

## Сравнение Claude Opus и GPT-5.5 в генерации кода

---

# Битва AI-кодеров: Claude Opus 4.7 vs GPT-5.5 (Codex Desktop)

### Ключевые тезисы:

- Эксперимент по адаптации реального продакшен-стартеркита (7600 строк кода) под стоматологический кейс.
  - Обе модели (Claude Opus 4.7 и GPT-5.5 в Codex Desktop) получили **один идентичный промт** без подсказок.
  - Оценка по 8 критериям (безопасность, тесты, архитектура и др.), максимум 80 баллов.
  - **GPT-5.5 оказался в ~2 раза быстрее**, но итоговое качество кода по результатам перекрёстного код-ревью оказалось очень близким.
  - Лучший стек на 2026 год — **комбинация подписок** на обе модели для генерации и взаимной проверки кода.
- 





## Условия эксперимента

**Цель:** Адаптировать готовый продакшен-стартеркит AI-продавца под стоматологическую клинику «Улыбка+».






- **Стартеркит:** 7600 строк кода, 47 файлов. Включает админку, виджет, тесты, защиту от промт-инъекций (Guard Rails), интеграции с CRM.
- **Промт:** Изучить файлы проекта и адаптировать промт бота, базу знаний, визуал. **Не трогать ядро (Engine), Guard Rails и безопасность.**
- **Методология:** Один промт → генерация → тестирование функционала → **перекрёстное код-ревью** (каждая модель проверяет код конкурента).

# Результаты генерации и тестирования

## GPT-5.5 (Codex Desktop)


-  **Скорость:** Выполнил задачу примерно в **2 раза быстрее** Claude.
-  **Фронтенд (виджет):** Сделал аккуратно, визуально приятно. Функционал работает.
-  **Безопасность:** Защита от трёх типов промт-инъекций сработала корректно (блокировка после 3 попыток).
-  **Админ-панель:** Сделана, но с **критичными недочётами:**
  - Главная страница раскрывает всю внутреннюю информацию.
  - Часть кнопок не работает или работает некорректно (например, в диалогах).
  - События промт-инъекций не записывались в лог.
-  **Выбор моделей:** Добавил несколько LLM на выбор (включая дорогую GPT-5.1), что соответствует требованиям стартеркита.

## Claude Opus 4.7 (Claude Code)

-  **Скорость:** Работал **значительно дольше** (более 30 минут).
-  **Фронтенд (виджет):** Внешне похож на результат GPT, но дизайн сайта, созданного «зачем-то», слабый.
-  **Безопасность:** Защита от инъекций сработала безупречно, логирование ведётся.
-  **Админ-панель:** Реализована **качественнее:**
  - Чистый, минималистичный интерфейс.
  - Вся статистика и функционал (перехват диалога, блокировка) работают корректно.
  - Поля настроек редактируемы.
-  **Главные минусы:**
  - Навязчивое использование смайликов в интерфейсе.
  - В настройках по умолчанию выбрана **дорогая модель Claude 4.6**, не добавлены более дешёвые альтернативы, что может привести к большим расходам.

## Перекрёстное код-ревью (по 8 критериям)

Каждая модель оценила код конкурента по шкале от 0 до 10 по каждому из 8 критериев.

Критерий	GPT-5.5 о коде Claude	Claude о коде GPT-5.5
Безопасность	6/10	6/10
Модульность	7/10	8/10
Тесты	7/10	6/10
Документация	6/10	5/10
Обработка ошибок	7/10	8/10
Готовность к продакшену	6/10	7/10
Качество кода	7/10	8/10
Функциональность	6/10	6/10
 ИТОГО	51/80	53/80

**Вывод по ревью:** Разница всего в 2 балла. Модели нашли друг у друга схожие критические проблемы (секреты в коде, слабые тесты). Ни один код не был признан идеально готовым к продакшену.

## Автоматизация код-ревью: Плагин CodeEx

Для Claude Code существует официальный плагин **CodeEx**, который автоматизирует проверку кода:

- **CodeEx Adversarial Review** : GPT критикует план кода от Claude до начала написания.
- **CodeEx Rescue** : Полный аудит кодовой базы (безопасность, производительность).
- **CodeEx Code Review** : Проверка готового кода.
- **CodeEx Watchdog** : Автоматическая проверка каждого коммита.

**Идеальная схема:** Claude генерирует код → GPT (через CodeEx) проверяет и ищет ошибки.



## Итоговые выводы и рекомендации

1. **Скорость vs. Качество:** GPT-5.5 быстрее, Claude Opus 4.7 в этом тесте сделал более качественную админ-панель.
2. **Реальная разница — в инструментах:** Как LLM модели близки по результату. Главное отличие — экосистема (Claude Code vs. Codex Desktop). Claude Code пока предлагает более продвинутые инструменты для разработки (навигация, память, субагенты).
3. **Лучший стек на 2026 год:**
  - **Для новичка:** Начать с GPT-5.5 (Codex Desktop, \$20/мес).
  - **Для продвинутого пользователя/профессионала:** Комбинация **Claude Code (\$100/мес) + подписка на GPT (\$20/мес)** для взаимной проверки кода. Это дает две мощные модели за \$120.
  - **Максимальная эффективность:** Claude Code Enterprise (\$200/мес) + GPT для ревью.
4. **Ни одна модель** не создаст идеальный продакшен-код с одного промта. Они дают **рабочее MVP**, которое нужно итеративно дорабатывать.
5. **Не привязывайтесь к одному инструменту.** Лучшие разработчики используют сильные стороны разных систем.