

**Олейникова Олеся Олеговна**

студентка

**Ступина Алёна Александровна**

д-р техн. наук, профессор

**Потапов Борис Васильевич**

д-р техн. наук, профессор

ФГБОУ ВО «Российский государственный  
аграрный университет – МСХА им. К.А. Тимирязева»  
г. Москва

## **СОЦИАЛЬНО-ЭКОНОМИЧЕСКИЕ ЭФФЕКТЫ АВТОМАТИЗАЦИИ РАЗМЕТКИ ДАННЫХ: СИНТЕТИЧЕСКИЕ СРЕДЫ КАК ФАКТОР ТРАНСФОРМАЦИИ ЦИФРОВОГО ТРУДА**

**Аннотация:** в статье анализируется влияние технологий синтеза данных и доменной адаптации (*Domain Adaptation*) на структуру рынка цифрового труда и барьеры входа для инновационных стартапов. Показано, что использование физически точных виртуальных сред (*NVIDIA Isaac Gym*) в сочетании с состязательными нейросетями (*DANN*) позволяет достичь точности классификации редких событий в 87,1% без привлечения ручного труда разметчиков. Обосновывается тезис о том, что синтетические данные становятся новым экономическим активом, способствующим демократизации ИИ, снижению эксплуатационных рисков и трансформации профессии «оператор данных» в «инженера симуляций».

**Ключевые слова:** социология труда, цифровая экономика, синтетические данные, человеческий капитал, доменная адаптация, инновационное предпринимательство, автоматизация.

**Введение.** Фундаментальной проблемой современной цифровой экономики является «кризис разметки»: создание качественных датасетов для обучения ИИ

требует колоссальных затрат ручного труда, что порождает новую форму социального неравенства – концентрацию технологий у корпораций, способных оплачивать краудсорсинговые платформы. В сферах, связанных с экстремальными ситуациями (биомеханика аварий, редкие патологии, критические отказы техники), сбор реальных данных физически невозможен или этически неприемлем [1]. Возникает феномен «длинного хвоста» (long-tail distribution), где 90% данных описывают рутину, а критически важные 10% отсутствуют. Цель работы – оценить социально-экономический потенциал инженерного конвейера, обеспечивающего *zero-shot transfer* (перенос без примеров) моделей из синтетических сред в реальность.

1. *Методология.* От ручного труда к физическому моделированию – вместо найма тысяч разметчиков предлагается генерация синтетических данных на платформе NVIDIA Isaac Gym. Симулятор воспроизводит лагранжеву динамику твердых тел, гравитацию и инерцию, автоматически выдавая идеальные ground-truth аннотации (3D-координаты, оптический поток, фазы движения) [2]. Для преодоления разрыва между симуляцией и реальностью (Sim-to-Real Gap) применяются следующие методы.

1. *Визуальная доменная рандомизация:* случайное изменение освещения, шумов сенсоров и текстур, что учит ИИ игнорировать поверхностные признаки.

2. *Состязательная адаптация (DANN):* использование слоя инверсии градиентов (GRL), который заставляет сеть извлекать инвариантные признаки, общие для синтетики и реальности [3].

Таблица 1

Социально-экономическое сравнение парадигм создания датасетов

| Критерий                 | Традиционный подход (Ручная разметка) | Синтетический подход (Симуляция + DANN)    |
|--------------------------|---------------------------------------|--------------------------------------------|
| Трудоемкость             | Высокая (тысячи человеко-часов)       | Нулевая (автоматическая генерация)         |
| Доступ к редким событиям | Ограничен (ожидание инцидентов)       | Абсолютный (моделирование любых сценариев) |
| Стоимость 10k примеров   | > \$50,000 (аутсорсинг)               | < \$500 (облачные GPU-вычисления)          |

|                        |                                                 |                                                             |
|------------------------|-------------------------------------------------|-------------------------------------------------------------|
| Влияние на рынок труда | Создание прекариата (низкооплачиваемые кликеры) | Создание высокотехнологичных рабочих мест (инженеры физики) |
| Этика и приватность    | Риски утечки биометрии и ПДн                    | Полное отсутствие реальных людей в кадре                    |

2. *Эмпирические результаты.* Демократизация доступа к ИИ Эксперимент по классификации 10 фаз биомеханического движения (включая такие редкие, как «экстремальная гиперэкстензия» или «рывок») показал, что модель ViT-B/16, обученная *исключительно* на синтетике с применением DANN, достигает сбалансированной точности (ВAcc) 87,1% на реальных видеопотоках. Оракул (модель, обученная на дорогих реальных данных), показывает 92,4% [4].

