

Секция «1.1 Цифровая трансформация и искусственный интеллект в государственном администрировании 3.0: от автоматизации к принятию интеллектуальных решений»

**Нейрокогнитивный анализ доверия человека к искусственному интеллекту в экономической деятельности**

**Научный руководитель – Левченко Андрей Александрович**

*Шарыгин Никита Александрович*

*Студент (магистр)*

Поволжский институт управления имени П.А. Столыпина - филиал «Российской академия народного хозяйства и государственной службы при Президенте Российской Федерации», Факультет магистратуры и аспирантуры, Саратов, Россия

*E-mail: sharyginnik@mail.ru*

В настоящее время, когда искусственный интеллект все интенсивнее внедряется в различные сферы жизни человека, от медицины до образования, уровень доверия к нему становится все более острой темой дискуссий и исследований. В контексте искусственного интеллекта доверие человека к такой модели - это прежде всего доверие к результату ее работы.

Доверие к результату деятельности в свою очередь - это устойчивость модели, способность к обобщению, отсутствие уязвимостей и множество других факторов [1, с. 75]. К сожалению, активное развитие моделей искусственного интеллекта не сопоставимы с темпами разработки методов измерения такого типа доверия. Перед исследователями по всему миру поставлена задача не только теоретически определить, что такое доверие, но и практически измерить и внедрить в реальную деятельность организаций подобное наблюдение.

Одним из первых академических экспериментальных исследований можно считать работу Винокурова Ф.Н и Садовской Е.Д. в Московском государственном институте имени М.В. Ломоносова “Экспериментальное сравнение доверия искусственному интеллекту и человеку в экономических решениях”. В рамках экономической игры участники принимали решение инвестировать денежные средства в ту или иную фирму, опираясь на совет эксперта в лице человека или ПО [2 с. 90]. Такой дизайн исследования показывает, как доверие меняется в процессе диалога с автоматизированной системой, и как ошибки или успехи модели влияют на последующие решения.

Комплексные экспериментальные дизайны, которыми пользуются в настоящее время исследователи, предполагают комбинирование трех видов измерения: шкал самоотчета, экономических задач и дополнительных объективных индикаторов (eye-tracking, время принятия решения, сердечный ритм и т.д.). Такой подход позволяет не просто оценить субъективные факторы, но и зафиксировать, как доверие влияет на поведение человека, а также зафиксировать, как и на каком этапе зародилось это самое доверие.

Так исследователями из Пекинского университета химической технологии и Пекинского союзного университета, На Чен, Яо Цзу, Цзин Сун был произведен эксперимента при помощи электроэнцефалографа. В рамках эксперимента изучалось доверие человека к автономному автомобилю. Участник помещался в изолированное помещение, окруженное экранами, на которых транслировалась поездка на автомобиле. В экспериментальных ситуациях участники смотрели различные видеоролики с «проблемой вагонетки» (Фут, 1967), в которых автономные транспортные средства принимали разные моральные решения. Чтобы проверить, правильно ли участники эксперимента определили моральный тип транспортного средства в различных ситуациях, после просмотра видео им было предложено ответить на два вопроса, подтверждающих или опровергающих, осознание морального поступка транспортного средства. В результате опроса, математических вычислений

и съема данных с ЭЭГ, а именно значения P300 формировалось заключение о мнении человека относительно модели искусственного интеллекта. [3, с. 6].

Цель моего исследования в рамках выпускной квалификационной работы - выявить поведенческие и нейрокогнитивные маркеры доверия человека при помощи совместного использования электроэнцефалографии и технологии айтрекинга.

Методически работа строится как многоуровневое эмпирическое исследование. На первом этапе участнику исследования предлагается пройти опрос для определения его материального самоощущения. Это позволяет определить, опирается ли человек на свое материальное положение или личные установки при ответе на экономические вопросы.

Второй этап исследования, предполагает показ заранее заготовленных карточек с вопросом, связанным с ним изображением, отдельным советом от эксперта в лице человека, советом от эксперта в лице ИИ. Участник принимает решение путем выбора варианта ответ “да” или “нет” (рисунок 1).

В процессе выполнения задания, одновременно регистрируется длительность зрительной активности в зонах интереса и показатели мозговой деятельности участника, что позволяет выявлять скрытые процессы, такие как сомнение или уверенность, не всегда видимые внешне.

Научная новизна и практическая значимость исследования заключаются в разработке комплексной модели анализа доверия человека к искусственному интеллекту, объединяющей экономическое поведение, визуальное внимание и нейрофизиологические показатели.

Особую значимость работа приобретает в условиях цифровой трансформации экономики России и активном развитии отечественных систем искусственного интеллекта в финансовом секторе.

### Источники и литература

- 1) Д. Е. Намиот, Е. А. Ильюшин Об оценке доверия к системам Искусственного интеллекта // International Journal of Open Information Technologies. 2025. №3
- 2) Винокуров Ф.Н., Садовская Е.Д. Экспериментальное сравнение доверия искусственному интеллекту и человеку в экономических решениях // Экспериментальная психология. 2023. Том 16. № 2. С. 87–100
- 3) Чэнь Н., Цзу Ю. и Сун Дж. (2023) Исследование влияния и механизма морального соответствия человека и транспортного средства на доверие к автономным транспортным средствам. Front. Psychol. 14:1071872

### Иллюстрации



Рис. : Дизайн карточки с вопросом