



ОСНОВЫ ЭТИЧЕСКИ ОРИЕНТИРОВАННОГО ИСПОЛЬЗОВАНИЯ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА ПРИ ПОДГОТОВКЕ СПЕЦИАЛИСТОВ

З.В. Пунчик, А.М. Зеневич

Белорусский государственный экономический университет, Беларусь

Аннотация. В статье исследуются признанные на международном уровне базовые инициативы в области использования этически ориентированного искусственного интеллекта, рассматриваются примеры положительного опыта подготовки лиц, принимающих решения в области систем искусственного интеллекта. Сделан вывод о необходимости изучения этических принципов искусственного интеллекта при профессиональной подготовке специалистов не только ИТ-сферы, но и других сфер, где профессиональные компетенции так или иначе связаны с принятием решений на основе использования современных технологий обработки информации. В частности, в статье предложены темы для включения в профильные учебные программы при подготовке специалистов в бакалавриате и магистратуре БГЭУ.

Ключевые слова. Искусственный интеллект; этика искусственного интеллекта; этические принципы искусственного интеллекта; стандарты искусственного интеллекта; рекомендации об изучении этики искусственного интеллекта.

FUNDAMENTALS OF ETHICALLY ORIENTED USE OF ARTIFICIAL INTELLIGENCE IN TRAINING SPECIALISTS

Z. Punchik, A. Zenevich

Belarusian State Economic University, Belarus

Abstract. The article examines internationally recognized basic initiatives in the use of ethically oriented artificial intelligence, examines examples of positive experience in training decision makers in the field of artificial intelligence systems. It is concluded that it is necessary to study the ethical principles of artificial intelligence in the professional training of specialists not only in the IT sphere, but also in other areas where professional competencies are somehow related to decision-making based on the use of modern information processing

technologies. In particular, the article suggests topics for inclusion in specialized training programs in the preparation of specialists in the bachelor's and master's degrees of the Belarusian State Economic University.

Keywords. Artificial intelligence; ethics of artificial intelligence; ethical principles of artificial intelligence; artificial intelligence standards; recommendations on the study of the ethics of artificial intelligence.

Введение

Искусственный интеллект (далее – ИИ), понимаемый как свойство компьютерных информационных систем выполнять творческие функции, которые традиционно считались прерогативой человека, в настоящее время бурно развивается, а технологии ИИ широко шагнули в повседневную жизнь. Несмотря на несомненные преимущества ИИ, его использование сопряжено с возникновением множества этических проблем, которые по мере расширения сфер применения приобретают все более глубокий и весомый по последствиям характер. По мнению экспертов в области разработки последствия применения ИИ настолько серьезны, что по масштабам влияния они сравнимы с распространением ядерных технологий и антропогенным влиянием на изменение климата, так как все они, оставленные без ограничений, способны негативно повлиять на весь мир [1].

Согласно определению, данному президентом Российской ассоциации искусственного интеллекта Осиповым Г.С. (1948 – 2020), наука под названием «искусственный интеллект» входит в комплекс компьютерных наук, а создаваемые на ее основе технологии относятся к информационным технологиям. Такой подход означает, что технологии ИИ развиваются по собственным законам [2]. В рамках этой парадигмы можно согласиться с утверждением российских исследователей о том, что существует две взаимосвязанные проблемы использования ИИ с точки зрения этики: согласование работы ИИ с существующими в обществе ценностными установками и формализация этих установок [3]. При этом, и та, и другая проблема являются весьма сложными, и принципиально не могут быть решены только специалистами в области информационных технологий. Урегулирование этих проблем требует активного привлечения представителей социально-гуманитарных наук, способных оценить последствия применения технологий ИИ для отдельных индивидов, стран и общества в целом. Однако сейчас в многочисленной русскоязычной литературе технологического характера по ИИ, предназначенной профессионалам в данной сфере, практически не рассматриваются этические проблемы ИИ и при их возникновении специалист

руководствуется лишь интуитивным пониманием этических норм. При профессиональной подготовке специалистов эта проблема остается также «за кадром». Создавшаяся ситуация обуславливает актуальность изучения этических аспектов ИИ при подготовке не только ИТ-профессионалов, но и всех специалистов, каким-либо образом имеющих дело с системами ИИ [4].

Теоретические основы

Актуальность, мультидисциплинарность, сложность и многогранность вопроса этики ИИ, обусловленная в том числе, спецификой национальных подходов и значительно отличающихся ценностей разных культур, привела к тому, что в 2010-е годы в мире началась активная коллективная работа в этом направлении. В результате такой работы создаются некоммерческие организации, различные подразделения на базе ведущих университетов мира, исследующие проблемы этики ИИ, постоянно появляются новые, разнообразные по форме инициативы: проекты этических кодексов различного уровня (универсальные, корпоративные, профессиональные), фреймворки, гайдлайны и другие.

