

Научная статья
УДК 004.89
ББК 32.813
С 37
DOI: 10.53598/2410-3225-2023-4-331-61-72

Концептуальные основы реализации человеко-машинного взаимодействия (Рецензирована)

Владимир Сергеевич Симанков¹, Стефан Владимирович Онищенко²

¹ Кубанский государственный технологический университет, Краснодар, Адыгейский государственный университет, Майкоп, Россия, vs@simankov.ru

² Адыгейский государственный университет, Майкоп, Россия, osv@adygnet.ru

Аннотация. В работе представлен обзор исследований в области организации человеко-машинных систем, определена возможная схема построения системного интеллекта как инструмента для реализации процессов принятия решений в сложных информационных системах с множеством различных интеллектуальных агентов, подразумевающая комплексное взаимодействие различных видов интеллектов в зависимости от имеющегося уровня неопределенности в системе на конкретном этапе ее функционирования.

Ключевые слова: искусственный интеллект, интеллектуальная система, система, системный интеллект, совместное принятие решений, человеко-машинное взаимодействие, человеческий интеллектуальный агент

Original Research Paper

Conceptual bases for realization of human-machine interaction

Vladimir S. Simankov¹, Stefan V. Onishchenko²

¹ Kuban State Technological University, Krasnodar, Adyghe State University, Maykop, Russia, vs@simankov.ru

² Adyghe State University, Maykop, Russia, osv@adygnet.ru

Abstract. The paper presents an overview of research in the field of organization of human-machine systems, defines a possible scheme for building system intelligence as an instrument for implementing decision-making processes in complex information systems with many different intellectual agents, implying a complex interaction of various types of intelligence depending on the existing level of non-certainty in the system at a specific stage of its functioning.

Keywords: artificial intelligence, intelligent system, system, system intelligence, collaborative decision making, human-machine interaction, human intelligent agent

Введение

В последнее десятилетие во всем мире наблюдается небывалый рост исследовательского интереса к области, связанной с искусственным интеллектом [1, 2]. Был пройден путь по самым различным направлениям: от кибернетических систем с использованием простых обратных связей и обобщенных когнитивных архитектур до моделей обучения с подкреплением, которые организуют знания и итеративно совершенствуют сами себя с каждой новой итерацией [3], в результате чего удалось получить новые модели, способные использовать глубокие закономерности в коллективных знаниях для генерации новых идей [4]. По мере того, как масштабы и распространение систем, работающих на базе искусственного интеллекта, стремительно растут, объем ин-

формации и темпы изменений опережают ограниченные когнитивные способности людей. Следовательно, люди все больше вынуждены полагаться на физические и цифровые машины, которые упрощают общие коллективные действия – увеличивают память, управляют вниманием и помогают координировать наши коллективные решения.

Много лет идут различные споры и дискуссии относительно появления и применения сильного искусственного интеллекта (ИИ) [5], однако в настоящий момент исследователи считают, что в ближайшее время данная технология не может быть реализована и стоит обратиться в область реализации эффективного человеко-машинного взаимодействия [6]. В данной работе рассмотрены существующие исследования в направлении организации эффективных систем человеко-машинного взаимодействия и определен возможный механизм взаимодействия интеллектуальных агентов в рамках понятия системного интеллекта, введенного авторами в предыдущих работах [7, 8].

Материалы и методы

В последние годы объем исследований в области искусственного интеллекта стремительно вырос, появилось много новых источников литературы и всевозможных обзоров, представляющих различные взгляды на развитие данной области и на возможные новые технологии, направленные на изменение эффективности использования средств вычислительной техники. С одной стороны, это в значительной степени связано с увеличением мощности компьютеров, ускорением коммуникаций и ростом объема больших данных, что открывает широкое пространство для использования всех существующих технологий искусственного интеллекта. С другой стороны, развитие ИИ связано с высокими ожиданиями и надеждами, которые обусловлены возможностью реализации сильного ИИ, общего ИИ (AGI) и их многочисленных вариаций [9].

Чтобы понимать текущее состояние дел в области ИИ, важно различать такие понятия, как:

- слабый ИИ (AWI);
- узкий ИИ (ANI);
- общий ИИ (AGI);
- гибридный ИИ (HAI);
- сильный ИИ;
- искусственный суперинтеллект.

При этом все множество этих понятий может быть сведено к следующим трем основным категориям [10] (рис. 1):

– Artificial Narrow Intelligence (ANI) – тип интеллекта, присущий реальным интеллектуальным системам, предназначенным для выполнения конкретного (достаточно узкого) круга задач. Подобные интеллектуальные агенты «запрограммированы» на выполнение поставленных условий, что не позволяет им формализовать неизвестную задачу;

– Artificial General Intelligence (AGI) – подобный тип интеллектуальных агентов чаще всего интерпретируется исследователями как наиболее близкий к человеческому, способному самостоятельно разрешить более широкий спектр задач в сравнении с ANI;

– Artificial Super Intelligence (ASI) – согласно работе [11], подобные интеллектуальные агенты делятся на три типа:

- скоростной ASI – подобные агенты решают задачи быстрее человека;
- коллективный ASI – подобные агенты являются эквивалентом группе людей, принимающих решения;
- качественный ASI.

Позже эта концепция была пересмотрена различными авторами [12, 13], в результате чего было определено общепринятое на настоящий момент разделение искусственного интеллекта на два типа: сильный и слабый [14].

