

Введение. В настоящее время векторная графика является удобным инструментом в дизайне, технической графике и создании научных иллюстраций, так как использует примитивы: линии, окружности и пути. Векторная графика основана на языке разметки XML, что позволяет использовать большие языковые модели (БЯМ) для работы с ней. Основное преимущество такого подхода — взаимодействие с программой на естественном языке без обязательного изучения технических деталей. Недавние работы продемонстрировали возможность реализации данного метода [1].

Основная часть. БЯМ применяются для решения следующих задач:

- 1) Безусловная генерация векторной графики или генерация на основе запроса на естественном языке [2], позволяющая создавать графические элементы.
- 2) Редактирование векторной графики: добавление и удаление отдельных элементов, замена числовых и категориальных значений. Например, сжатие изображений посредством удаления элементов, не влияющих на растеризованный вид, а также оптимизация как отдельных путей, так и внутренней структуры XML кода, для уменьшения размера файлов и повышения их читаемости [3].
- 3) Распознавание моделью деталей векторной графики, категоризация, выделение количественных и качественных значений, что может быть полезно для автоматизированного семантического поиска и контекстного анализа [4].

Во всех трёх задачах используются различные комбинации входных и выходных форматов данных: (текст, вектор) – генерация, (вектор, вектор) – редактирование, (вектор, текст) – распознавание и представление в виде естественного языка.

Создание модели, способной эффективно решать все вышеуказанные задачи, откроет широкие возможности для работы с векторной графикой. Однако в связи с этим требуется разработка единой методики оценки качества работы БЯМ с векторной графикой.

Выводы. Проведен анализ существующих подходов к генерации векторной графики с использованием больших языковых моделей, рассмотрены открытые наборы данных и способы апробации качества генерации. Предложен подход к созданию универсальной методики оценки эффективности БЯМ при работе с векторной графикой.

Список использованных источников:

1. Bubeck S. et al. Sparks of artificial general intelligence: Early experiments with gpt-4 // arXiv preprint arXiv:2303.12712. – 2023.
2. Wu R., Su W., Ma K., Liao J. IconShop: Text-Guided Vector Icon Synthesis with Autoregressive Transformers // ACM Transactions on Graphics (TOG). 2023. Т. 42, № 6. С. 1–14.
3. Nishina K., Matsui Y. SVGEditBench: A Benchmark Dataset for Quantitative Assessment of LLM's SVG Editing Capabilities // arXiv preprint. 2024. № 2404.13710. URL: <https://arxiv.org/abs/2404.13710> (дата обращения: 12.02.2025).
4. Cai M., Huang Z., Li Y., Ojha U., Wang H., Lee Y.J. Leveraging large language models for scalable vector graphics-driven image understanding // arXiv preprint. 2023. № 2306.06094. URL: <https://arxiv.org/abs/2306.06094> (дата обращения: 12.02.2025).