Поскольку вопросы, касающиеся этически ориентированного ИИ, актуальны для всего мирового сообщества остро ощущается необходимость разработки основополагающего документа, который, с одной стороны, приемлем для всех стран, а с другой – позволяет каждой стране сохранять национальные особенности этических норм. В разработке таких документов за последние два-три года происходит переход от дискуссий и координации принципов применения ИИ к первым сбалансированным практическим шагам. Как результат – авторитетными международными организациями опубликовано несколько рамочных документов в сфере этики ИИ. Рассмотрим некоторые из них.

1. Глобальный нормативный документ «Рекомендация об этике искусственного интеллекта», определяющий базовые принципы и подходы к решению этических проблем применения систем ИИ во всем мире, принят ЮНЕСКО в ноябре 2021 года. В документе подчеркивается, что излагаемые принципы должны применяться при одновременном соблюдении соответствующих национальных, региональных и международных правовых норм, и определяются следующие этические принципы деятельности на всех этапах жизненного цикла систем ИИ: соразмерность и непричинение вреда; безопасность и защищенность; справедливость и отказ от дискриминации; устойчивость; право на неприкосновенность частной жизни и защита данных; подконтрольность и подчиненность человеку; прозрачность и объяснимость;

ответственность и подотчетность; осведомленность и грамотность; многостороннее и адаптивное управление и взаимодействие [5].

2. Стандарт ISO/IEC TR 24368, разработанный в 2022 году международной организацией по стандартизации (ISO), отмечает важность знания различными группами, и, в первую очередь, профессионалами ИТ, этических и социальных проблем, связанных с системами и приложениями ИИ¹. В качестве одного из основных источников возникновения этических проблем фиксируется недостаточное понимание социальных последствий внедрения технологий ИИ после их развертывания.

3. «Глобальная инициатива IEEE по этическим рекомендациям для искусственного интеллекта и автономных систем» («IEEE Global Initiative») сформулировала концепцию «этически ориентированного ИИ». Эта инициатива нацелена на разработку стандартов, адресованных, в первую очередь, профессионалам в области ИИ, и ориентирована на решение конкретных вопросов на стыке технологических и этических точек зрения². Подход IEEE заключается в том, что стандарты носят рекомендательный характер. В настоящее время «IEEE Global Initiative» осуществляет работу над группой специализированных стандартов, часть из которых разработана и доступна для ознакомления по подписке всем желающим:

- 7000-2021 - Процесс стандартной модели IEEE для решения этических проблем при проектировании системы;

- 7001-2021 - Стандарт IEEE прозрачности автономных систем;

- 7002-2022 - Стандарт IEEE для процесса конфиденциальности данных;

- 7005-2021 - Стандарт IEEE для прозрачного управления данными работодателя;

- 7007-2021 - Онтологический стандарт IEEE для этических систем робототехники и автоматизации;

- 7010-2020 - Рекомендуемая IEEE практика оценки влияния автономных и интеллектуальных систем на благополучие человека.

IEEE считает, что одна из главных проблем, препятствующих этически выверенному применению технологий ИИ, заключается в том, что в настоящее время при профессиональной подготовке специалистов изучение этических аспектов практически не предусматривается учебными планами и программами.

¹ ISO/IEC TR 24368:2022 Information technology – Artificial intelligence – Overview of ethical and societal concerns [Электронный ресурс]. URL: <https://www.iso.org/standard/78507.html?browse=tc> (дата обращения: 01.08.2023).

² The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems . [Электронный ресурс]. URL: <https://standards.ieee.org/industry-connections/ec/autonomous-systems/> – (дата обращения: 01.08.2023).

Анализ этих документов позволяет согласиться с мнением, высказанным российскими исследователями [6], о том, что многие формулировки в них для русскоговорящей (и тем более белорусско-говорящей) аудитории режут слух и выглядят неуклюже и непривычно. Такое изложение в основополагающих международных документах принципов этически ориентированного использования ИИ делает их самостоятельное изучение практически невозможным, что определяет необходимость включения соответствующих вопросов в учебные планы подготовки специалистов и при обучении преподавателей.

Результаты и их обсуждение

На рынке труда уже сейчас, а особенно в ближайшем будущем, будут востребованы те, кто достаточно глубоко понимает этические аспекты систем ИИ. Таким образом, при подготовке лиц, принимающих решения в области ИИ, необходимо предусмотреть изучение основных принципов и положительного опыта разработки и применения этически ориентированного ИИ.

