

Какими нас делает Искусственный Интеллект: беззаботными или безработными?

Е. Д. Садовская¹, Ф. Н. Винокуров²

В данной работе обсуждаются результаты эксплораторного этапа (первого из трех) исследования социальных представлений молодежи об Искусственном интеллекте (ИИ). Наша работа основана на теории социальных представлений Московиси[11]. Были проведены интервью ($N = 14$), наблюдения ($N = 14$) и собраны свободные ассоциации на тему ИИ ($N = 127$). Цель данного этапа - обозначить карту общих значений, связанных с ИИ. Наиболее ярко выраженными оказались темы, описанные ниже.

- 1) Представления о будущем с ИИ: одновременно сосуществуют сценарии катастрофы с безработицей и сценарии свободного, творческого и беззаботного будущего для человека.
- 2) Размышления о “не-человечности” ИИ.
- 3) Деление сфер применения на допустимые для использования ИИ и недопустимые.

Следующими шагами данного исследования станут анализ медиа на выявление новых и подтверждение проявившихся тем; анализ позиционирования этих тем.

Ключевые слова: Искусственный интеллект, Взаимодействие Человек-ИИ, Теория социальных представлений

1. Введение

Внедрение технологии Искусственного Интеллекта (ИИ) в ежедневную рутину происходит стремительно. ИИ используется в сферах транспорта[8], медицины[9], развлечений, психотерапии[10]. Теперь, когда ИИ так сильно проник в жизни обычных людей, его активно начали рассматривать и сквозь призму психологической науки. В работах на стыке психологии и ИИ появляются новые направления. *Human-AI*

¹ Садовская Екатерина Дмитриевна — аспирант каф. социальной психологии, факультета психологии МГУ имени М.В. Ломоносова, e-mail: ed.sadovskaya@gmail.com.

Sadovskaya Ekaterina Dmitrievna — graduate student, Lomonosov Moscow State University, Faculty of Psychology, Chair of Social Psychology.

² Винокуров Федор Никитич — к.п.н., ст.н.с. Кафедры Социальной психологии, факультета психологии МГУ имени М.В. Ломоносова, e-mail: VinokurovFN@my.msu.ru .

Vinokurov Fyodor Nikitich — PhD, Senior Researcher, Lomonosov Moscow State University, Faculty of Psychology, Chair of Social Psychology.

Interaction[4], *Отношение* к ИИ у россиян, перспективы психологических исследований. Технофобия и Технофилия и ее предикторы[2, 3], возможности применения технологии ИИ для диагностики в клинической психологии и психиатрии[7]. Стереотипы и установки, отраженные в данных, на основе которых обучается ИИ[5]

В нашей работе мы рассматриваем ИИ в контексте социальной психологии, как объект социального представления. ИИ уже не новая технология, но именно сейчас, в эпоху Big Data, он появляется в центре повседневной социальной жизни человека и порождает социальные представления.

В рамках Летней Психологической Школы МГУ 2021 — “Мастерская UX-исследований в киберпсихологии” проектной группой “Доверие Искусственному интеллекту (ИИ)” был проведен эксплораторный этап (первый из трех этапов) анализа социальных представлений на тему ИИ у молодежи, основанный на теории социальных представлений[11].

2. Метод

Следуя стратегии исследования социальных представлений в Женевской школе[6] на первом этапе важно отобразить дискуссию, где присутствует вышуклость и попытаться реконструировать общее поле этой дискуссии[1].

Наше исследование проходило в три этапа:

- 1) Сбор свободных ассоциаций на “Искусственный интеллект”.
- 2) Интервью на тему восприятия ИИ
- 3) Наблюдение в ситуации решения задач во взаимодействии с технологией, основанной на ИИ. Конкретной технологией в данном этапе был Голосовой помощник(ГП): Алиса (Яндекс), Маруся (Mail.ru) и Siri (Apple). Данный выбор обусловлен распространенностью знания о ГП и доступностью в использовании (не требуется специальных знаний и оплаты). В рамках наблюдения фиксировались реакции и рассуждения на тему взаимодействия с ГП и выполнения с помощью него определенного набора команд: поиск новостей, управление музыкой, уточнение вопроса, попытка диалога и игра.