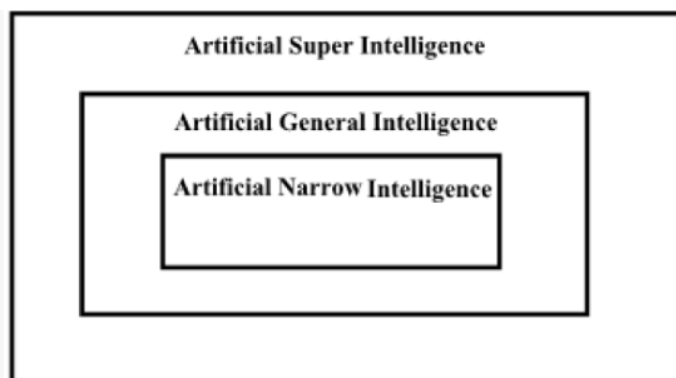


Рис. 1. Диаграмма Венна для определений искусственного интеллекта
 Fig. 1. Venn diagram for artificial intelligence definitions

Помимо этого, существуют различные концепции, предвосхищающие появление новых парадигм в рамках искусственного интеллекта и организацию систем активного логического вывода, рассмотренные коллективом авторов в работе [15]:

- systemic intelligence – как современный этап развития ИИ;
- sentient intelligence – активный вывод, основой которого будет распространение убеждений и «реакция» на сенсорные воздействия;
- sophisticated intelligence – ИИ, основанный на некотором опыте о представлениях мироустройства;
- sympathetic (sapient) intelligence – ИИ, способный находиться в некотором симбиозе с пользователями, учитывать их особенности, совместно воспринимать перспективы;
- shared (super) intelligence – искусственный суперинтеллект, позволяющий объединить множество интеллектуальных агентов в единую сеть, представляющий планетарный масштаб.

Текущее состояние и развитие слабого ИИ, а также концепции будущего сильного ИИ и искусственного сверхинтеллекта были получены из науки, культуры и философии на основе изучения мыслительных способностей человека и функций его мозга [16] (рис. 2).



Рис. 2. Особенности человека и «машины»
 Fig. 2. Human and “machine” features

В определении понятия гибридного интеллекта существуют некоторые различия, поскольку он может определяться либо как сочетание нескольких интеллектуаль-

ных методов [17], либо как искусственный интеллект, погруженный в гибридную реальность, где человек перестает быть наблюдателем и становится когнитивной частью системы [18]. Аналогично происходит ошибочное рассмотрение понятий общего ИИ, искусственного суперинтеллекта и сильного ИИ как схожих понятий [19].

Для более точного понимания концепции интеллектуального человеко-машинного взаимодействия авторы данной работы в предыдущих статьях [7, 8] ввели понятие системного интеллекта, позволяющего объединить в единую систему оценки все интеллекты, содержащиеся в системе, в результате чего можно объединить оба понятия гибридного интеллекта в одно целое. Однако понятие системного интеллекта все же подразумевает участие человека в работе системы, хотя это и не обязательный шаг, поскольку она может функционировать только с использованием методов ИИ, что определяется в зависимости от уровня существующей неопределенности исходной информации, поступающей на вход рассматриваемой информационной системы.

С развитием технологий искусственного интеллекта интеллектуальные машины все чаще наделяются человеческими способностями, такими как автономное принятие решений, рассуждение, активное взаимодействие и осознание ситуации. Интеллектуальные машины могут действовать как равные с людьми и сотрудничать с ними для выполнения задач принятия решений. Способность сотрудничать с людьми стала показателем уровня интеллекта машины и определяет сферу и глубину ее применения [20]. Системы ИИ способны решать конкретные задачи, возможно, гораздо лучше, чем это сделали бы люди. Системы AGI способны решать множество задач различного характера с приемлемым уровнем производительности:

- распознавание образов;
- помощь при покупке за счет формирования рекомендаций на основе предпочтений;
- построение прогноза погоды;
- осуществление контроля транспорта;
- постановка медицинских диагнозов;
- обработка речи;
- прочесывание цифровых следов;
- прогнозирование вероятности совершения преступления.

Рассмотренные выше типы интеллектуальных агентов можно объединить под понятием человекоподобного интеллекта, основанного на существующих знаниях об устройстве человеческого мозга и человеческой сущности (рис. 3).

Рассмотрим более подробно ряд проблем, присущих «человекоподобным» интеллектуальным агентам [21, 22], которые могут быть разрешены или уменьшены за счет непосредственного вовлечения человека в работу интеллектуальной системы.

Определение и формулирование проблемы. Данный этап является очень важным и должен быть реализован в автономных системах ИИ, основанных на принципе подобия человеческому интеллекту. Однако существует ряд задач, которые трудно формализуемы даже людьми [23], поэтому они не могут быть переложены для исполнения интеллектуальными системами. В настоящее время формулирование подобных проблем остается за разработчиками системы, а не является частью функций интеллектуального агента системы. Для решения этой проблемы иногда используются агенты, основанные на знаниях, или различные механизмы вывода [24].

Проблемы данных. Как известно, современные интеллектуальные модели основываются на большом количестве входных данных, использование которых позволяет правильно настроить систему для решения конкретных задач. Часть подобных проблем может быть разрешена в системе в автоматическом режиме [25]. Но некоторые из них трудноразрешимы без участия человека: неоднородность, недостаточность, несбалансированность, ненадежность данных – вот некоторые из аспектов, значительно повы-

шающих степень неопределенности в системе [26], которые не могут быть полностью разрешены без участия человеческого интеллектуального агента.



Рис. 3. Проблемы в ИИ и их основные причины

Fig. 3. Problems in AI and their main causes

Проблемы надежности. В некоторых работах [27] обсуждается разница между такими понятиями, как точность и надежность модели, которая может быть оценена посредством оценки результатов работы системы группой экспертов-аналитиков. Проблемы с надежностью модели могут возникнуть в случае непредвиденного и аномального изменения входных данных (возможно в результате деятельности злоумышленников), что не может быть отслежено самой системой в автономном режиме функционирования.

Объяснимость ИИ. Несмотря на развитие подобных методов организации интеллектуальных агентов [28, 29], данный аспект имеет важное значение в критически важных задачах, требующих понятного механизма принятия решений, что не может быть обеспечено за счет использования только лишь моделей ИИ. В таком случае идеальным решением становится вариант системы с непосредственным участием человеческого агента, который принимает непосредственное и окончательное решение на основе полученной информации от интеллектуальной системы, либо же корректирует процесс принятия решения машинным агентом.

