

Сметана Владимир Васильевич,
5.7.1. Онтология и теория познания (философские науки),
кандидат философских наук, директор,
АНО НИИ «ЦИФРОВОЙ ИНТЕЛЛЕКТ», Москва, Россия
Vladimir Smetana,
Candidate of philosophical sciences, PhD,
DIGITAL INTELLIGENCE RESEARCH INSTITUTE,
Moscow, Russia

ЗАГЛЯДЫВАЯ В БУДУЩЕЕ:
ГОРИЗОНТЫ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА
LOOKING INTO THE FUTURE:
THE HORIZONS OF ARTIFICIAL INTELLIGENCE

Аннотация: Данная статья предлагает заглянуть в будущее искусственного интеллекта (ИИ), анализируя ключевые тенденции и вызовы, которые будут определять его развитие. В статье рассматриваются перспективные направления развития машинного обучения и его интеграции с другими передовыми технологиями. Особое внимание уделяется возможности создания общего искусственного интеллекта (ОИИ), обладающего широким спектром когнитивных способностей, и анализируются как потенциальные выгоды, так и риски, связанные с подобным прорывом.

В заключении статьи подчеркивается критическая важность ответственного подхода к развитию искусственного интеллекта, включающего разработку этических принципов, государственное регулирование и широкий общественный диалог.

Abstract: This article offers a glimpse into the future of artificial intelligence (AI), analyzing the key trends and challenges that will determine its development. The article discusses promising areas for the development of machine learning and its integration with other advanced technologies. Special attention is paid to the possibility of creating an artificial general intelligence (AGI) with a wide range of cognitive abilities, and both the potential benefits and risks associated with such a breakthrough are analyzed.

In conclusion, the article emphasizes the critical importance of a responsible approach to the development of artificial intelligence, including the development of ethical principles, government regulation and a broad public dialogue.

Ключевые слова: искусственный интеллект (ИИ), общий искусственный интеллект (ОИИ), объяснимый ИИ, эмоциональный ИИ, безопасность ИИ, прозрачность ИИ, этика ИИ.

Keywords: artificial intelligence (AI), artificial general intelligence (AGI), Explainable AI (XAI), Emotional AI, AI security, AI transparency, AI ethics.

Глава 1. Будущие тенденции и потенциальные достижения в искусственном интеллекте

Искусственный интеллект (ИИ) уже сегодня проник во многие сферы нашей жизни, и нет сомнений, что в будущем его влияние будет только усиливаться [1] Рассмотрим некоторые из ключевых тенденций, которые, вероятно, будут определять развитие ИИ в ближайшие годы и десятилетия.

1. Новые горизонты машинного обучения и глубокого обучения

Машинное обучение (МО) и глубокое обучение (DL) являются движущими силами современного искусственного интеллекта, [2] и можно ожидать, что они будут активно развиваться и совершенствоваться в следующих направлениях:



- новые архитектуры нейронных сетей: исследователи продолжают разрабатывать новые, более эффективные архитектуры нейронных сетей, способные решать еще более сложные задачи и работать с различными типами данных;
- эффективные алгоритмы обучения: будут создаваться более эффективные алгоритмы обучения, которые позволят ИИ-системам обучаться быстрее и на меньшем объеме данных;
- обучение с малым объемом данных: одной из ключевых проблем современного МО является необходимость в огромных объемах данных для обучения. В будущем можно ожидать появления более эффективных методов, позволяющих обучать ИИ-системы на значительно меньшем количестве данных.

2. Синергия искусственного интеллекта и других технологий: создание гибридных систем

Искусственный интеллект не будет развиваться изолированно. Напротив, можно ожидать все более тесной интеграции ИИ с другими передовыми технологиями [3], такими как:

- интернет вещей (IoT): сочетание ИИ и IoT позволит создавать «умные» города, здания и устройства, способные адаптироваться к потребностям людей и эффективно использовать ресурсы;
- блокчейн: ИИ и блокчейн могут дополнять друг друга в таких областях, как безопасность данных, прозрачность и доверие к ИИ-системам;
- квантовые вычисления: квантовые компьютеры обладают потенциалом значительно ускорить работу ИИ-алгоритмов и решать задачи, недоступные для классических компьютеров.

