

## **Исследование создания «искусственных респондентов» на основе LLM для пользовательских исследований**

**Половинко А.С.**

**Научный руководитель – старший преподаватель Балканский А.А.**

<sup>1</sup>Университет ИТМО

aleksa.polovinko@yandex.com

### **Введение**

В современной практике UX-исследований подбор респондентов и проведение глубинных интервью требуют много времени и материальных средств. Поэтому использование больших языковых моделей LLM выглядит всё более перспективным для создания синтетических респондентов, которые могут имитировать ответы реальных пользователей в исследованиях. Однако до сих пор нет единого мнения, как именно формулировать запрос к нейросети, чтобы получать результаты достаточного качества для формирования достоверных выводов [1-2]. Поэтому компаниям сложно применять LLM в пользовательских интервью. Непонятно, какие сведения о человеке необходимо использовать в промте, чтобы ИИ-респондент был максимально точным. В рамках работы были изучены существующие подходы к созданию ИИ-респондентов, выделен и обоснован перечень характеристик и черт человека, значимых для корректной имитации его поведения нейросетью в пользовательских исследованиях, а также способы измерения семантической близости текстов. Сравнились два подхода, а именно генерация на основе чёткого набора данных и генерация на основе неструктурированного интервью.

### **Основная часть**

В ходе исследования были выделены и протестированы два принципиально разных подхода к описанию респондента в промте. Первый подход основан на атрибутивном описании. Он опирается на психологию черт и задаёт жёсткую структуру личности через валидизированные опросы. Для формирования профиля использовались вопросы European Social Survey (социокультурный контекст), методика Big Five-2-S [3] (личностные особенности), Subjective Knowledge Scale (субъективная компетентность) и ATI-S (отношение к технологиям). Второй подход строится на коротком монологе. Предполагается, что индивидуальные особенности лучше всего проявляются в естественной речи человека [4]. Для обоих подходов были разработаны наборы вопросов, с помощью которых собирались данные о реальных людях. Также был сформирован инструмент оценки на основе существующих метрик и моделей представления текста (BERTScore, E5 и Sentence-BERT), необходимый для сравнения ответов искусственного респондента с эталонными ответами реального человека. Он рассчитывает семантическую схожесть, которая является целевой метрикой качества искусственного респондента.

### **Выводы**

Проведенный эксперимент показал, что для задачи моделирования ответов по семантическому сходству атрибутивное описание оказалось эффективнее короткого монолога. Монолог в текущем виде добавляет шум и смещает контекст, иногда вводит неактуальные ассоциации.

### Литература

1. Bisbee J. и др. Synthetic Replacements for Human Survey Data? The Perils of Large Language Models // Political Analysis. – 2024.
2. Lutz M. и др. The Prompt Makes the Person(a): A Systematic Evaluation of Sociodemographic Persona Prompting for Large Language Models // arXiv preprint arXiv:2507.16076. – 2025.
3. Aleksandra Sorokovikova, Sharwin Rezagholi, Natalia Fedorova, Ivan P. Yamshchikov. LLMs Simulate Big5 Personality Traits: Further Evidence // ACL Anthology. – 2024.
4. Park J.S., Zou C.Q., Shaw A., Hill B.M., Cai C., Morris M.R., Willer R., Liang P., Bernstein M.S. Generative Agent Simulations of 1,000 People // arXiv preprint arXiv:2411.10109. – 2024.