Рациональность ИИ. Одним из важных аспектов, учитываемых при создании интеллектуального агента, является рациональность. Вычисление рациональности играет важную роль в распределенном машинном обучении, многоагентных системах, теории игр и пр. [30, 31]. Согласно исследованию [32], рациональность может быть разбита на четыре основные категории:

- совершенная рациональность – агенты с таким типом рациональности способны генерировать максимально успешное поведение/решение за счет использования доступной им информации;
- расчетная рациональность – подобный интеллектуальный агент может определить совершенно рациональное решение, учитывая только изначально доступную информацию, обладающую достаточной степенью полноты;
- метаярность рациональности – агент с этим типом рациональности может выбрать оптимальную комбинацию «вычисления-последовательность-плюс действие».

При этом ограничение состоит в том, что должно быть выбрано действие, в соответствии с которым производится вычисление;

– ограниченная рациональность – агент с таким типом рациональности может вести себя успешно исходя из имеющейся информации и вычислительных ресурсов.

На практике же становится понятно, что возможно создание агента только с одним из типов рациональности, а наличие неопределенности в исходной информации не позволяет реализовать интеллектуальную систему с идеальной рациональностью.

Обман ИИ. Поскольку в недавнем прошлом были разработаны генеративные модели ИИ, одной из основных современных проблем использования технологий ИИ является вводящее в заблуждение совпадение концепций причинности и корреляции. Но такие корреляции не всегда обусловлены причинно-следственной связью. Традиционный ANI не может преобразовать корреляцию в причинно-следственную связь и не может обеспечить высокий уровень прозрачности или доверия [33], в результате чего можно получать заведомо ложные результаты при абсолютно штатной работе системы.

Для решения описанных выше проблем необходимо привлечь дополнительных интеллектуальных агентов в виде людей, которые, выступая в роли эксперта, смогут провести оценку и при необходимости коррекцию результатов, полученных интеллектуальной системой. Поскольку часть задач является слабоформализуемой и не пригодна для решения за счет использования интеллектуальных систем, возникает необходимость прямого вовлечения человека в процесс решения задачи и формирования конечных решений.

В литературе в последнее время стали появляться различные исследования [20, 34–36], рассматривающие новые концепции в возможной организации систем с вовлечением интеллектуальных агентов различной природы с целью реализации наиболее эффективной вычислительной системы. Интерес в данной области обусловлен тем фактом, что создание искусственного интеллекта, подобного человеческому, в ближайшей перспективе не видится возможным из-за различных проблем и ограничений, присутствующих искусственному интеллекту и системам на его основе [37].

Авторы этих работ также предприняли некоторые попытки для описания возможных механизмов взаимодействия человека и интеллектуальной системы с использованием различных технологий. Так, в работе [20] группа авторов, рассматривая вопросы реализации возможных механизмов взаимодействия человека и машины, отмечает необходимость перехода от системы ведущего и ведомого к коллегиальному совместному принятию решений (рис. 4), что порождает сотрудничество, а не простое выполнение последовательности заданных операций.

В работе авторы не выделяют четкого механизма, за счет использования которого можно было бы реализовать подобное взаимодействие. При этом ими отмечено, что существующие исследования в данной области можно условно разделить на три большие категории:

– человеко-машинные системы. В данных работах представляются подходы, предназначенные для реализации интеллектуальных систем, направленных на совместное принятие решений [38, 39];

– механизмы интеграции интеллекта. Сюда относятся как работы, изучающие доверие и объяснимость ИИ, так и описывающие фундаментальные принципы и технологии для слияния информации [40–42];

– совместное принятие решений. Работы в этом направлении изучают способы распределения полномочий между машинами и людьми [43] и режимы их взаимодействия [44].

Исходя из этого, можно предположить, что механизм взаимодействия должен основываться на интерфейсе, построенном за счет применения самых современных

технологических решений (рис. 5).

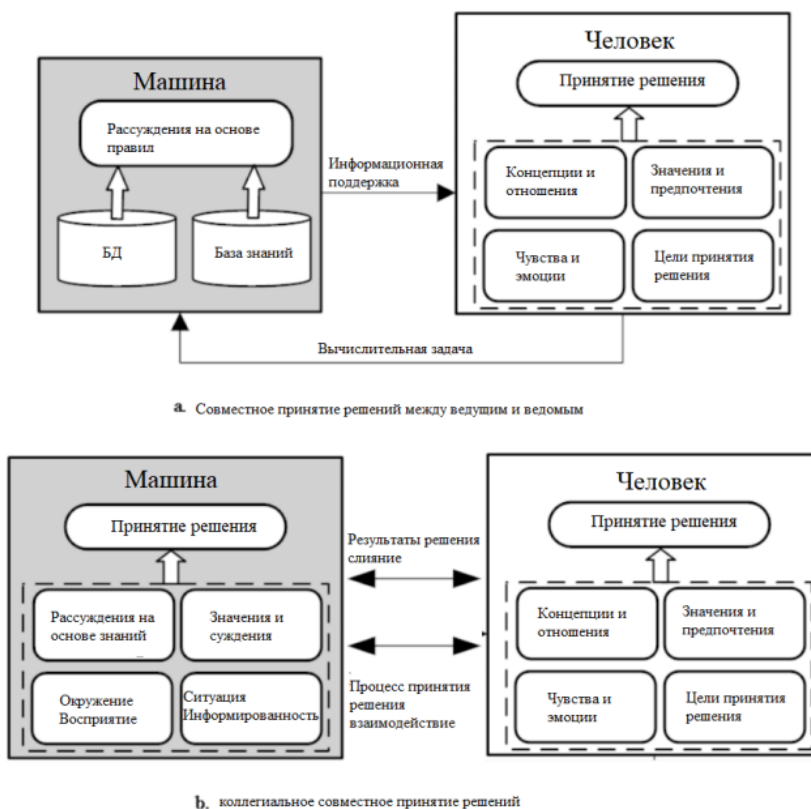


Рис. 4. Переход отношений между человеком и машиной от ведущего-ведомого к равноправным лицам, принимающим решения

