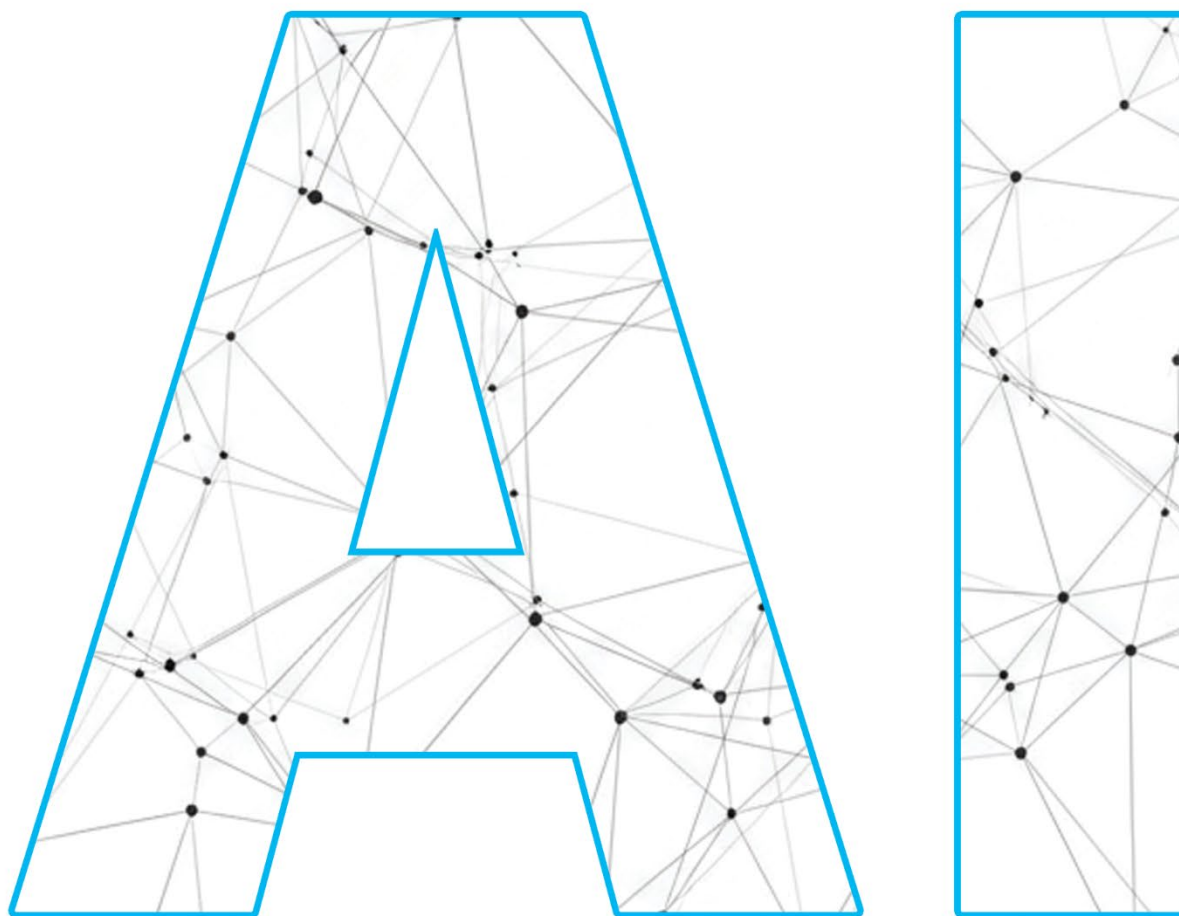


**ПРОМПТ ИНЖИНИРИНГ
ОТ НОВИЧКА ДО АРХИТЕКТОРА МНОГОСЛОЙНЫХ СИСТЕМ**

Методология **СТРАТУМ**

инженерный подход к управлению нейросетями



Игорь Степанов

aiprompting.ru



Хватит промптить Начинай проектировать

Индустрия искусственного интеллекта застряла в «стохастической петле». Мы тратим тысячи часов, пытаясь «уговорить» языковые модели выдать верный результат, но раз за разом сталкиваемся с галлюцинациями, потерей логики и деградацией внимания нейросети.

Методология Стратум предлагает системный уход от традиционного промпт-инжиниринга. Это не сборник советов, а жесткий инженерный протокол, созданный для промышленного применения. Внедряя строгую архитектуру управления, эта технология превращает непредсказуемый «черный ящик» в детерминированную машину.

80% методов промптинга умрут с приходом AGI. Стратум — это про оставшиеся 20%, которые станут единственным навыком архитектора: целеполагание, декомпозиция, безопасность, интеграция. Эта книга не о том, как уговорить модель говорить правду. Эта книга о том, как построить систему вокруг любой модели — даже самой умной.



Игорь Степанов — архитектор программных систем с более чем 30-летним стажем в IT. Выпускник классической советской инженерной школы, фундамент которой — строгая математическая логика и структурная целостность.