3. *Обсуждение.* Трансформация человеческого капитала Внедрение синтетических данных ведет к структурным сдвигам на рынке труда.

1. *Исчезновение «цифровых потогонен»:* спрос на низкоквалифицированных операторов разметки (data labelers) будет падать. Это позитивный социальный эффект, освобождающий ресурсы от монотонного труда.

2. *Рост спроса на «Инженеров симуляций»:* возникает потребность в специалистах на стыке физики, математики и социологии, способных проектировать виртуальные миры, адекватные реальности. Это стимулирует развитие STEM-образования.

3. *Снижение барьеров для МСБ:* малый бизнес может позволить себе внедрение систем компьютерного зрения для контроля качества или безопасности, не инвестируя миллионы в сбор данных. Это *оздоравливает* конкурентную среду [5].

*Заключение.* Синтетические данные и методы доменной адаптации перестают быть узкоспециальным инженерным инструментом и становятся макроэкономическим фактором. Они решают проблему дефицита данных, снижают стоимость инноваций и трансформируют рынок труда, заменяя эксплуатацию рутинного человеческого интеллекта машинным моделированием физических законов. Государственная политика в сфере ИИ должна стимулировать создание открытых библиотек физических симуляторов как инфраструктуры национального значения.

### ***Список литературы***

1. Потапов Б.В. Применение методов машинного обучения для прогнозирования прочностных характеристик горных пород / Б.В. Потапов // Известия Тульского государственного университета. Науки о Земле. – 2026. – №2. – С. 406–411. EDN VSQHDC
2. Прогнозирование тепловых процессов в подземных горных сооружениях с использованием рекуррентных нейронных сетей / М.В. Журавлев, О.А. Антамошкин, М.Н. Степанцевич, Е.Ф. Малыха // Известия Тульского государственного университета. Науки о Земле. – 2025. – №3. – С. 205–211. EDN LZFRCG
3. Пчелинцева С.В. Применение методов компьютерного зрения для автоматизированной диагностики деформаций и разрушений горных выработок / С.В. Пчелинцева, М.В. Журавлев // Известия Тульского государственного университета. Науки о Земле. – 2026. – №1. – С. 359–364. EDN KOUAZF
4. Shcherban P.S. Detection of defects in steel oil and gas operating pipelines using a neural network, based on the results of optical diagnostics / P.S. Shcherban, B.V. Potapov, A.N. Karagadyan // Research on Engineering Structures and Materials. – 2025. – P. 3051–3068.
5. Обнаружение аномалий в данных промышленных сенсорных систем добычи на основе алгоритмов кластеризации / С.О. Курашкин, Л.В. Красовская, М.В. Журавлев, Т.В. Бирюкова // Известия Тульского государственного университета. Науки о Земле. – 2025. – №4. – С. 382–388. EDN WMCTCH
6. Применение методов машинного обучения для автоматизированной классификации магнитных аномалий... / В.С. Тынченко, М.В. Журавлев, Н.А. Ягудева [и др.] // Геология и Геофизика Юга России. – 2026. – Т. 16. №2. – С. 108–119. DOI 10.46698/VNC.2026.48.42.001. EDN CBGZAI
7. Kazakovtsev D.V. Intelligent Sensor Data Processing Methods for Real-Time Control of Mechatronic Manufacturing Systems / D.V. Kazakovtsev, B.V. Potapov // AISEI. – IEEE, 2026. – P. 1230–1235.

8. Stepantsevich M.N. Federated Learning for Privacy-Preserving Energy Consumption Prediction in Smart Buildings / M.N. Stepantsevich, B.V. Potapov // AISEI. – IEEE, 2026. – P. 1182–1187.

9. Функциональное моделирование бизнес-процессов как основа проектирования информационных систем предприятия / М.В. Журавлев, В.Ю. Панченко, С.В. Кукарцева, А.Р. Глинская // Социально-экономические процессы современного общества. – Чебоксары: Среда, 2026. – С. 163–167. EDN LVNZDY