Fig. 4. Transition of human-machine relationship from lead-slave to equal decision-makers

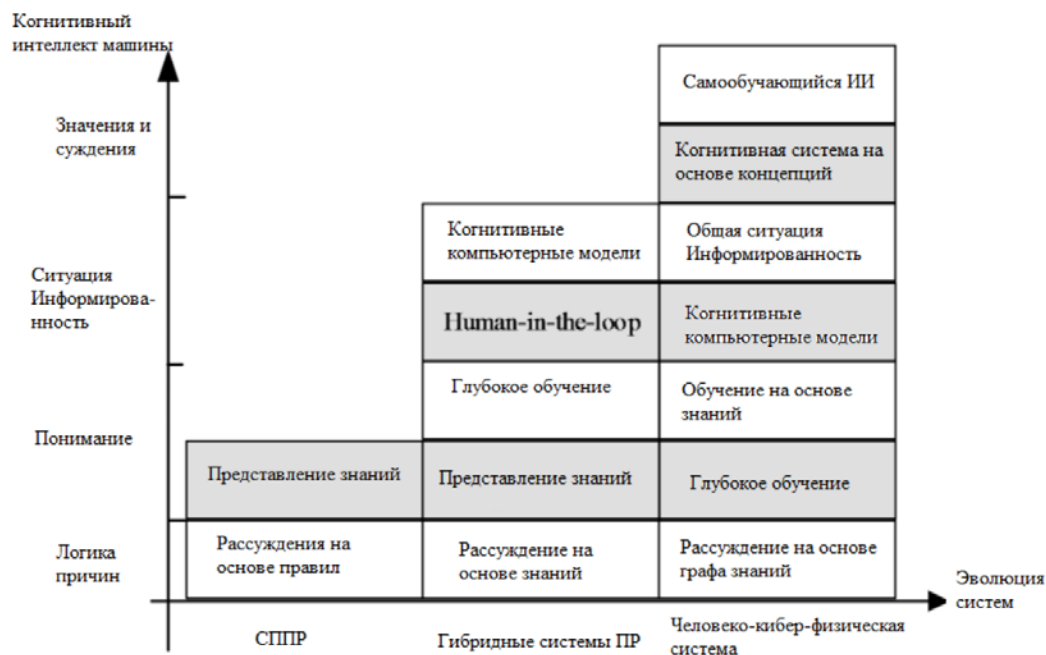


Рис. 5. Инструменты для реализации человеко-машинных интеллектуальных систем для совместного принятия решений

Fig. 5. Tools for implementing man-machine intelligent systems for joint decision-making

Используя подобный перечень технологий, можно реализовать интеллектуальную систему, позволяющую принимать человеку непосредственное участие в процессе формирования итогового решения и выступать в качестве эксперта в процессе формирования решения самой системой. При этом важным аспектом является уровень неопределенности в системе, поскольку именно она является решающим фактором, который позволяет определить необходимость привлечения человека для решения конкретной задачи. На рисунке 6 представлена схема подобного совместного формирования решений человеком и машиной.

Основываясь на определении исходного уровня неопределенности в вычислительной системе, мы можем оптимизировать процесс принятия решений и работы человеческого интеллектуального агента, поскольку при низком значении неопределенности интеллектуальная система способна самостоятельно решить формализованную задачу с высокой степенью доверия [45]. Таким образом, это упрощает взаимодействие человека и системы, позволяя человеку включаться в работу только при необходимости и принимать участие в формировании решения только в случае необходимости. При этом в любом из случаев дополнительной задачей, возлагаемой на человека в подобной системе, является осуществление контроля над ее функционированием и проверка формируемого решения на основе своих личных навыков и жизненного опыта.

В качестве инструмента подобного взаимодействия необходимо использовать когнитивный интерфейс, позволяющий человеку оказывать непосредственное влияние на процесс функционирования системы, производить корректировку параметров, критериев и постановку самой задачи, поскольку данные задачи на настоящий момент времени относятся к типу неформализуемых и не могут быть алгоритмизированы для решения в четком виде. В качестве дополнительного инструмента в данном случае может быть применена технология генеративного искусственного интеллекта (например, GhatGPT), которая позволяет генерировать различные идеи при правильной постановке вопросов [46], но при этом человек должен обязательно проверить корректность сгенерированного запроса.