Международный опыт такой подготовки уже существует. К примеру, с 2021 года в Кембриджском университете по программе AI Ethics and Society осуществляется подготовка магистров по исследованиям в области этики ИИ. Учебная программа – по заявлению ее разработчиков – является уникально междисциплинарной, основанной на последних научных разработках, что позволяет сформировать у обучающихся критические навыки, знания и аналитические умения, необходимые для выявления и решения этических проблем, возникающих на практике в результате применения ИИ.

Лундский университет на платформе онлайн-образования coursera.org предоставляет бесплатный онлайн курс «Artificial Intelligence: Ethics & Societal Challenges», позволяющий осмыслить этические дилеммы, связанные с ИИ.

В рамках реализации европейского «Плана по цифровизации образования 2021-2027» в 2022 году разработан документ «Этические рекомендации по использованию искусственного интеллекта и данных в преподавании и обучении для преподавателей»³. В документе рассматривается четыре направления использования ИИ и данных в образовательном процессе. Первое – использование ИИ для обучения студентов, например, интеллектуальное обучение без вмешательства преподавателя, изучение языка с использованием обучающих приложений на основе искусственного интеллекта. Второе направление – использование ИИ для поддержки обучения студентов,

³ European Commission, Directorate-General for Education, Youth, Sport and Culture, Ethical guidelines on the use of artificial intelligence (AI) and data in teaching and learning for educators, Publications Office of the European Union, 2022, [Электронный ресурс]. URL: <https://data.europa.eu/doi/10.2766/153756>– (дата обращения: 25.10.2023).

например, в рамках группового обучения системы ИИ комментируют или дают рекомендации о том, как группа работает вместе, отслеживая уровень взаимодействия между членами группы. Третье – использование ИИ для поддержки преподавателей, например, использование систем ИИ для автоматической оценки письменных работ обучающихся, агентов ИИ (чат-ботов) для ответов на часто задаваемые вопросы обучающихся. Четвертое направление – использование ИИ для поддержки диагностики и планирования в образовательной системе, например, интеллектуальный анализ образовательных данных. По каждому из направлений даются практические рекомендации по использованию ИИ и данных в педагогической практике.

Также интересен опыт Волгоградского государственного технического университета по разработке программ бакалавриата и программ магистратуры для направлений подготовки 09.03.01 и 09.04.01 «Информатика и вычислительная техника» а также на повышение квалификации педагогических работников образовательных организаций высшего образования в сфере искусственного интеллекта⁴. В частности, образовательная программа магистратуры «Системы искусственного интеллекта в топливно-энергетическом комплексе» по профилям «Искусственный интеллект в цифровой экономике», «Анализ данных и интеллектуальные технологии в ТЭК», включает теоретическую и практическую подготовку по разработке и эксплуатации систем ИИ и технологий бизнес-аналитики предприятия, направленную на формирование компетенции в области разработки систем этически ориентированного ИИ.

В Республике Беларусь вопрос применения принципов этически ориентированного ИИ находится в стадии обсуждения. В настоящее время признается необходимость более полного и систематического включения этики ИИ в учебные программы вузов, колледжей и различных курсов. Для координации усилий в роли регулятора разработки и использования систем ИИ в республике в ближайшем будущем планирует выступать Национальная академия наук. Она, в качестве практического шага реализации этого плана, с 2022 года организует форум IT-Академграда «Искусственный интеллект в Беларуси», в рамках которого широко обсуждаются вопросы этики ИИ.

Анализ зарубежных учебных планов подготовки специалистов по ИИ, представленных в сети Интернет, показывает, что многие из них уже содержат

⁴ ИИ-2021 [Электронный ресурс]. URL: [https:// www.vstu.ru/obrazovanie/ii2021](https://www.vstu.ru/obrazovanie/ii2021) (дата обращения: 25.10.2023)

требование формирования компетенции – уметь объяснить этические и социальные аспекты применения ИИ в конкретных случаях и оценить их влияние на практике.

На взгляд авторов, с учетом проведенного анализа общепризнанных норм этически ориентированного использования ИИ, имеющегося мирового опыта и с учетом индивидуальной цифровой компетентности преподавателей целесообразно внести следующие предложения по изучению теоретических основ этического применения ИИ при подготовке специалистов в Белорусском государственном экономическом университете (БГЭУ):

- при изучении студентами дисциплины государственного компонента «Информационные технологии», рассмотреть вопрос «Этичное применение ИИ в современном мире»;

- для студентов специальности «Экономическая информатика» и магистрантов специальности «Экономика» рассмотреть ряд теоретических вопросов (представлены в таблице).

