
Оглавление

От издательства	7
О научном редакторе русского издания	7
Предисловие	8
Для кого эта книга?	9
Условные обозначения	10
Использование исходного кода примеров	11
Благодарности	11
Глава 1. Введение.....	13
Паттерны проектирования для GenAI	13
Приложения на основе базовых моделей	14
Агентский искусственный интеллект	24
Детальное управление.....	26
Контекстное обучение	32
Постобучение.....	34
Организация остальной части книги	39
Глава 2. Управление стилем контента	43
Паттерн 1. «Маскирование логитов»	44
Паттерн 2. «Грамматика».....	60
Паттерн 3. «Перенос стиля»	74
Паттерн 4. «Обратная нейтрализация».....	85
Паттерн 5. «Оптимизация контента»	95
Итоги.....	112
Глава 3. Добавление знаний (часть 1)	115
Паттерн 6. «Базовый RAG».....	117
Паттерн 7. «Семантическая индексация»	131
Паттерн 8. «Индексация в масштабе»	148
Итоги.....	158
Глава 4. Добавление знаний (часть 2)	159
Паттерн 9. «Поиск с учетом индекса»	159
Паттерн 10. «Постобработка узлов»	170
Паттерн 11. «Достоверная генерация»	179
Паттерн 12. «Глубокий поиск»	198
Итоги.....	210

Глава 5. Расширение возможностей модели	213
Границы способностей к рассуждению	213
Паттерн 13. «Цепочка рассуждений»	216
Паттерн 14. «Дерево рассуждений»	229
Паттерн 15. «Настройка адаптера»	243
Паттерн 16. «Развитие инструкций» (Evol-Instruct)	259
Итоги	275
Глава 6. Повышение надежности	277
Паттерн 17. «LLM-судья»	277
Паттерн 18. «Рефлексия»	288
Паттерн 19. «Внедрение зависимостей»	296
Паттерн 20. «Оптимизация промптов»	303
Итоги	312
Глава 7. Агенты, способные действовать	314
Паттерн 21. «Вызов инструментов»	314
Паттерн 22. «Выполнение кода»	329
Паттерн 23. «Мультиагентное сотрудничество»	334
Итоги	351
Глава 8. Преодоление ограничений	353
Паттерн 24. «Малая языковая модель»	353
Паттерн 25. «Кэширование промптов»	371
Паттерн 26. «Оптимизация инференса»	380
Паттерн 27. «Тестирование деградации»	388
Паттерн 28. «Долговременная память»	400
Итоги	412
Глава 9. Установка защитных механизмов	415
Паттерн 29. «Генерация шаблонов»	415
Паттерн 30. «Сборка и переформатирование»	420
Паттерн 31. «Самопроверка»	423
Паттерн 32. «Защитные механизмы» (Guardrails)	433
Итоги	441
Глава 10. Композитные агентные рабочие процессы	443
Агентный рабочий процесс	443
Итоги	460
Об авторах	461
Иллюстрация на обложке	462