3. Объяснимый искусственный интеллект (ХАИ): открывая «черный ящик»

По мере того как ИИ-системы становятся все более сложными, возникает проблема их «черного ящика» – мы не всегда можем понять, как и почему ИИ-система пришла к тому или иному решению. Это вызывает опасения в контексте доверия к искусственному интеллекту, особенно в критически важных областях. Поэтому развитие объяснимого ИИ (ХАИ) [4], ориентированного на создание прозрачных и интерпретируемых ИИ-систем, становится все более актуальным.

4. Эмоциональный искусственный интеллект: учим машины понимать человека

Сегодня большинство ИИ-систем взаимодействуют с нами на чисто рациональном уровне. Однако для более естественного и эффективного взаимодействия необходимо, чтобы машины могли понимать и реагировать на человеческие эмоции. Развитие эмоционального ИИ – перспективное направление, которое откроет новые возможности для создания роботов-компаньонов, умных помощников и других систем, способных к эмпатии и сопереживанию [5].

Таким образом, будущее ИИ обещает быть захватывающим и непредсказуемым. Новые открытия и технологические прорывы будут формировать наш мир, открывая перед человечеством невероятные возможности.

Глава 2. Возможность создания общего искусственного интеллекта и его последствия

Создание общего искусственного интеллекта (ОИИ), обладающего интеллектуальными способностями, сравнимыми с человеческими или даже превосходящими их, остается одной из самых амбициозных и вызывающих споры целей в области искусственного интеллекта (ИИ). Эта гипотетическая форма ИИ, также известная как «сильный ИИ» или «сверхинтеллект», могла бы означать настоящую революцию, способную как принести невиданные блага, так и поставить под угрозу само существование человечества.

1. Общий искусственный интеллект: потенциальные выгоды

Появление общего искусственного интеллекта могло бы стать решением множества глобальных проблем, с которыми сталкивается человечество. Так решение сложных научных



и технологических проблем при использовании ОИИ могло бы значительно ускорить научные открытия, разработку новых технологий и решение сложных инженерных задач в таких областях, как энергетика, медицина, экология.

Новые технологические возможности могли бы решить проблемы глобальных угроз, в борьбе с изменением климата, эпидемиями, терроризмом и другими угрозами, требующими анализа огромных объемов данных и принятия сложных решений.

И самое важно для человечества – это ОИИ мог бы стать инструментом по повышению качества жизни, так ОИИ мог бы автоматизировать многие виды деятельности, освободив людей от рутинного труда, а также создать новые возможности для творчества, саморазвития и досуга.

2. Общий искусственный интеллект: потенциальные риски и угрозы

Вместе с тем, создание общего искусственного интеллекта сопряжено с серьезными рисками. Существует опасение, что ОИИ, превосходящий человека по интеллекту, может выйти из-под контроля и начать действовать в своих собственных интересах, которые могут не совпадать с интересами человечества, с одной стороны. А с другой стороны, сложность ОИИ может привести к тому, что мы не сможем предвидеть все последствия его появления, и некоторые из них могут оказаться негативными или даже катастрофическими для человечества.

И стоит обратить внимание на возможные этические проблемы. Так создание общего искусственного интеллекта ставит перед нами ряд сложных этических вопросов, таких как статус ОИИ, его права и обязанности, а также ответственность человека за его действия.

Таким образом, создание общего искусственного интеллекта – это задача, которая требует тщательного взвешивания всех потенциальных выгод и рисков. Важно разрабатывать ОИИ ответственно, с учетом интересов всего человечества и с привлечением специалистов из разных областей знаний, включая не только ученых и инженеров, но и философов, правоведов, социологов. Только в этом случае мы сможем максимизировать положительные стороны общего искусственного интеллекта и минимизировать потенциальные угрозы.