Сертифицированный эксперт по управлению проектами AFW Wirtschaftsakademie Bad Harzburg. С 2004 года проектирует сложные архитектуры систем товарных потоков, логистики и промышленной автоматизации.

В 2025 году перенес строгий инженерный подход в сферу искусственного интеллекта. Создатель прикладной методики PPEF (Practical Prompt Engineering Framework), автор системы высокоэффективных запросов GRIDFORT и архитектурной методологии СТРАТУМ.

ISBN 978-5-600-05136-2



9 785600 051362

Когда берешь в руки книгу по промпт-инжинирингу, подсознательно ожидаешь увидеть очередной сборник волшебных фраз вроде «действуй как эксперт» или «сделай глубокий вдох перед ответом». Но труд Игоря Степанова оказался совершенно иного калибра. Это не сборник заклинаний для общения с чат-ботом, а настоящий инженерный манифест, который безжалостно срывает с нейросетей флер мистики и магии, обнажая их математическую, вероятностную суть. С первых страниц автор дает понять одну критически важную вещь: искусственный интеллект ничего не придумывает и не понимает в человеческом смысле, он лишь вычисляет следующий токен, и именно поэтому управлять им нужно не уговорами, а жесткими архитектурными рамками.

Книга очень грамотно ведет читателя от простого к сложному, начиная с базовых вещей, где объясняется природа контекстного окна и почему нейросеть страдает внезапной амнезией или начинает поддакивать пользователю. Именно на этом этапе автор вводит RPEF — прикладную методику создания промптов, которая берет хаотичный набор разрозненных техник и выстраивает их в строгую систему тактических приемов. Мы проходим путь от банальных прямых запросов к методам структурированного мышления, учимся использовать цепочку мыслей, декомпозицию и метод возвращения назад. Очень круто разобраны способы самопроверки и рефлексии, а также совершенно неочевидные приемы вроде эмоционального или даже «просящего» промптинга, заставляющие бездушную статистическую машину выдавать более точные ответы. Венцом этой тактической базы выступает фреймворк GRIDFORT — отличная конструкция, позволяющая изящно упаковать роль, контекст, формат и жесткие ограничения в один мощный монолитный запрос. И самое ценное здесь — честность автора: он не продает базовый промптинг как таблетку от всех болезней, а прямо говорит, что этот подход идеально закрывает 90 процентов повседневных задач, однако неизбежно начинает сбивать и захлебываться от информационной энтропии при попытке впихнуть в него сложную бизнес-логику.

И вот здесь чтение становится по-настоящему интересным, так как автор переходит к своей методологии СТРАТУМ и концепции многослойной архитектуры. Идея отказаться от одного гигантского и нестабильного промпта-простыни в пользу конвейера из изолированных микросервисов, где каждый слой отвечает только за извлечение, анализ, верификацию или синтез, кажется глотком свежего воздуха. Вместо того чтобы надеяться на случайный удачный ответ модели, нам предлагают строить предсказуемые механизмы, где ошибки отлавливаются специальными узлами-судьями до того, как испортят финальный результат.

Отдельного уважения заслуживает смелость автора погрузить читателя в хардкорную физику процесса обучения на примере архитектуры DeepSeek. Пошаговое объяснение того, как текст превращается в векторы, как работают матрицы эмбедингов, механизмы внимания и обратное распространение ошибки, полностью переворачивает восприятие ИИ. После этих глав ты начинаешь физически осознавать, почему параметры API вроде температуры, ядерной выборки или штрафов за повторы работают именно так, понимая, что это прямое вмешательство в математику, а не просто ползунки в интерфейсе. Отсюда происходит максимально логичный переход к практической работе с OpenAI API, который рассматривается здесь не как рядовая библиотека, а как полноценный промышленный протокол управления. Автор наглядно показывает, что на этом уровне происходит смена парадигмы: ты перестаешь быть просто собеседником чат-бота и становишься оператором вычислительной среды, где весь диалог превращается в строгий JSON-объект. Детальный разбор того, как правильно выстраивать массив сообщений с их ролевой иерархией, как жестко фиксировать формат ответа через JSON Schema и как давать модели «руки» через вызов инструментов (Function Calling), дает разработчику

реальные рычаги управления конвейером. А финальный разбор протокола MCP окончательно переводит разговор из плоскости лингвистики в сферу серьезной системной интеграции и корпоративной безопасности, показывая, как изолировать данные и заставить модель работать с базами данных без риска утечек.

После прочтения этой книги возвращаться к обычному написанию длинных текстов в окно ChatGPT уже просто не хочется, потому что приходит ясное понимание всей хрупкости такого подхода. Текст требует вдумчивости, местами заставляет напрячь мозги, продираясь сквозь JSON-схемы и логику маршрутизации, но это невероятно полезно. По сути, перед нами подробная инструкция для тех, кто перерос стадию развлечений с нейросетями и готов стать настоящим архитектором, способным встраивать искусственный интеллект в реальные бизнес-процессы, делая его безопасным, контролируемым и надежным инструментом.