

ВТОРОЕ ИЗДАНИЕ

# Разработка приложений на базе GPT-4 и ChatGPT

*интеллектуальные чат-боты,  
генераторы контента и многое другое*

*Оливье Келен, Мари-Алис Блете*

**SPRINT** 2025  
book

Выпущено  
при поддержке

**КРОК**

---

# Краткое содержание

Предисловие .....	11
<b>Глава 1.</b> Обзор GPT-4 и ChatGPT .....	15
<b>Глава 2.</b> Глубокое погружение в API OpenAI .....	53
<b>Глава 3.</b> Навигация по приложениям с поддержкой LLM: возможности и проблемы .....	116
<b>Глава 4.</b> Продвинутые стратегии интеграции LLM с OpenAI .....	170
<b>Глава 5.</b> Расширение возможностей LLM с помощью фреймворков, плагинов и многого другого .....	233
<b>Глава 6.</b> Собираем все воедино .....	282
Глоссарий .....	290
Приложение. Инструменты, библиотеки и фреймворки .....	299

---

# Оглавление

Предисловие .....	11
Условные обозначения.....	12
Использование примеров кода .....	13
От издательства .....	13
Благодарности.....	14
<b>Глава 1. Обзор GPT-4 и ChatGPT.....</b>	<b>15</b>
Введение в LLM .....	16
Изучение основ языковых моделей и NLP.....	16
Ключевые особенности трансформера и его роль в обработке естественного языка .....	19
Суть процесса токенизации и прогнозирования в моделях GPT .....	23
Интеграция «зрения» в программу LLM .....	24
Краткая история: от GPT-1 до GPT-4.....	27
GPT-1 .....	27
GPT-2 .....	28
GPT-3 .....	29
От GPT-3 до InstructGPT.....	29
GPT-3.5, Codex и ChatGPT.....	33
GPT-4 .....	34
Развитие ИИ в сторону мультимодальности .....	38
Сценарии использования LLM и примеры продуктов .....	40
Be My Eyes.....	40
Morgan Stanley .....	41
Khan Academy.....	41
Duolingo .....	42
Yabble .....	43
Waymark .....	43
Inworld AI.....	44

Галлюцинации ИИ: ложная информация.....	44
Раскрытие потенциала GPT с помощью дополнительных возможностей.....	48
Резюме .....	51
<b>Глава 2. Глубокое погружение в API OpenAI .....</b>	<b>53</b>
Основные понятия .....	54
Модели, доступные в API OpenAI.....	55
Базовое семейство моделей GPT.....	56
InstructGPT (устарело).....	56
GPT-3.5.....	56
GPT-4 .....	57
Использование моделей GPT с помощью OpenAI Playground .....	58
Библиотека OpenAI Python .....	62
Ключ доступа и API OpenAI.....	63
Hello World.....	65
Работа с моделями завершения чата.....	67
Параметры ввода для конечной точки завершения чата .....	68
Настройка параметров temperature и top_p .....	73
Формат выходного результата для конечной точки завершения чата .....	76
Функция компьютерного зрения Vision .....	77
Требование вывода в формате JSON .....	81
Использование других моделей завершения текста .....	85
Параметры ввода для конечной точки завершения текста .....	86
Формат результата вывода для конечной точки завершения текста .....	88
Соображения .....	88
Ценообразование и ограничения по токенам .....	88
Безопасность и конфиденциальность: осторожно!.....	89
Другие API и функционал OpenAI.....	90
Вложения .....	90
Модель модерации .....	93
Преобразование текста в речь.....	97
Перевод речи в текст .....	99
API изображений.....	103
Резюме (и шпаргалка) .....	114

<b>Глава 3. Навигация по приложениям с поддержкой LLM:</b>	
возможности и проблемы.....	116
Разработка приложений.....	116
Управление ключами API.....	117
Безопасность и конфиденциальность данных.....	119
Принципы проектирования архитектуры программного обеспечения.....	120
Интеграция возможностей LLM в ваши проекты.....	121
Разговорные возможности.....	121
Возможности обработки языка.....	122
Возможности взаимодействия человека и компьютера.....	124
Объединение возможностей.....	125
Примеры проектов.....	126
Проект 1. Создание решения для генератора новостей.....	126
Проект 2. Краткое изложение видеороликов YouTube.....	129
Проект 3. Создание ассистента для видеоигры Zelda BOTW.....	134
Проект 4. Персональный помощник — человеко-компьютерный интерфейс.....	141
Проект 5. Организация документов — обработка языка.....	149
Проект 6. Анализ настроений — обработка языка.....	150
Управление затратами.....	158
Уязвимости в приложениях с поддержкой LLM.....	160
Анализ ввода и вывода.....	161
Неизбежность атак.....	162
Работа с внешним API.....	162
Обработка ошибок и непредвиденных проблем с задержкой.....	163
Тарифные лимиты.....	164
Улучшение отзывчивости и пользовательского опыта.....	165
Резюме.....	169
<b>Глава 4. Продвинутое стратегии интеграции LLM с OpenAI.....</b>	<b>170</b>
Промт-инжиниринг.....	170
Разработка эффективных запросов с ролями, контекстами и задачами.....	172
Пошаговое мышление.....	178
Реализация обучения на примерах.....	180
Итеративная доработка с учетом отзывов пользователей.....	182
Повышение эффективности промт-инжиниринга.....	188

Тонкая настройка .....	191
Начало работы .....	192
Тонкая настройка с помощью API OpenAI .....	196
Тонкая настройка с помощью веб-интерфейса OpenAI .....	200
Тонкая настройка приложений .....	202
Генерация и тонкая настройка искусственных данных для email-маркетинга .....	205
Стоимость тонкой настройки .....	213
RAG .....	214
Простая реализация RAG .....	214
Продвинутая реализация RAG .....	215
Ограничения RAG .....	221
Выбор между стратегиями .....	222
Сравнение стратегий .....	222
Оценка результатов .....	226
От стандартного приложения к решению на базе LLM .....	227
Чувствительность к запросам .....	227
Недетерминизм .....	228
Галлюцинации .....	229
Резюме .....	231

<b>Глава 5. Расширение возможностей LLM с помощью фреймворков, плагинов и многого другого .....</b>	<b>233</b>
Фреймворк LangChain .....	233
Библиотеки LangChain .....	235
Динамические запросы .....	236
Агенты и инструменты .....	237
Память .....	242
Вложения .....	243
Фреймворк LlamaIndex .....	247
Демонстрация: RAG в десяти строках кода .....	248
Принципы LlamaIndex .....	249
Индивидуальные настройки .....	250
Плагины GPT-4 .....	252
Общие сведения .....	254
API .....	255
Манифест плагина .....	256
Спецификация OpenAPI .....	258
Описания .....	259

Пользовательские GPT .....	260
API ассистента .....	266
Создание API ассистента .....	268
Управление разговором с вашим API ассистента .....	269
Вызов функции .....	274
Ассистенты на веб-платформе OpenAI .....	277
Резюме .....	280
<b>Глава 6. Собираем все воедино .....</b>	<b>282</b>
Основные тезисы .....	282
Собираем все вместе: сценарий реализации ассистента .....	284
Шаг 1. Идея .....	284
Шаг 2. Определение требований .....	285
Шаг 3. Создание прототипа .....	286
Шаг 4. Улучшение, итерация .....	286
Шаг 5. Создание надежного решения .....	288
Извлеченные уроки .....	289
Глоссарий .....	290
Приложение. Инструменты, библиотеки и фреймворки .....	299
Об авторах .....	302
Иллюстрация на обложке .....	303