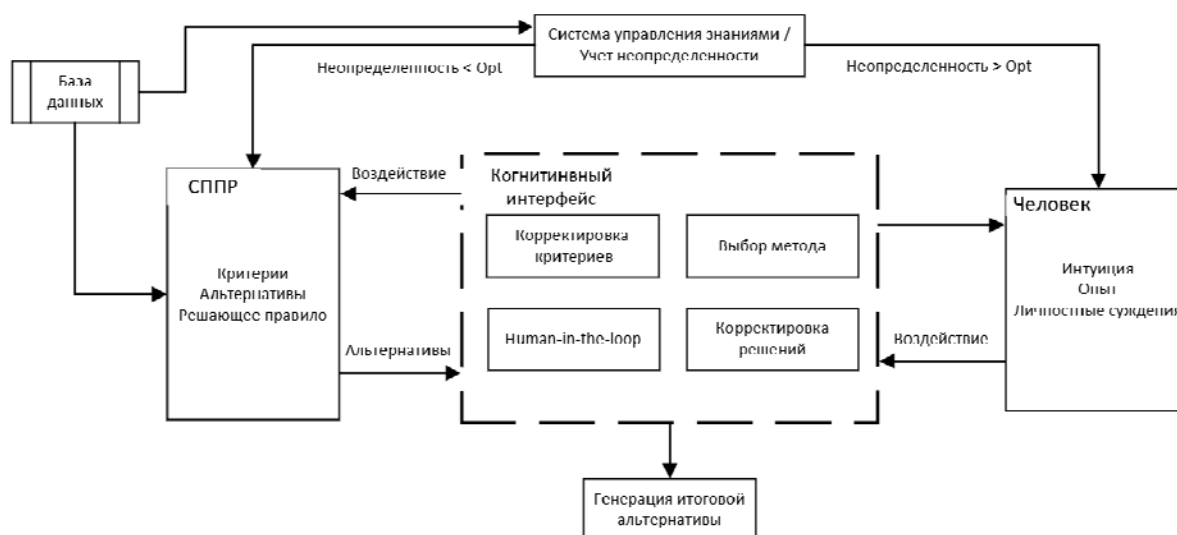


Рис. 6. Схема взаимодействия человека и машины при реализации системного интеллекта в рамках интеллектуальной вычислительной системы

Fig. 6. Diagram of interaction between a person and a machine in the implementation of system intelligence within the framework of an intelligent computing system

Заключение

В работе были рассмотрены последние достижения в области развития человеко-машинного взаимодействия, анализ которых доказывает, что это направление исследо-

ваний вызывает большой интерес у научного сообщества в свете невозможности создания сильного искусственного интеллекта в ближайшей перспективе. Различные авторы описывают концепции возможного совместного принятия решений компьютером и человеком, однако они не затрагивают технологии и конкретные методы, которыми это взаимодействие могло бы быть реализовано. Был предложен вариант организации взаимодействия человека и машины, который может использоваться при реализации системного интеллекта в вычислительной системе, отличительной особенностью которого является учет уровня неопределенности в системе, на основе чего принимается решение о необходимости привлечения человека в процесс ее функционирования на определенном этапе. Подобное взаимодействие может быть реализовано за счет создания когнитивного интерфейса, который использует различные технологии (например, Human-In-Loop), что позволит организовать гибкую систему, которая предоставит возможности человеку оказывать влияние на работу интеллектуальной системы и выступать в роли дополнительного интеллектуального агента.

Примечания

1. Andrist Sean, Bohus Dan. Accelerating the Development of Multimodal, Integrative-AI Systems with Platform for Situated Intelligence. 2020. ArXiv: 2010.06084v1. URL: <https://doi.org/10.48550/arXiv.2010.06084>
2. Bohus Dan, Andrist Sean, Jalobeanu Mihai. Rapid development of multimodal interactive systems: a demonstration of platform for situated intelligence. 2017. DOI: 10.1145/3136755.3143021
3. Silamut A., Petsangsri S. Self-directed learning with knowledge management model to enhance digital literacy abilities // Education and Information Technologies. 2020. Vol. 25. P. 4797–4815. URL: <https://doi.org/10.1007/s10639-020-10187-3>
4. Generative Model Using Knowledge Graph for Document-Grounded Conversations / B. Kim, D. Lee, D. Kim [et al.] // Applied Sciences. 2022. Vol. 12, Iss. 7. P. 3367. DOI: 10.3390/app12073367
5. Artificial intelligence: A powerful paradigm for scientific research / Xu Yongjun [et al.] // The Innovation. 2021. Vol. 2, Iss. 4. P. 100179. URL: <https://doi.org/10.1016/j.xinn.2021.100179>
6. The concept of “interaction” in debates on human-machine interaction / S. Schleidgen, O. Friedrich, S. Gerlek [et al.] // Humanities and Social Sciences Communications. Palgrave Communications. 2023. Vol. 10, Iss. 1. P. 551. DOI: 10.1057/s41599-023-02060-8
7. An Approach to the Definition of System Intelligence in the Management of Complex Systems / V.S. Simankov, S.V. Onishchenko, P.Yu. Buchatskiy, S.V. Teploukhov / 26th International Conference on Soft Computing and Measurements (SCM). Saint Petersburg, 2023. P. 159-162. DOI: 10.1109/SCM58628.2023.10159122
8. Симанков В.С., Онищенко С.В. Методологические основы системного интеллекта // Вестник Адыгейского государственного университета. Сер.: Естественно-математические и технические науки. 2023. Вып. 1 (316). С. 21-30. DOI: 10.53598/2410-3225-2023-1-316-21-30
9. Lyre H. The state space of artificial intelligence // Minds and Machines. 2020. Vol. 30. P. 325-347.
10. Saghiri A.M. A Survey on Challenges in Designing Cognitive Engines // Proceedings of the 2020 6th International Conference on Web Research (ICWR). Tehran, 2020. P. 165–171.
11. Boström N. Superintelligence: Paths, Dangers, Strategies. Oxford: Oxford University Press, 2014. 328 p.
12. Chollet F. On the measure of intelligence. ArXiv 2019, arXiv: 1911.01547. URL: <https://doi.org/10.48550/arXiv.1911.01547>
13. Yampolskiy R. V. Human is not equal to AGI. ArXiv 2020, arXiv: 2007.07710. URL: <https://doi.org/10.48550/arXiv.2007.07710>
14. Searle J.R. Minds, brains, and programs // Behavioral and Brain Sciences. 1980. No. 3. P. 417–424.
15. Designing ecosystems of intelligence from first principles / Karl J. Friston [et al.]. 2022. ArXiv preprint arXiv: 2212.01354. URL: <https://doi.org/10.48550/arXiv.2212.01354>
16. Sundval S. “Artificial intelligence” in Critical Terms in Futures Studies. Cham, Switzerland: Palgrave, 2019. P. 29–34.
17. Гаврилов А.В. Гибридные интеллектуальные системы. Новосибирск: Изд-во НГТУ, 2003. 164 с.
18. Perko I. Hybrid reality development – Can social responsibility concepts provide guidance? // Kybernetes. 2021. Vol. 50, Iss. 3. P. 676–693.
19. Fjelland R. Why general artificial intelligence will not be realized // Humanities & Social Sciences Communications. 2020. Vol. 7, Iss. 10. P. 1–9. DOI: <https://doi.org/10.1057/s41599-020-0494-4>
20. Ren Minglun, Chen Nengying, Qiu Hui. Human-machine Collaborative Decision-making: An Evolutionary Roadmap Based on Cognitive Intelligence // International Journal of Social Robotics. 2023. Vol. 15. P. 1101–1114. DOI: 10.1007/s12369-023-01020-1
21. Drawbacks of Artificial Intelligence and Their Potential Solutions in the Healthcare Sector / B. Khan, H. Fatima, A. Qureshi [et al.] // Biomed Mater Devices. 2023. Vol. 8. P. 1–8. DOI: 10.1007/s44174-