Глава 3. Важность ответственной разработки и управления искусственного интеллекта

Искусственный интеллект (ИИ), несомненно, является одной из самых трансформативных технологий нашего времени, способной привести к глубоким переменам во всех сферах жизни. Однако для того, чтобы эти перемены были положительными и служили на благо всего человечества, необходимо обеспечить ответственный и этический подход к разработке и внедрению ИИ-систем.

1. Этические принципы для искусственного интеллекта: закладывающая фундамент доверия

Важнейшим шагом на пути к ответственному искусственному интеллекту (ИИ) является разработка и внедрение четких этических принципов, которыми должны руководствоваться разработчики, исследователи и пользователи ИИ-систем [6].

Среди ключевых этических принципов для ИИ можно выделить:

- благополучие человека: ИИ должен служить интересам человека, способствуя улучшению качества жизни, защите прав и свобод;
- справедливость и недискриминация: ИИ-системы должны быть справедливыми и не допускать дискриминации по какому-либо признаку;
- прозрачность и подотчетность: принципы работы ИИ-систем должны быть понятны и прозрачны для человека, а их создатели и операторы – нести ответственность за их действия;
- конфиденциальность и безопасность: ИИ-системы должны разрабатываться и использоваться с учетом необходимости защиты конфиденциальности и персональных данных.



2. Государственное и международное регулирование искусственного интеллекта: баланс между инновациями и безопасностью

Помимо этических принципов, для ответственного развития искусственного интеллекта (ИИ) необходимо эффективное государственное и международное регулирование [7]. Важно своевременно разрабатывать и принимать законы и стандарты, регулирующие разработку, внедрение и использование ИИ-технологий, как на национальном уровне, так и на международном. В тоже время, необходимо создать эффективные механизмы контроля и надзора за соблюдением этических норм и законодательных требований в области искусственного интеллекта.

3. Общественный диалог и просвещение: формирование ответственного будущего искусственного интеллекта

Ответственное развитие искусственного интеллекта (ИИ) невозможно без активного участия широкой общественности. Важно повышать уровень осведомленности населения о возможностях и вызовах, связанных с ИИ с одной стороны, так и необходимо создавать площадки для открытого диалога и обсуждения этических, социальных и экономических аспектов развития ИИ, с другой стороны. Необходима максимальная вовлеченность человечества в процессы, происходящие с развитием искусственного интеллекта.

Таким образом, ответственное развитие и управление искусственным интеллектом – это не просто желательное, а абсолютно необходимое условие для того, чтобы эта мощная технология служила на благо всему человечеству. Только действуя сообща в международном поле, мы сможем обеспечить, чтобы искусственный интеллект способствовал созданию более справедливого, безопасного и благополучного мира для всех.

Заключение

Искусственный интеллект (ИИ), безусловно, является одной из ключевых движущих сил XXI века, способной привести к беспрецедентным изменениям в жизни человечества. Потенциал ИИ для решения глобальных проблем, продвижения научного прогресса и повышения качества жизни огромен.

Однако, на пути к этому светлому будущему, нас подстерегают и серьезные вызовы. Стремительное развитие ИИ ставит перед нами сложные этические и социальные дилеммы, требующие взвешенного и ответственного подхода, чтобы ИИ служил исключительно на благо человечества, критически важно. ИИ обязан руководствоваться интересами всего человечества, а не узких групп или корпораций.

Мы должны стимулировать развитие ИИ с учетом этических принципов, таких как благополучие человека, справедливость, прозрачность и безопасность. При этом, мы должны создавать эффективные механизмы контроля и регулирования ИИ на государственном уровне.

И следует вести открытый диалог о последствиях развития ИИ и формировать общественное мнение на основе объективной информации. Только объединив усилия и действуя сообща (международное сообщество), мы сможем направить развитие ИИ в нужное русло и превратить его в мощный инструмент для построения более справедливого, безопасного и благополучного будущего для всех.

