展已經取得了很多成果。無論如何，研究人類及其認知過程的不同學科之間已經建立了富有成效的互動——以前從未有過這樣的情況。與此同時，認知科學本身整體也發生了巨大的變化。本著認知象徵主義和功能主義精神對認知的理解已被其他思想所取代：聯結主義、所謂“具身心靈”的思想、認知植根於生命生物過程的概念、需要考慮到文化體現在認知活動中的作用、對人腦的理解不僅僅是自然形成，而是一種特殊的生物製品、意識對認知過程的不可還原性等等。比 60 年前想像的還要複雜。使用人工智能模型來理解它的可能性和局限性是當今激烈爭論的主題。

如今，許多認知科學理論家將其未來與基於預期（預測）信息處理（預測處理）理論的認知過程解釋聯繫起來。假設認知主體從世界中提取信息，其行為受到環境未來可能狀態的某種模型的指導，並根據所獲得的經驗調整該模型。為了理解信息處理過程，使用了人工智能領域的一些想法。但這個概念只不過是我們傑出科學家彼得·庫茲米奇·阿諾欣的功能係統理論的一個版本：所需未來的模型、行動結果的接受者等等！因此，系統理論與認知科學的現代趨勢以及人工智能領域的發展息息相關。

必須說，系統論和系統分析的問題是有國內良好傳統的。早在上世紀 20 年代初，我們的哲學家 and 科學家 A.A. 博格丹諾夫發表了一篇題為“Tektology（一般組織科學）”的著作。在這項工作中，提出了某些超前於時代的系統思想。博格丹諾夫的思想並沒有立即被接受，但在很久以後，大約 40 年後，在 L. Bertalanffy 的影響下，他們開始發展“系統通論”，並得到了廣泛的讚賞。這些發展是在哲學家、邏輯學家、生物學家和經濟學家的努力下在我國開始的。上世紀 80 年代和 90 年代，世界和我國開始對複雜組織和自我發展系統的研究，並持續至今。與此同時，當今系統的發展與認知研究的聯繫越來越緊密。而後者在很大程度上決定了整個科學知識的發展。

我想談談最近引起專家激烈爭論的另一個問題。這是一個意識問題。

人們一直認為，與所有其他生物相比，心靈是一個人的主要優勢。我們的生物物種被稱為智人（Homo Sapiens）——理性的人，並非沒有道理。人們認為認知伴隨著意識：如果不了解思維過程，就無法思考。因此，當 19 世紀

中葉德國偉大的科學家（物理學家、生理學家、心理學家、哲學家）G·亥姆霍茲向科學界提出將知覺理解為無意識推論時，引起了一陣憤慨：這是不可能的。他的反對者說，這個想法——推理，就像任何其他類型的智能活動一樣，如果沒有意識的參與就無法進行！與此同時，今天所有認知科學領域的專家都承認“認知無意識”的存在。事實證明，大腦所進行的大部分巨大的認知工作（分析、綜合、概括、假設、各種類型的推論等）都是在沒有意識參與的情況下發生的。在這方面，出現了這樣的問題：意識在這種情況下是什麼以及它在生命世界的進化和認知主體的生命中扮演什麼角色。

如今，認知科學領域的大多數專家都認為，意識不能簡化為某些認知功能，甚至不能簡化為一組認知功能。人工智能系統可能有很多這樣的功能，但它很難歸因於意識的存在。什麼是意識？有觀點認為，這是認知系統對自身狀態的一種自我報告。但是，如果不伴隨意識（當某些水平的大腦活動代表並控制其他人的工作時），這種自我報告是不可能的嗎？還有一種觀點：意識是一種特殊的體驗，是一種伴隨認知活動的現象場。然而，無論如何，問題出現了：如果如此多的認知過程在沒有意識參與的情況下成功發生，那麼為什麼還需要意識，它的作用是什麼（一位哲學家寫了一篇文章，題為：“為什麼不是所有的認知過程？過程在黑暗中進行？”）？是否有可能通過研究認知主體的外部行為及其大腦的功能來客觀地研究意識？一般來說，是否存在意識的客觀標記，或者總是有必要轉向主體並向他詢問主觀體驗？那麼動物意識的研究又該怎麼辦呢？畢竟，不可能去詢問他們的經歷！如果我們假設不僅人類，其他生物也有意識，那麼我們如何通過外在跡象來識別它的存在呢？我們通常認為狗、貓、馬、猿和其他哺乳動物有意識，因為它們在某些方面與我們相似。但蟑螂或蠕蟲有它嗎？這些都是認知科學領域爭論激烈的棘手問題。如果沒有這些問題的答案，就不可能考慮設計人工意識——而今天就有這樣的嘗試。

人工智能系統在生活和科學知識中的運用，以及認知科學的深入發展，創造了人類學革命和人類學挑戰的局面。

大數據的機器處理產生人類無法做出的預測。但通常預測（基於簡單的觀察或在科學理論的幫助下）預示著其解釋和理解的可能性。機器預測並不伴隨著對其如何獲得的解釋和理解（人不理解這一點，機器也不理解，但後者不需要它）。在這種情況下，我們面對的是一個與我們陌生的心靈。這就是理性，但不是迄今為止我們所熟悉的、使我們成為“智人”的理性。此外，在某些情況下，機器可能會產生錯誤的預測——盲目信任其建議是有風險的。在科學研究中使用大數據的機器處理使一些專家認為不再需要科學理論（在此基礎上可以理解和解釋所研究的現象），就像不再需要科學假設和實驗一樣。更需要。如果我們認真對待這個想法，那麼我們習慣的科學就沒有未來。

使用人工智能來控制技術系統（特別是無人駕駛車輛）以及法律（所謂的“機器可讀法律”）和其他生活領域，尖銳地引發了許多倫理問題：在這些新條件下維持自主人的角色，道德代理人對其行為的後果負責，創建不會傷害人的“道德”智能系統，信任人工智能建議的標準。這與關於知識、理性、意識和科學理論結構的問題一樣，都是基本的哲學問題，不僅具有巨大的研究意義，而且在今天被證明是極其實用的。

偉大的哲學家康德曾經寫下這句話來回答“什麼是人？”這個問題。必須回答另外三個問題：“我能知道什麼？”、“我應該做什麼？”、“我能期待什麼？”今天我們必須重新回答所有這些問題。一個人的未來取決於這個答案。