Выборка на первом этапе выборка составила 127 человек, 80% женщин, 20% мужчин. Возраст от 16 до 29. На втором и третьем - 14 человек. К участию привлекались по большей части студенты психологи в возрасте от 18 до 26 лет. 11 женщин и 3 мужчины.

3. Результаты

В ходе интервью и наблюдения были выделены следующие темы, затрагиваемые респондентами:

- соотношение человек и ИИ, может ли ИИ имитировать эмпатию и эмоции;
- разделение сфер применения технологии ИИ на те, в которых респондент готов к взаимодействию с ним и те, в которых не готов;
- представление будущего в двух полярных сценариях при одних и тех же вводных. На примере работы:
 - ИИ захватывает человечество, отнимает работу, оставляет людей без профессий
 - ИИ служит верным помощником, берет на себя рутинные задачи и дает пространство для творчества человечества

Таблица 1. Топ-10 слов-ассоциаций на словосочетание “Искусственный интеллект”

Слово	Частота употребления
умный	19
современный	17
компьютер	16
робот	15
будущее	14
сложный	13
опасный	11
быстрый	11
компьютерный	11
роботизированный	10

В ходе анализа частотности свободных ассоциации всего было собрано 349 слов-ассоциаций (встретились один и более раз при подсчетах). Наиболее часто встречались “умный”, “современный”, “компьютер”, “робот” и “будущее” (19, 17, 16, 15 и 14 раз соответственно). В целом, наблюдается очень большой разброс вариативности ассоциаций. В ассоциациях так же, как и в интервью всплывает тема невозможности ИИ приблизится к человечности. ИИ - робот, неживой, холодный, компьютерный, технический, искусственный, опасный. Причем некоторые слова данной категории могут нести отталкивающую коннотацию по аналогии

живой-мертвый: холодный, неживой, опасный. Отдельная намечающаяся категория ассоциаций - наполнение, из чего он состоит: нейросети, программирование. Другая наметившаяся в ассоциациях тема - способности и: умный, быстрый, логичный.

В ассоциациях также можно заметить похожий на результаты интервью тренд на два полюса: страшное/пугающее и полезное/помогающее. Одновременно встречаются ассоциации с положительным отношением и с отрицательным. Что говорит нам о той же социокогнитивной выпуклости термина. В текущих реалиях нас посещают фантазии о том, каким будет будущее с ИИ - беззаботным или безработным?

4. Выводы

Целью первого эксплораторного этапа нашего исследования было собрать карту общих значений, связанных с ИИ. Были обозначены следующие основные темы: представления о будущем с ИИ, его “не-человечности” и способностях, делении сфер на допустимые для использования ИИ и недопустимые.

Следующими шагами данного исследования планируются: анализ медиа на выявление новых и подтверждение проявившихся тем; анализ позиционирования этих тем.

Список литературы

- [1] Бовина И. Б., “Стратегии исследования социальных представлений”, *Социологический журнал*, 2011, №3.
- [2] Нестик Т. А., “Внедрение технологий искусственного интеллекта в повседневную жизнь: перспективы исследований”, *Цифровое общество в культурно-исторической парадигме*, ред. Т.Д. Марцинковская, В.Р. Орестовая, О.В. Гавриченко, РГГУ, 2018, 29–35.
- [3] Солдатова Г. У., Нестик Т. А., “Отношение к интернету среди интернет-пользователей: технофобы и технофилы”, *Вестник Московского государственного областного университета*, 2016, №1, Серия: Психологические науки, 54–61.
- [4] Arrieta A. B. et al., “Explainable Artificial Intelligence (XAI)”, *Information Fusion*, **58** (2020), 85–115.
- [5] Caliskan A., Bryson J. J., Narayanan A., “Semantics derived automatically from language corpora contain human-like biases”, *Science*, **356**:6334 (2017), 183–186.
- [6] Clémence A., *Les méthodes des sciences humaines*, Sous la dir S. Moscovici, F. Buschini, Presses Universitaires de France, Paris, 2003.
- [7] Dwyer D. B., Falkai P., Koutsouleris N., “Machine learning approaches for clinical psychology and psychiatry”, *Annual review of clinical psychology*, **14** (2018), 91–118.