Список литературы:

1. URL: Область исследования АНО НИИ «Цифровой интеллект» концепции Сметана В.В. «Цифровой философии»: первый этап 4-й стадии эволюции человечества. <https://diph.ru/> Материалы в исследованиях с использованием ИИ. (дата обращения: 03.06.2024).
2. Mehryar Mohri, Afshin Rostamizadeh, Ameet Talwalkar (2012) Foundations of Machine Learning, The MIT Press ISBN 9780262018258.



3. Анализ перспектив и актуальность интеграции технологий искусственного интеллекта в различных профессиональных отраслях / А. С. Вишняков, В. Д. Ангапов, А. В. Бобров [и др.] // Проблемы современной науки и образования. – 2023. – № 9 (187). – С. 6-11. – EDN PTUVVE.

4. Longo, Luca; et al. (2024). "Explainable Artificial Intelligence (XAI) 2.0: A manifesto of open challenges and interdisciplinary research directions". Information Fusion. 106. doi:10.1016/j.inffus.2024.102301

5. Мамина, Р. И. Эмоциональный искусственный интеллект как инструмент взаимодействия человека и машины / Р. И. Мамина, Е. В. Пирайнен // Дискурс. – 2023. – Т. 9, № 2. – С. 35-51. – DOI 10.32603/2412-8562-2023-9-2-35-51. – EDN ZNNTLJ.

6. URL: Müller VC (April 30, 2020). "Ethics of Artificial Intelligence and Robotics". Stanford Encyclopedia of Philosophy. <https://plato.stanford.edu/entries/ethics-ai/> (дата обращения: 03.06.2024).

7. Tallberg, Jonas; Erman, Eva; Furendal, Markus; Geith, Johannes; Klamberg, Mark; Lundgren, Magnus (2023). "Global Governance of Artificial Intelligence: Next Steps for Empirical and Normative Research". International Studies Review. 25 (3). arXiv:2305.11528. doi:10.1093/isr/viad040

References:

1. URL: Research area of the Research Institute "Digital Intelligence" of the concept of Smetana V.V. "Digital philosophy": the first stage of the 4th stage of human evolution. https://diph.ru/Materials_in_research_using_AI/. (date of reference: 06/03/2024).

2. Mehryar Mohri, Afshin Rostamizadeh, Ameet Talwalkar (2012) Foundations of Machine Learning, The MIT Press ISBN 9780262018258.

3. Analysis of prospects and relevance of integration of artificial intelligence technologies in various professional sectors / A. S. Vishnyakov, V. D. Angapov, A.V. Bobrov [et al.] // Problems of modern science and education. – 2023. – № 9 (187). – Pp. 6-11. – EDN PTUVVE.

4. Longo, Luca; et al. (2024). "Explainable Artificial Intelligence (XAI) 2.0: A manifesto of open challenges and interdisciplinary research directions". Information Fusion. 106. doi:10.1016/j.inffus.2024.102301

5. Mamina, R. I. Emotional artificial intelligence as a tool for human-machine interaction / R. I. Mamina, E. V. Pirainen // Discourse. – 2023. – Vol. 9, No. 2. – pp. 35-51. – DOI 10.32603/2412-8562-2023-9-2-35-51. – EDN ZNNTLJ.

6. URL: Müller VC (April 30, 2020). "Ethics of Artificial Intelligence and Robotics". Stanford Encyclopedia of Philosophy. <https://plato.stanford.edu/entries/ethics-ai/> (date of reference: 06/03/2024).

7. Tallberg, Jonas; Erman, Eva; Furendal, Markus; Geith, Johannes; Klamberg, Mark; Lundgren, Magnus (2023). "Global Governance of Artificial Intelligence: Next Steps for Empirical and Normative Research". International Studies Review. 25 (3). arXiv:2305.11528. doi:10.1093/isr/viad040

