

ЦИФРОВЫЕ ПОМОЩНИКИ: ЭТИЧЕСКИЕ ПРОБЛЕМЫ

Цифровые помощники становятся частью окружающей нас действительности дома, в автомобиле, на работе. Предполагается, что в скором времени цифровых помощников станет больше, чем людей. Иногда мы можем даже не догадываться, что агент по обслуживанию клиентов, с которым мы разговаривали, является цифровым помощником. Цифровые помощники собирают информацию о наших взаимодействиях и многое узнают о нас. В статье анализируются этические принципы прозрачности, справедливости, непричинения вреда, ответственности и подотчетности, конфиденциальности для цифровых помощников. Акцентируется внимание на том, какие этические соображения должны учитывать разработчики и пользователи при разработке цифровых помощников и взаимодействии с ними. Предлагаются практические рекомендации, позволяющие понять этические последствия создания и функционирования цифровых помощников.

Ключевые слова:

искусственный интеллект, цифровая этика, цифровые помощники, этика.

Шляпников В.В. Цифровые помощники: этические проблемы // Общество. Среда. Развитие. – 2022, № 2. – С. 78–80. – DOI 10.53115/19975996_2022_02_078–080

© Шляпников Виктор Валерьевич – кандидат философских наук, доцент, Санкт–Петербургский университет ГПС МЧС России, Санкт–Петербург; e–mail: shlyapnikovv@mail.ru

Многие компании и организации быстрыми темпами внедряют цифровых помощников для автоматизации коммуникации, обслуживания клиентов, рекомендаций по продуктам и услугам. По некоторым оценкам мировой рынок цифровых помощников вырастет с 16,8 млрд долларов в 2021 году до 43,8 млрд долларов к 2026 году [9]. По другим прогнозам, количество используемых цифровых помощников к 2023 году достигнет 8 млрд устройств по всему миру [7]. Цифровые помощники предлагают круглосуточное обслуживание и удовлетворяют потребности клиентов в ответах в режиме реального времени. Они позволяют компаниям и организациям сократить расходы, повысить лояльность к бренду и лучше обслуживать клиентов. Однако цифровые помощники создаются людьми с применением технологий искусственного интеллекта, и их использование имеет широкий спектр этических последствий. В статье анализируются этические принципы, содержащие ключевые этические соображения, необходимые для понимания разработчиками цифровых помощников и их пользователями.

Цифровые помощники – это интеллектуальные приложения, которые могут вы-

полнять задачи или предоставлять услуги пользователю, реагируя на его заказы или запросы. Многие цифровые помощники могут понимать человеческую речь и реагировать на нее с помощью синтезированных голосов. Пользователи могут использовать голосовые команды, чтобы попросить своего цифрового помощника ответить на вопросы, управлять бытовой техникой или воспроизведением мультимедиа, выполнять другие важные действия, например, создавать списки дел и планировать встречи в календарях. В сфере интернета вещей цифровые помощники – популярный сервис для общения с пользователями на основе голосовых команд [16].

Этические проблемы, связанные с различными аспектами использования цифровых помощников, в последнее время находятся в центре внимания многих исследований [2; 3; 8; 11; 13; 17; 18]. В статье о глобальном ландшафте руководящих принципов этики искусственного интеллекта [10] авторы обнаружили пять этических принципов, которые считаются важными во всем мире: прозрачность, справедливость, непричинение вреда, ответственность и подотчетность, конфиденциальность. Эти

принципы можно интерпретировать и в контексте цифровых помощников, акцентируя внимание на том, какие соображения должны учитывать разработчики и пользователи при разработке цифровых помощников и взаимодействии с ними [11].

Прозрачность. При работе с цифровыми помощниками прозрачность относится к информированию потребителей о том, с кем они общаются (то есть с цифровым помощником, а не с реальным человеком), и о том, как данные потребителей будут храниться, анализироваться и использоваться [14].

Компании и организации завоевывают доверие потребителей, будучи прозрачными и честными в общении. Если цифровой помощник выдает себя за человека, это может привести к плохому опыту и недоверию к компании или организации. Особенно вредным это может быть при взаимодействии с потребителями в таких деликатных сферах, как здравоохранение или финансы.

Кроме того, потребители также должны иметь возможность связаться с реальным человеком, запросить обратный звонок или отправить электронное письмо, если им неудобно взаимодействовать с цифровым помощником.

Справедливость. В случае с цифровыми помощниками справедливость относится прежде всего к приоритету интересов потребителя и предоставлению беспристрастных рекомендаций.