Таблица – Примерный перечень изучаемых вопросов этически ориентированного использования ИИ для специальности Экономическая информатика

Table – An approximate list of the studied issues of ethically oriented use of AI for the specialty of Economic Informatics

№	Перечень вопросов	Краткое содержание	Уровень/специальность	Учебная дисциплина
1.	Этичное применение ИИ в современном мире	Этические проблемы, связанные с применением систем ИИ: ответственность ИИ, прозрачность ИИ; предвзятость алгоритмов; надежность ИИ; приватность при применении технологий ИИ.	Бакалавриат/ Экономическая информатика	Бизнес-аналитика
2	Регулирование ИИ в мировой практике	Национальные стратегии ИИ, принципы применения ИИ в конкретных сферах	Магистратура/ Экономика	Технологии интеллектуального анализа данных
3.	Стандарты на разработку этического ИИ	Стандарты, регулирующие этически обусловленное проектирование интеллектуальных и автономных систем (стандарты IEEE)	Бакалавриат/ Экономическая информатика	Бизнес-аналитика
4.	Оценка воздействия алгоритмических систем на права человека	Подходы к оценке воздействия алгоритмических систем на права человека		Системы искусственного интеллекта

Формирование компетенции этически ориентированного использования ИИ является актуальной задачей при подготовке специалистов. Изучение базовых теоретических основ, учитывающих конкретный контекст образования и профессиональной подготовки, обеспечивают основу формирования такой компетенции, необходимость в которой постоянно возрастает не только в профессиональной деятельности, но и в повседневной жизни.

Список использованных источников (на языке оригинала)

1. AI & Ethics: The Grand Challenge for Our Generation [Электронный ресурс]. URL: <https://www.ibm.com/blogs/ibm-canada/2021/05/ai-ethics-the-grand-challenge-for-our-generation> (дата обращения: 01.08.2023).
2. Осипов Г.С. Искусственный интеллект: состояние исследований и взгляд в будущее [Электронный ресурс]. URL: <http://www.raai.org/about/persons/osipov/pages/ai/ai.html> (дата обращения: 01.08.2023).
3. Гаспарян Д.Э., Стырин Е.М. Прикладные проблемы внедрения этики искусственного интеллекта в России: отраслевой анализ и судебная система; Нац. исслед. ун-т «Высшая школа экономики». М.: Изд. дом Высшей школы экономики, 2020. 16 с.
4. Пунчик З.В., Зеневич А.М. Этические аспекты работы с цифровыми данными. *Экономический рост Республики Беларусь: глобализация, инновационность, устойчивость*: материалы XV Междунар. науч.-практ. конф. (Минск, 19-20 мая 2022 г.). Минск: БГЭУ, 2022. С. 413-414.
5. Пунчик З.В., Зеневич А.М. Этически ориентированное использование искусственного интеллекта как необходимая компетенция подготовки специалистов ИТ-сферы. *Научные труды Белорусского государственного экономического университета*. Минск: БГЭУ, 2023. Вып. 16. С. 363-370.
6. Леушина В.В., Карпов В.Э. Этика искусственного интеллекта. *Философия и общество*. 2022;3(104): 135.

References (на английском языке)

1. AI & Ethics: The Grand Challenge for Our Generation [Electronic resource]. Available at: <https://www.ibm.com/blogs/ibm-canada/2021/05/ai-ethics-the-grand-challenge-for-our-generation>. Access date: 01.08.2023.

2. Osipov G.S. Artificial intelligence: the state of research and a look into the future [Electronic resource]. Available at: <http://www.raai.org/about/persons/osipov/pages/ai/ai.html> Access date: 01.08.2023. (In Russian)
3. Gasparyan D. E., Styrin, E. M. Applied problems of implementing the ethics of artificial intelligence in Russia: industry analysis and the judicial system. Nats. research. un-t "Higher School of Economics". M.: Publishing House of the Higher School of Economics, 2020. 16 p. (In Russian)
4. Punchik Z.V., Zenevich A.M. Ethical aspects of working with digital data. *Economic growth of the Republic of Belarus: Globalization, Innovation, sustainability: materials of the XV International Scientific and Practical Conference* (Minsk, May 19-20, 2022). Minsk: BSEU, 2022. P. 413-414. (In Russian)
5. 4. Punchik Z.V., Zenevich A.M. Ethically oriented use of artificial intelligence as a necessary competence for training IT specialists. *Scientific works of the Belarusian State University of Economics*. Minsk: BSEU, 2023. Vol. 16. P. 363-370. (In Russian)
6. Leushina V.V., Karpov V.E. Ethics of artificial intelligence. 2022;3(104):135. (In Russian)