023-00063-2

22. A Survey of Artificial Intelligence Challenges: Analyzing the Definitions, Relationships, and Evolutions / A.M. Saghiri, S.M. Vahidipour, M.R. Jabbarpour [et al.] // Applied Sciences. 2022. Vol. 12, Iss. 8. URL: <https://doi.org/10.3390/app12084054>

23. Linz P. An Introduction to Formal Languages and Automata. Burlington: Jones & Bartlett Learning, 2006. 415 p.

24. Toward programs with common sense / D.B. Lenat, R.V. Guha, K. Pittman [et al.] // Communications of the ACM. 1990. Vol. 33, Iss. 8. P. 30–49. DOI: 10.1145/79173.79176

25. Baig M.I., Shuib L., Yadegaridehkordi E. Big Data Tools: Advantages and Disadvantages // Journal of Soft Computing and Decision Support Systems. 2019. Vol. 6, No. 6. P. 14–20.

26. Enhancing robustness of machine learning systems via data transformations / A.N Bhagoji, D. Cullina, C. Sitawarin, P. Mittal // Proceedings of the 2018 52nd Annual Conference on Information Sciences and Systems (CISS). Princeton, NJ, USA, 21–23 March 2018. Princeton, 2018. P. 1–5. DOI: 10.1109/CISS.2018.8362326

27. Rozsa A., Günther M., Boulton T.E. Are accuracy and robustness correlated? // Proceedings of the 2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA). Anaheim, CA, USA, 18–20 December 2016. Anaheim, 2016. P. 227–232. DOI: 10.1109/ICMLA.2016.0045

28. On fairness and calibration / G. Pleiss, M. Raghavan, F. Wu [et al.] // Advances in Neural Information Processing Systems. 2017. Vol. 30. 10 p.

29. Explainable AI in industry / K. Gade, S.C. Geyik, K. Kenthapadi [et al.] // Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. Anchorage, AK, USA, 4–8 August 2019. Anchorage, 2019. P. 3203–3204.

30. Halpern J.Y., Pass R. Algorithmic rationality: Game theory with costly computation // Journal of Economic Theory. 2015. Vol. 156. P. 246–268.

31. Gigerenzer G., Selten R. Rethinking rationality. Bounded Rationality: The Adaptive Toolbox. Cambridge: MIT Press, 2001. Vol. 1. P. 37–50.

32. Russell S.J. Rationality and intelligence // Artificial Intelligence. 1997. Vol. 94. P. 57–77.

33. Pearl J. Theoretical impediments to machine learning with seven sparks from the causal revolution. ArXiv: 1801.04016, 2018. URL: <https://doi.org/10.48550/arXiv.1801.04016>

34. Toward Integrated Human-machine Intelligence for Civil Engineering: An Interdisciplinary Perspective / C. Zhang, J. Kim, J. Jeon [et al.]. 2021. ArXiv, abs/2107.13498. URL: <https://doi.org/10.48550/arXiv.2107.13498>

35. Wei Liu. Integrated Human-Machine Intelligence. Elsevier, 2023. 298 p. URL: <https://doi.org/10.1016/B978-0-323-99562-7.05001-4>

36. Khramov V.V. Development of a human-machine interface based on hybrid intelligence // Современные информационные технологии и ИТ-образование. 2020. Т. 16, № 4. С. 893-900. URL: <https://doi.org/10.25559/SITITO.16.202004.893-900>

37. Sahar Sultan. Limitations of Artificial Intelligence: Thesis. Rochester Institute of Technology. 2021. 49 p. URL: <https://scholarworks.rit.edu/theses/10975>

38. A trustworthy human-machine framework for collective decision making in food-energy-water management: the role of trust sensitivity / S. Uslu, D. Kaur, S.J. Rivera [et al.] // Knowledge-Based Systems. 2021. Vol. 213. 41 p. DOI: 10.1016/j.knosys.2020.106683

39. Yun Y., Ma D., Yang M. Human-computer interaction-based decision support system with applications in data mining // Future Generation Computer Systems. 2021. Vol. 114. P. 285–289.

40. Multi-sensor fusion for body sensor network in medical human-robot interaction scenario / K. Lin, Y. Li, J. Sun [et al.] // Information Fusion. 2020. Vol. 57. P. 15–26.

41. Interpretable machine learning: fundamental principles and 10 grand challenges / C. Rudin, C. Chen, Z. Chen [et al.] // Statistics Surveys. 2022. Vol. 16. P. 1–85. DOI: 10.1214/21-SS133

42. Vinuesa R., Sirmacek B. Interpretable deep-learning models to help achieve the Sustainable Development Goals // Nature Machine Intelligence. 2021. Vol. 3, Iss. 11. P. 926. URL: <https://doi.org/10.48550/arXiv.2108.10744>

43. Toward human-vehicle collaboration: review and perspectives on human-centered collaborative automated driving / Y. Xing, C. Lv, D. Cao, P. Hang // Transportation Research. Part C: Emerging Technologies. 2021. Vol. 128. P. 103199. DOI: 10.1016/j.trc.2021.103199

44. Machines as teammates: a research agenda on AI in team collaboration / I. Seeber, E. Bittner, R.O. Briggs [et al.] // Information & Management. 2020. Vol. 57, Iss. 2. P. 103174.

