

Университет ИТМО, факультет программной инженерии и компьютерной техники  
Двухнедельная отчётная работа по «Информатике»: аннотация к статье

Дата прошедшей лекции	Номер прошедшей лекции	Название статьи/главы книги/видеолекции	Дата публикации (не старше 2022 года)	Размер статьи (от 400 слов)	Дата сдачи
10.09.2025	1	Non-Standard Binary Representations and the Stern Sequence	14.08.2023	~3,800	24.09.2025
24.09.2025	2	LLM-Guided Search for Deletion-Correcting Codes	01.04.2025	~10000	08.10.2025
08.10.2025	3	Intelligent Interface: Enhancing Lecture Engagement with Didactic Activity Summaries	20.06.2024	~7500	22.10.2025
22.10.2025	4	Advancing vision-language models in front-end development via data synthesis	03.03.2025	~12000	05.11.2025
	5				
	6				
	7				

Выполнил(а) Эллити Мохамед Эмал Ахмед Авад, № группы P3131, оценка             
Фамилия И.О. студента не заполнять

Прямая полная ссылка на источник или сокращённая ссылка (bit.ly, tr.im и т.п.)  
Прямая ссылка на источник: <https://arxiv.org/abs/2503.01619>

Теги, ключевые слова или словосочетания (минимум три слова)  
Генерация фронтенд-кода ,Мультимодальные модели, Синтез данных, Компьютерное зрение и NLP

Перечень фактов, упомянутых в статье (минимум четыре пункта)

☐ Предложен **автоматизированный пайплайн синтеза данных** для обучения VLM в задачах генерации фронтенд-кода.

☐ Введены три метода синтеза: **Evolution-Based**, **Waterfall-Model-Based** и **Additive Development**.

☐ Разработана модель **Flame** (на базе SigLIP и DeepSeek-Coder).

☐ Создан бенчмарк **Flame-React-Eval** с 80 тестами для генерации React-кода по изображению.

☐ Показано, что **интерпретация изображения перед генерацией** повышает качество кода.

Позитивные следствия и/или достоинства описанной в статье технологии (минимум три пункта)  
Высокое качество: Flame превосходит GPT-4o, Gemini и другие SOTA-модели в генерации React-кода.  
Гибкость: пайплайн поддерживает одиночные и множественные изображения на входе.  
Открытость: код, данные и модели доступны для воспроизводимости и дальнейшего развития.  
Практическая ценность: бенчмарк оценивает синтаксис и визуальное соответствие через эмбединги DINOv2.

Негативные следствия и/или недостатки описанной в статье технологии (минимум три пункта)


☐ **Зависимость от данных:** качество генерации сильно зависит от синтезированных layout-описаний.

☐ **Ограничение на React:** методы не подходят для Vue, Angular и других фреймворков.

☐ **Однозадачность:** пайплайн не поддерживает итерации или диалоговые сценарии.

☐ **Высокие требования:** обучение требует 24×H800 GPU и больших объёмов данных.

Ваши замечания, пожелания преподавателю или анекдот о программистах<sup>1</sup>



\_\_\_\_\_ны не влияет на оценку