- [8] Fridman L. et al, *Mit autonomous vehicle technology study: Large-scale deep learning based analysis of driver behavior and interaction with automation*, **1**, 2017, arXiv: 1711.06976.
- [9] Longoni C., Bonezzi A., Morewedge C. K., “Resistance to medical artificial intelligence”, *Journal of Consumer Research*, **46**:4 (2019), 629–650.
- [10] Lucas G. M. et al., “It’s only a computer: Virtual humans increase willingness to disclose”, *Computers in Human Behavior*, **37** (2014), 94–100.
- [11] Moscovici S., *La Psychanalyse: Son image et son public*, Presses Universitaires de France, Paris, 1961.

Future with AI: careless or workless **Sadovskaya E.D., Vinokurov F.N.**

This paper discusses the results of the exploratory stage (the first of three) of the study of social social perceptions of young people about Artificial Intelligence (AI). Our work is based on Moskovisi’s theory of social perceptions[11]. Interviews ($N = 14$) and observations ($N = 14$) were conducted, and free associations were collected on the topic of AI ($N = 127$). The purpose of this step is to map the common meanings associated with AI. The most prominent were the topics described below.

- 1) Ideas about the future with AI: the scenarios of a catastrophe with unemployment and free, creative and carefree future for humans coexist at the same time.
- 2) Reflections on the “non-humanity” of AI
- 3) The division of the areas of application of the use of AI into acceptable and unacceptable.

The next steps of this study will be the analysis of the media to identify new topics, confirmation and analysis of the positioning of these topics.

Keywords: Artificial Intelligence, Human-AI interaction, Social representations

References

- [1] Bovina I. B., “Strategii issledovaniya social’nyh predstavlenij”, *Sociologicheskij zhurnal*, 2011, №3 (in Russian).
- [2] Nestik T. A., “Vnedrenie tekhnologij iskusstvennogo intellekta v povsednevnyu zhizn’: perspektivy issledovaniy”, *Cifrovoe obshchestvo v kul’turno-istoricheskoy paradigme*, eds. T.D. Marcinkovskaya, V.R. Orestovaya, O.V. Gavrichenko, Russian State University for the Humanities, 2018, 29–35 (in Russian).

- [3] Soldatova G. U., Nestik T. A., “Otnoshenie k internetu sredi internet-pol’zovatelej: tekhnofoby i tekhnofily”, *Vestnik Moskovskogo gosudarstvennogo oblastnogo universiteta*, 2016, № 1, Seriya: Psihologicheskie nauki, 54–61 (in Russian).
- [4] Arrieta A. B. et al., “Explainable Artificial Intelligence (XAI)”, *Information Fusion*, **58** (2020), 85–115.
- [5] Caliskan A., Bryson J. J., Narayanan A., “Semantics derived automatically from language corpora contain human-like biases”, *Science*, **356**:6334 (2017), 183–186.
- [6] Clémence A., *Les méthodes des sciences humaines*, Sous la dir S. Moscovici, F. Buschini, Presses Universitaires de France, Paris, 2003.
- [7] Dwyer D. B., Falkai P., Koutsouleris N., “Machine learning approaches for clinical psychology and psychiatry”, *Annual review of clinical psychology*, **14** (2018), 91–118.
- [8] Fridman L. et al, *Mit autonomous vehicle technology study: Large-scale deep learning based analysis of driver behavior and interaction with automation*, **1**, 2017, arXiv: 1711.06976.
- [9] Longoni C., Bonezzi A., Morewedge C. K., “Resistance to medical artificial intelligence”, *Journal of Consumer Research*, **46**:4 (2019), 629–650.
- [10] Lucas G. M. et al., “It’s only a computer: Virtual humans increase willingness to disclose”, *Computers in Human Behavior*, **37** (2014), 94–100.
- [11] Moscovici S., *La Psychanalyse: Son image et son public*, Presses Universitaires de France, Paris, 1961.