

Применение синтетических респондентов в социологических и маркетинговых исследованиях

Научный руководитель – Маркеева Анна Валерьевна

Жуликов Кирилл Олегович

Студент (магистр)

Московский государственный университет имени М.В.Ломоносова, Социологический факультет, Кафедра экономической социологии и маркетинга, Москва, Россия

E-mail: Kir.prog2002@gmail.com

Актуальность обращения к синтетическим респондентам в маркетинговых и социологических исследованиях задается тремя процессами: ростом стоимости и организационной сложности полевых опросов, хроническими проблемами качества онлайн-выборок (выгорание панелей, неаутентичные ответы, боты) и распространением генеративного искусственного интеллекта, воспроизводящего речевые и поведенческие паттерны, внешне неотличимые от человеческих. В результате синтетизация респондента становится методологической развилкой: ускорение исследовательского цикла сочетается с рисками подмены социального факта статистически правдоподобной имитацией [10; 11].

В прикладной перспективе синтетические респонденты востребованы на ранних стадиях исследовательского контура: генерация гипотез, диагностика анкеты, предварительное тестирование формулировок, моделирование реакции на концепты продукта или коммуникации. Обзоры фиксируют сдвиг от «полной» синтетизации к гибридным схемам, где «затравочный» массив реальных интервью используется для настройки модели и последующего наращивания синтетических ответов [7; 8]. Однако устойчивость выводов зависит от контекста и типа измерения, поэтому требуется обязательная эмпирическая валидация [7].

Для социологической методологии принципиально, что синтетический респондент – вычислительная конструкция, чья «субъектность» задается корпусами данных и инженерией «персоны». Отсюда необходимость различать синтетические данные как эвристику/симулятор и данные как основание репрезентативного вывода. Сопоставления синтетических массивов с массовыми обследованиями показывают пригодность синтетики для ограниченного класса задач, но предупреждают от подмены репрезентативных опросов и тем более нормативных решений [9]. Значимость контроля качества цифровых опросов (параданные, поведенческие индикаторы, паттерны ответов) возрастает, поскольку синтетические агенты способны адаптироваться к стандартным «фильтрам внимания» [3; 2].

Синтетизация меняет структуру рисков: к ошибкам выборки и неответам добавляется «алгоритмическое загрязнение» – скрытое включение синтетических агентов в онлайн-опросы, размывающее границу между данными о социальных действиях и модельно реконструированными ответами [11]. Отдельный вызов – социальная и культурная смещенность: синтетические ответы воспроизводят нормы обучающих массивов, повышая вероятность унификации и вытеснения маргинальных позиций. Российская литература описывает этот сдвиг как переход от цифрового к алгоритмическому неравенству, где различия закрепляются не только доступом к технологиям, но и способами категоризации и нормирования смыслов [4; 1].

Институционально применение синтетических респондентов в России сопряжено с режимом персональных данных и требованиями к обезличиванию. Синтетизация не отменяет законности исходных данных и ответственности оператора; последние изменения связаны с регулированием оборота обезличенных данных и уточнением процедур обезличивания [5], а также с закреплением требований и методов обезличивания в подзаконном акте

[6]. Перспективной представляется смешанная архитектура исследований: синтетические ответы – как симулятор и ускоритель пилотажа, человеческие данные – как основание валидности и интерпретации социальных механизмов [9; 2; 4].

Источники и литература

- 1) Былевский П. Г. Социально-культурные риски мультимодальных больших генеративных моделей «искусственного интеллекта» (GenAI) // Культура и искусство. – 2024. – №. 6. – С. 213-224.
- 2) Корытникова Н. В. Параданные как показатели качества данных онлайн-опроса: опыт классификации // Социологические исследования. – 2021. – №. 3. – С. 111-120.
- 3) Попов Д. С. Потенциал использования параданных в качественных исследованиях и контроле качества онлайн-опросов // Мониторинг общественного мнения: экономические и социальные перемены. 2025. [Электронный ресурс]. URL: <https://monitoringjournal.ru/index.php/monitoring/article/view/2880> (дата обращения: 02.03.2026).
- 4) Мартыненко Т. С., Добринская Д. Е. Социальное неравенство в эпоху искусственного интеллекта: от цифрового к алгоритмическому разрыву // Мониторинг общественного мнения: экономические и социальные перемены. – 2021. – №. 1. – С. 171-192.
- 5) Российская Федерация. Федеральный закон от 08.08.2024 № 233-ФЗ «О внесении изменений в Федеральный закон “О персональных данных” ...» // Официальный интернет-портал правовой информации. 2024. URL: <https://publication.pravo.gov.ru/document/0001202408080031> (дата обращения: 02.03.2026).
- 6) Роскомнадзор. Приказ от 19.06.2025 № 140 «Об утверждении требований к обезличиванию персональных данных и методов обезличивания...» : [Электронный ресурс]. 2025. URL: <https://normativ.kontur.ru/document?documentId=500957&moduleId=1> (дата обращения: 02.03.2026).
- 7) Ipsos. Ipsos Views: Synthetic data – From hype to reality : [Электронный ресурс]. 2024. URL: <https://www.ipsos.com/sites/default/files/ct/publication/documents/2024-08/Ipsos%20Views%20Synthetic%20Data.pdf> (дата обращения: 02.03.2026).
- 8) Ipsos. Ipsos Views: The Power of Product Testing with Synthetic Data : [Электронный ресурс]. 2025. URL: https://www.ipsos.com/sites/default/files/ct/news/documents/2025-11/IpsosViews_ThePowerOfProductTestingWithSyntheticData.pdf (дата обращения: 02.03.2026)
- 9) Verasight. Can Large Language Models Replicate Survey Data? Synthetic omnibus survey report: [Электронный ресурс]. 2026. URL: <https://www.verasight.io/reports/synthetic-omnibus-survey> (дата обращения: 02.03.2026).
- 10) Verian Group. Synthetic Sample in Social Research: significant limitations and risks : [Электронный ресурс]. 2025. URL: <https://www.veriangroup.com/news-and-insights/synthetic-sample-in-social-research> (дата обращения: 02.03.2026).
- 11) Westwood S. J. The potential existential threat of large language models to online survey research // Proceedings of the National Academy of Sciences. 2025. Vol. 122. Art. e2518075122. DOI: 10.1073/pnas.2518075122. URL: <https://www.pnas.org/doi/10.1073/pnas.2518075122> (дата обращения: 02.03.2026).