

Обзор Gemini 3.5 Flash и Antigravity с Google I/O 2024

Google I/O 2024: Обзор Gemini 3.5 Flash и Antigravity

Ключевые тезисы

- **Gemini 3.5 Flash** — новая, более интеллектуальная, но и более дорогая модель от Google.
- **Antigravity 2.0** — новое автономное приложение для кодинга, похожее на Cursor, Codeium и другие аналоги.
- Модель не предназначена для сложного кодинга, её сильная сторона — выполнение рутинных задач и работа в автономных агентах.
- Ценовая политика Gemini 3.5 Flash ставит под вопрос её оптимальность для длительных задач с большим контекстом.

Gemini 3.5 Flash: Характеристики и бенчмарки

- **Цена выросла в 3 раза:** \$1.5 за 1М входных и \$1 за 1М выходных токенов.
- **Контекст:** 1 миллион токенов.
- **Производительность в тестах:**
 - В индексе интеллекта (Artificial Analysis Intelligent Index) находится рядом с **GPT-4o** и чуть ниже **Gemini 3.1 Pro**.
 - В тестах на **кодинг** показывает результаты ниже многих конкурентов (Xiaomi, Kimi, Spark), что подтверждает её ориентацию не на разработчиков.
 - Хорошо показывает себя в задачах для **автономных агентов**, находясь в одном ряду с Orus и GPT-4o.
 - Соотношение "интеллект/потраченные токены" хуже, чем у конкурентов: для решения задачи ей требуется больше токенов.



Antigravity 2.0 App: Первое впечатление

- **Что это:** Автономное десктоп-приложение, форк VS Code с интегрированным Gemini.
- **Интерфейс:** Очень простой и знакомый — сайдбар с проектами (директориями) и чатами внутри них.
- **Настройки:** Можно отключить постоянные запросы разрешений (permissions), выставив `unrestricted` и `always proceed` для проекта (аналог bypass в Cursor).
- **Вывод:** Приложение выглядит **сырым и урезанным**, без уникальных функций и интеграций, которые есть у аналогов. Также присутствуют **строгие лимиты** на использование (кредитная система).



Практический тест: Создание лендинга и дашборда

Цель: Проверить автономность работы Gemini 3.5 Flash в Antigravity на фронтенд-задачах.

Задача: Создать лендинг и дашборд по одному исчерпывающему промпту с требованием вернуть чистый HTML.

Результаты Gemini 3.5 Flash:

- **Лендинг:** Модель проигнорировала запрос на HTML и сгенерировала сложный стек (React, TypeScript, Framer Motion). Результат — обычный generic-лендинг **без адаптивности** и с кривой светлой темой.
- **Дашборд:** Та же проблема. Получена страница с аляповатыми компонентами и без корректной светлой темы.

Сравнение с Orus 4.7:

- Orus **точно выполнил инструкцию**, вернув чистый HTML.
- Результаты (лендинг и дашборд) были **более детализированными, динамичными и проработанными** даже с одного промпта.

Выводы и инсайты

1. **Gemini 3.5 Flash** — мощная мультимодальная модель, но её **ниша неочевидна**. Для больших задач дороговата, для сложного кодига — не предназначена.
2. Основная гипотеза применения — **автономные агенты** (типа OpenClow, Гермес) для рутинных операций, но требует проверки на эффективность по токенам.
3. **Antigravity App** на текущий момент — **сырой и ограниченный продукт** без конкурентных преимуществ.
4. Внутри Google команды уже переведены на использование собственных моделей (Gemini Pro), что в перспективе может улучшить их качество.
5. Ожидания теперь связаны с анонсом **Gemini 3.5 Pro**, которая должна составить конкуренцию Opus 4.7 и GPT-4o.