Разработчики должны регулярно тестировать цифровых помощников на соответствие показателям справедливости, публиковать результаты этих тестов для потребителей, а также предоставлять потребителям возможность оставлять отзывы о процессе общения с цифровыми помощниками. Чем лучше цифровой помощник адаптируется к интересам потребителей и дает справедливые рекомендации, тем более популярным он станет среди потребителей.

Непричинение вреда. Цифровые помощники часто являются первыми, с кем общаются потребители при поиске информации об определенном товаре или услуге, и, если их ответы не помогают, это приводит к разочарованию и оскорблению со стороны потребителей.

Разработчикам необходимо проектировать процесс общения с учетом того, что цифровые помощники могут подвергнуться оскорблениям. Разговор должен строиться таким образом, чтобы потребители получали положительный ответ, а если возникнет такая необходимость, по запросу переводились на более полезный канал, например, общение с реальным человеком или электронную почту [15].

Еще одним важным аспектом является пол цифрового помощника. Во многих случаях цифровые помощники по умолчанию имеют женский голос или образ. Как показывают многочисленные исследования [1; 4–6; 12], это может способствовать укреплению гендерных стереотипов и формированию отношения потребителя к конкретной ситуации. Разработчики могут предоставить потребителям возможность выбора персонажа цифрового помощника и изменения языка, голоса, тона ответов, характерных для выбранного персонажа.

Ответственность и подотчетность. Это означает раскрытие потребителям информации о юридической ответственности и праве собственности на данные.

Возникает много споров о том, кто в конечном итоге несет ответственность – цифровые помощники на основе искусственного интеллекта или люди, которые их создали. Как правило, условия соглашения об оказании услуг, с которыми потребители должны согласиться перед использованием цифровых помощников, определяют ограничения, обязанности и ответственность в соответствии с нормативными актами.

Потребители должны иметь возможность выбора: поделиться своими данными или сохранить их в тайне. Если бизнес-модель цифрового помощника основывается на предоставлении бесплатных услуг и использовании потребительских данных для рекламы, то это также должно быть прозрачно для потребителя.

Конфиденциальность. Конфиденциальность означает, что ваша личная информация хранится в тайне и передается только с вашего согласия. Многие страны приняли законы и правила для защиты конфиденциальности персональных данных. Что касается цифровых помощников, конфиденциальность часто упоминается в связи со следующими вопросами:

- У кого имеется доступ к записям разговоров?
- Используются ли записи для профилирования потребителей?
- Передаются ли записи рекламодателям?
- Анонимизируются ли данные потребителей перед тем, как поделиться ими?
- Используются ли записи для совершенствования программных моделей искусственного интеллекта?
- Где хранятся записи?
- В течение какого времени хранятся записи?
- Может ли потребитель удалить записи?

– Защищен ли канал связи?

Разработчики должны быть прозрачными в отношении политики конфиденциальности и публиковать ее в открытом доступе, чтобы потребители могли быть проинформированы о том, как хранится и защищается их информация. Это также поможет укрепить доверие к цифровым помощникам, и потребители будут охотнее делиться информацией, если будут знать, что их будут обслуживать лучше.

Можно констатировать, что по мере того как цифровые помощники становятся частью нашей повседневной жизни, связанные с ними этические проблемы будут нарастать. Для общества в целом важно обсудить и согласовать этические принципы прозрачности, справедливости, непричинения вреда,

ответственности и подотчетности, конфиденциальности для цифровых помощников.

Этические принципы обычно основаны на хорошо обоснованных стандартах правильного и неправильного, которые предписывают, что должно делать с точки зрения прав, обязанностей, пользы для общества, справедливости или конкретных добродетелей. Цифровые помощники необходимо разрабатывать с учетом этических стандартов общества. Эти стандарты должны распространяться и на разработчиков, которые должны гарантировать, что взаимодействие цифровых помощников с пользователями будет способствовать честности, воздержанию от мошенничества или причинения вреда и обеспечивать право на неприкосновенность частной жизни.

Список литературы:

- [1] Abercrombie G., Curry A., Pandya M., Rieser V. Alexa, Google, Siri: What are Your Pronouns? Gender and Anthropomorphism in the Design and Perception of Conversational Assistants // Proceedings of the 3rd Workshop on Gender Bias in Natural Language Processing, Online, August 5, 2021, Association for Computational Linguistics. – 2021. – P. 24–33. – DOI: <https://doi.org/10.18653/v1/2021.gebnlp-1.4>
- [2] Acikgoz F., Perez-Vega R. The Role of Privacy Cynicism in Consumer Habits with Voice Assistants: A Technology Acceptance Model Perspective // International Journal of Human-Computer Interaction. – 2021. – P. 1–15. – DOI: <https://doi.org/10.1080/10447318.2021.1987677>
- [3] Bolton T., Dargahi T., Belguith S., Al-Rakhami M., Sodhro A. On the Security and Privacy Challenges of Virtual Assistants // Sensors. Vol. 21. – 2021, № 7. – 2312. – DOI: <https://doi.org/10.3390/s21072312>
- [4] Chin C., Robison M. How AI bots and voice assistants reinforce gender bias // Center for Technology Innovation – The Brookings Institution. – Интернет-ресурс. Режим доступа: <https://www.brookings.edu/research/how-ai-bots-and-voice-assistants-reinforce-gender-bias/>
- [5] Curry A., Rieser V. #MeToo Alexa: How Conversational Systems Respond to Sexual Harassment // Proceedings of the Second ACL Workshop on Ethics in Natural Language Processing, New Orleans, Louisiana, USA, June, 2018, Association for Computational Linguistics. – 2018. – P. 7–14. – DOI: <https://doi.org/10.18653/v1/W18-0802>
- [6] Curry A., Robertson J., Rieser V. Conversational Assistants and Gender Stereotypes: Public Perceptions and Considerations for Voice Personas // Proceedings of the Second Workshop on Gender Bias in Natural Language Processing, Barcelona, Spain (Online), December 13, 2020, Association for Computational Linguistics. – 2020. – P. 72–78
- [7] Digital Voice Assistants in Use to Triple to 8 Billion by 2023, Driven by Smart Home Devices // Juniper Research. – Интернет-ресурс. Режим доступа: <https://www.juniperresearch.com/press/digital-voice-assistants-in-use-to-8-million-2023>
- [8] Heinrichs J-H. Why Digital Assistants Need Your Information to Support Your Autonomy // Philosophy & Technology. – 2021. – № 34. – P. 1687–1705. – DOI: <https://doi.org/10.1007/s13347-021-00481-4>
- [9] Intelligent Virtual Assistant: Global Markets to 2026 // Research and Markets. – Интернет-ресурс. Режим доступа: <https://www.researchandmarkets.com/reports/5511337/intelligent-virtual-assistant-global-markets-to#rela0-4845914>
- [10] Jobin A., Ienca M., Vayena, E. The global landscape of AI ethics guidelines // Nature Machine Intelligence. – 2019, № 1. – P. 389–399. – DOI: <https://doi.org/10.1038/s42256-019-0088-2>
- [11] Kaul A. Virtual Assistants and Ethical Implications // Virtual Assistant (Soofastaei A., Ed.). – London, United Kingdom: IntechOpen, 2021. – P. 33–44. – DOI: <https://doi.org/10.5772/intechopen.95479>
- [12] Nass C., Moon Y., Green N. Are Machines Gender Neutral? Gender-Stereotypic Responses to Computers With Voices // Journal of Applied Social Psychology. Vol. 27. – 1997, № 10. – P. 864–876. – DOI: <https://doi.org/10.1111/j.1559-1816.1997.tb00275.x>
- [13] Porra J., Lacity M., Parks M. «Can Computer Based Human-Likeness Endanger Humanness?» – A Philosophical and Ethical Perspective on Digital Assistants Expressing Feelings They Can't Have // Information Systems Frontiers. – 2020, № 22. – P. 533–547. – DOI: <https://doi.org/10.1007/s10796-019-09969-z>
- [14] Shanbhag A. 5 Chatbot Code Of Ethics Every Business Should Follow // BotCore. – Интернет-ресурс. Режим доступа: <https://botcore.ai/blog/5-chatbot-code-of-ethics-every-business-should-follow/>
- [15] Shevat A. Hard questions about bot ethics // Medium. – Интернет-ресурс. Режим доступа: <https://medium.com/slack-developer-blog/hard-questions-about-bot-ethics-4f80797e34f0>
- [16] Soofastaei A. Introductory Chapter: Virtual Assistants // Virtual Assistant (Soofastaei A., Ed.). – London, United Kingdom: IntechOpen, 2021. – P. 1–9. – DOI: <https://doi.org/10.5772/intechopen.100248>
- [17] Suarez B. The Ethics of Digital Voice Assistants // Viterbi Conversations in Ethics. Vol. 5. – 2021, № 1. – Интернет-ресурс. Режим доступа: <https://vce.usc.edu/volume-5-issue-1/the-ethics-of-digital-voice-assistants/>
- [18] Wilson R., Ifimie I. Virtual assistants and privacy: An anticipatory ethical analysis // 2021 IEEE International Symposium on Technology and Society (ISTAS). – 2021. – P. 1–1. – DOI: <https://doi.org/10.1109/ISTAS52410.2021.9629164>