45. Lukyanenko R., Maass W., Storey V.C. Trust in artificial intelligence: From a Foundational Trust Framework to emerging research opportunities // Electronic Markets. 2022. Vol. 32. P. 1993–2020. URL: <https://doi.org/10.1007/s12525-022-00605-4>

46. Doshi Anil Rajnikant, Hauser Oliver. Generative Artificial Intelligence Enhances Creativity but Reduces the Diversity of Novel Content. 2023. 96 p. URL: <http://dx.doi.org/10.2139/ssrn.4535536>

References

1. Andrist Sean, Bohus Dan. Accelerating the Development of Multimodal, Integrative-AI Systems with Platform for Situated Intelligence. 2020. ArXiv: 2010.06084v1. URL: <https://doi.org/10.48550/arXiv.2010.06084>

2. Bohus Dan, Andrist Sean, Jalobeanu Mihai. Rapid development of multimodal interactive systems: a demonstration of platform for situated intelligence. 2017. DOI: 10.1145/3136755.3143021

3. Silamut A., Petsangsri S. Self-directed learning with knowledge management model to enhance digi-

- tal literacy abilities // Education and Information Technologies. 2020. Vol. 25. P. 4797–4815. URL: <https://doi.org/10.1007/s10639-020-10187-3>
4. Generative Model Using Knowledge Graph for Document-Grounded Conversations / B. Kim, D. Lee, D. Kim [et al.] // Applied Sciences. 2022. Vol. 12, Iss. 7. P. 3367. DOI: 10.3390/app12073367
5. Artificial intelligence: A powerful paradigm for scientific research / Xu Yongjun [et al.] // The Innovation. 2021. Vol. 2, Iss. 4. P. 100179. URL: <https://doi.org/10.1016/j.xinn.2021.100179>
6. The concept of “interaction” in debates on human–machine interaction / S. Schleidgen, O. Friedrich, S. Gerlek [et al.] // Humanities and Social Sciences Communications. Palgrave Communications. 2023. Vol. 10, Iss. 1. P. 551. DOI: 10.1057/s41599-023-02060-8
7. An Approach to the Definition of System Intelligence in the Management of Complex Systems / V.S. Simankov, S.V. Onishchenko, P. Yu. Buchatskiy, S.V. Teploukhov / 26th International Conference on Soft Computing and Measurements (SCM). Saint Petersburg, 2023. P. 159–162. DOI: 10.1109/SCM58628.2023.10159122
8. Simankov V.S., Onishchenko S.V. Methodological foundations system intelligence // The Bulletin of the Adyghe State University. Ser.: Natural-Mathematical and Technical Sciences. 2023. Iss. 1 (316). P. 21–30. DOI: 10.53598/2410-3225-2023-1-316-21-30
9. Lyre H. The state space of artificial intelligence // Minds and Machines. 2020. Vol. 30. P. 325–347.
10. Saghiri A.M. A Survey on Challenges in Designing Cognitive Engines // Proceedings of the 2020 6th International Conference on Web Research (ICWR). Tehran, 2020. P. 165–171.
11. Boström N. Superintelligence: Paths, Dangers, Strategies. Oxford: Oxford University Press, 2014. 328 p.
12. Chollet F. On the measure of intelligence. ArXiv 2019, arXiv: 1911.01547. URL: <https://doi.org/10.48550/arXiv.1911.01547>
13. Yampolskiy R. V. Human is not equal to AGI. ArXiv 2020, arXiv: 2007.07710. URL: <https://doi.org/10.48550/arXiv.2007.07710>
14. Searle J.R. Minds, brains, and programs // Behavioral and Brain Sciences. 1980. No. 3. P. 417–424.
15. Designing ecosystems of intelligence from first principles / Karl J. Friston [et al.]. 2022. ArXiv preprint arXiv: 2212.01354. URL: <https://doi.org/10.48550/arXiv.2212.01354>
16. Sundval S. “Artificial intelligence” in Critical Terms in Futures Studies. Cham, Switzerland: Palgrave, 2019. P. 29–34.
17. Gavrilov A.V. Hybrid intelligent systems. Novosibirsk: NSTU Publishing House, 2003. 164 p.
18. Perko I. Hybrid reality development – Can social responsibility concepts provide guidance? // Kybernetes. 2021. Vol. 50, Iss. 3. P. 676–693.
19. Fjelland R. Why general artificial intelligence will not be realized // Humanities & Social Sciences Communications. 2020. Vol. 7, Iss. 10. P. 1–9. DOI: <https://doi.org/10.1057/s41599-020-0494-4>
20. Ren Minglun, Chen Nengying, Qiu Hui. Human-machine Collaborative Decision-making: An Evolutionary Roadmap Based on Cognitive Intelligence // International Journal of Social Robotics. 2023. Vol. 15. P. 1101–1114. DOI: 10.1007/s12369-023-01020-1
21. Drawbacks of Artificial Intelligence and Their Potential Solutions in the Healthcare Sector / B. Khan, H. Fatima, A. Qureshi [et al.] // Biomed Mater Devices. 2023. Vol. 8. P. 1–8. DOI: 10.1007/s44174-023-00063-2
22. A Survey of Artificial Intelligence Challenges: Analyzing the Definitions, Relationships, and Evolutions / A.M. Saghiri, S.M. Vahidipour, M.R. Jabbarpour [et al.] // Applied Sciences. 2022. Vol. 12, Iss. 8. URL: <https://doi.org/10.3390/app12084054>
23. Linz P. An Introduction to Formal Languages and Automata. Burlington: Jones & Bartlett Learning, 2006. 415 p.
24. Toward programs with common sense / D.B. Lenat, R.V. Guha, K. Pittman [et al.] // Communications of the ACM. 1990. Vol. 33, Iss. 8. P. 30–49. DOI: 10.1145/79173.79176
25. Baig M.I., Shuib L., Yadegaridehkordi E. Big Data Tools: Advantages and Disadvantages // Journal of Soft Computing and Decision Support Systems. 2019. Vol. 6, No. 6. P. 14–20.
26. Enhancing robustness of machine learning systems via data transformations / A.N Bhagoji, D. Cullina, C. Sitawarin, P. Mittal // Proceedings of the 2018 52nd Annual Conference on Information Sciences and Systems (CISS). Princeton, NJ, USA, 21–23 March 2018. Princeton, 2018. P. 1–5. DOI: 10.1109/CISS.2018.8362326
27. Rozsa A., Günther M., Boulton T.E. Are accuracy and robustness correlated? // Proceedings of the 2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA). Anaheim, CA, USA, 18–20 December 2016. Anaheim, 2016. P. 227–232. DOI: 10.1109/ICMLA.2016.0045
28. On fairness and calibration / G. Pleiss, M. Raghavan, F. Wu [et al.] // Advances in Neural Information Processing Systems. 2017. Vol. 30. 10 p.
29. Explainable AI in industry / K. Gade, S.C. Geyik, K. Kenthapadi [et al.] // Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. Anchorage, AK, USA, 4–8 August 2019. Anchorage, 2019. P. 3203–3204.
30. Halpern J.Y., Pass R. Algorithmic rationality: Game theory with costly computation // Journal of Economic Theory. 2015. Vol. 156. P. 246–268.
31. Gigerenzer G., Selten R. Rethinking rationality. Bounded Rationality: The Adaptive Toolbox. Cambridge: MIT Press, 2001. Vol. 1. P. 37–50.
32. Russell S.J. Rationality and intelligence // Artificial Intelligence. 1997. Vol. 94. P. 57–77.
33. Pearl J. Theoretical impediments to machine learning with seven sparks from the causal revolution. ArXiv: 1801.04016, 2018. URL: <https://doi.org/10.48550/arXiv.1801.04016>
34. Toward Integrated Human-machine Intelligence for Civil Engineering: An Interdisciplinary Perspective / C. Zhang, J. Kim, J. Jeon [et al.]. 2021. ArXiv, abs/2107.13498. URL:

<https://doi.org/10.48550/arXiv.2107.13498>

35. Wei Liu. Integrated Human-Machine Intelligence. Elsevier, 2023. 298 p. URL:

<https://doi.org/10.1016/B978-0-323-99562-7.05001-4>

36. Khramov V.V. Development of a human-machine interface based on hybrid intelligence // Современные информационные технологии и ИТ-образование. 2020. Т. 16, № 4. С. 893-900. URL: <https://doi.org/10.25559/SITITO.16.202004.893-900>

37. Sahar Sultan. Limitations of Artificial Intelligence: Thesis. Rochester Institute of Technology. 2021. 49 p. URL: <https://scholarworks.rit.edu/theses/10975>

38. A trustworthy human-machine framework for collective decision making in food-energy-water management: the role of trust sensitivity / S. Uslu, D. Kaur, S.J. Rivera [et al.] // Knowledge-Based Systems. 2021. Vol. 213. 41 p. DOI: 10.1016/j.knosys.2020.106683

39. Yun Y., Ma D., Yang M. Human-computer interaction-based decision support system with applications in data mining // Future Generation Computer Systems. 2021. Vol. 114. P. 285-289.

40. Multi-sensor fusion for body sensor network in medical human-robot interaction scenario / K. Lin, Y. Li, J. Sun [et al.] // Information Fusion. 2020. Vol. 57. P. 15-26.

41. Interpretable machine learning: fundamental principles and 10 grand challenges / C. Rudin, C. Chen, Z. Chen [et al.] // Statistics Surveys. 2022. Vol. 16. P. 1-85. DOI: 10.1214/21-SS133

42. Vinuesa R., Sirmacek B. Interpretable deep-learning models to help achieve the Sustainable Development Goals // Nature Machine Intelligence. 2021. Vol. 3, Iss. 11. P. 926. URL: <https://doi.org/10.48550/arXiv.2108.10744>

43. Toward human-vehicle collaboration: review and perspectives on human-centered collaborative automated driving / Y. Xing, C. Lv, D. Cao, P. Hang // Transportation Research. Part C: Emerging Technologies. 2021. Vol. 128. P. 103199. DOI: 10.1016/j.trc.2021.103199

44. Machines as teammates: a research agenda on AI in team collaboration / I. Seeber, E. Bittner, R.O. Briggs [et al.] // Information & Management. 2020. Vol. 57, Iss. 2. P. 103174.

45. Lukyanenko R., Maass W., Storey V.C. Trust in artificial intelligence: From a Foundational Trust Framework to emerging research opportunities // Electronic Markets. 2022. Vol. 32. P. 1993-2020. URL: <https://doi.org/10.1007/s12525-022-00605-4>

46. Doshi Anil Rajnikant, Hauser Oliver. Generative Artificial Intelligence Enhances Creativity but Reduces the Diversity of Novel Content. 2023. 96 p. URL: <http://dx.doi.org/10.2139/ssrn.4535536>

Авторы заявляют об отсутствии конфликта интересов.

Статья поступила в редакцию 25.11.2023; одобрена после рецензирования 05.12.2023; принята к публикации 06.12.2023.

The authors declare no conflicts of interests.

The article was submitted 25.11.2023; approved after reviewing 05.12.2023; accepted for publication 06.12.2023.

© **В.С. Симанков,** С.В. Онищенко, 2023