Материалы к книге «ИИ без магии» — bez-magii.ru

Здесь лежат свежие версии приложений 4, 5, 7, 8 и 10 книги «ИИ без магии» (Александр Провоторов, издание первое, 2026) и пустые бланки из приложения 3. В книге — снимок на момент печати, здесь — то, что обновляется вслед за продуктами.

- **Тренажёр** — тридцать три упражнения в пяти блоках, с разбором после каждого блока, и отдельное упражнение «соберите презентацию».
- **Отрывные листы** — шесть листов, каждый на своей странице: их можно распечатать и повесить рядом с рабочим местом.
- **Бланки для печати** — пустые бланки из приложения 3, каждый на своей странице.
- **Карта инструментов** — обновляемая версия приложения 4 с датой последней сверки.
- **Своя модель за полчаса** — установка программы для запуска модели на своём компьютере.
- **Установка для продвинутых** — Claude и ChatGPT на своём компьютере, n8n и OpenClaw на своём сервере, с командами и файлами настроек.

Лаборатория «Кто есть кто в мире ИИ» — отдельная страница: двенадцать названий, вопрос по каждому, разбор ошибки, повтор неверных и итоговый счёт. Она открывается с телефона, устанавливать ничего не нужно.

Ссылки вида «глава 9» и «приложение 4» ведут в книгу. Места, которые стареют быстрее прочих, — подписи настроек на листе «Где что в окне» и карта инструментов — помечены датой последней сверки.

Тренажёр: тридцать три упражнения

Книга объясняла и показывала. Здесь вы работаете сами, причём на чужом материале: на своей задаче вы не знаете правильного ответа, а здесь знаете.

Ничего открывать не нужно: все тридцать три упражнения делаются карандашом, на блок уходит минут пятнадцать. В конце тренажёра есть ещё одно, отдельное, — про презентацию: для него нужен компьютер. Разбор каждого блока — сразу после него; не подглядывайте, пока не ответили на все пункты блока.

Блок 1. Найдите ошибку в ответе

Перед вами фрагменты ответов ИИ. В каждом есть ровно одно место, которое нельзя принимать на веру. Найдите его и назовите, по какому признаку вы его поймали.

1. «По данным исследования 2023 года, 73% компаний уже используют ИИ в повседневной работе».
2. «Согласно пункту 4.2 вашего договора, вы вправе расторгнуть его за 30 дней». *(Договор вы не прикладывали.)*
3. «Итого по таблице: 1 240 + 3 180 + 960 = 5 480 рублей».
4. «Это регулируется статьёй 14.3 — там прямо сказано, что так делать нельзя».
5. «Я перепроверила свой ответ — всё верно».
6. «В вашем городе такая услуга стоит примерно 3 500 рублей».

7. *(На расплывчатый вопрос «помоги с презентацией» приходит готовая структура на десять слайдов без единого уточняющего вопроса.)*
8. «Как писал профессор Иванов в книге „Экономика внимания“, выигрывает тот, кто считает не деньги, а минуты».
9. «Компания, скорее всего, планирует выйти на этот рынок в следующем квартале».
10. «На вашей квитанции двенадцать строк, все суммы я разобрала». *(Фотография была снята под углом, нижняя часть смазана.)*

Разбор блока 1

1. Число из исследования без автора, издания и ссылки. Признак из главы 9.
2. Ссылка на документ, которого перед моделью нет: она достроила правдоподобный пункт (глава 3).
3. Арифметика: сумма равна 5 380. Если продукт не запустил калькулятор, числа порождаются, а не вычисляются (глава 7), — просите показать слагаемые.
4. Номер статьи без названия документа и даты редакции. Номера законов — самое правдоподобное, что модель умеет выдумывать.
5. Самопроверка без внешней опоры ненадёжна: модель чаще подтверждает свой текст. Работает «найди три самых сомнительных места» (глава 9).
6. Свежее и узкое одновременно: цены зависят от даты и от вашего города. Нужен поиск (глава 8).
7. Ответ на нечёткий вопрос без единой встречной просьбы уточнить — признак того, что вопрос достроили за вас.
8. Цитата, которая слишком хорошо подтверждает мысль. Проверяется существованием книги, а не красотой фразы.

9. Уверенность в том, чего никто снаружи знать не может: чужие планы и внутренние решения.
10. Машина не сказала, что часть изображения не разобрана. Лечение: «сначала прочитай вслух то, что видишь» (глава 5).

Блок 2. Что не так с этой просьбой

В каждой просьбе не хватает чего-то из пяти мест главы 6 — или она нарушает правило из части третьей. Назовите, чего именно.

1. «Сделай анализ отчёта».
2. «Ты гениальный маркетолог мирового уровня, придумай нам слоган».
3. «Напиши покороче».
4. «Посчитай, сколько мы сэкономим за год».
5. «Найди актуальную статистику по нашему рынку».
6. «Перепиши, сделай лучше».
7. «Вот фотография моего паспорта, помоги заполнить анкету».
8. «Отвечай только по фактам, ничего не выдумывай». *(Больше в просьбе ничего нет.)*
9. «Составь мне план на год по неделям».
10. «Напиши статью в стиле такого-то автора, мы опубликуем её у себя в блоге».

Разбор блока 2

1. Нет ни одного из пяти мест, кроме материала. Что должно получиться? Для кого? Что сделает результат хорошим?
2. Комплимент вместо роли. «Ты редактор делового журнала» задаёт тон и приоритеты; «гений мирового уровня» не задаёт ничего.

3. Нет критерия («вдвое»? «до трёх абзацев»?) и не сказано, что **не** трогать.
4. Счёт без данных. Дайте цифры и попросите таблицу; считать словами — не работа модели (глава 7).
5. Нет требования ссылок и дат публикации. Без дат ответ со ссылками может быть годичной давности (глава 8).
6. «Лучше» — не правка. Правка называет, что именно не так: «убери канцелярит», «третий абзац мимо».
7. Документ, удостоверяющий личность, не отправляется никогда (глава 12). Задачу решают, описав анкету словами.
8. Пятое место заполнено, остальные четыре пусты. Запрет выдумывать без материала не даёт модели ничего, на что опереться.
9. Нет ограничений и нет указания, что делать с неизвестным. Вы получите правдоподобный план из ниоткуда.
10. Это область главы 15: общая манера никому не принадлежит, а конкретные тексты и узнаваемые обороты автора — принадлежат. Для блога назовите нужный тон и приёмы, а не имя автора, и проверьте, что в тексте нет чужих узнаваемых фраз.

Блок 3. Проверьте факт за две минуты

Для каждого утверждения выберите, какую из четырёх проверок главы 9 вы примените **первой** и почему. Иногда правильный ответ — вообще не проверка ИИ, а другое действие.

1. «Правила въезда в эту страну изменились в прошлом месяце».
2. «В приложенном вами договоре нет ни одного пункта про штрафы».
3. «Средний срок доставки по этому направлению — четыре дня».
4. «Этот автор написал книгу с таким названием в 2019 году».

5. «Сумма по вашей выгрузке за квартал — 148 300 рублей».

Разбор блока 3

1. Не одна из четырёх проверок, а поиск с датой публикации. Ответ зависит от сегодняшнего дня, а память модели закончилась раньше. Потом — открыть официальный источник самому.
2. Не проверка ИИ вообще: поиск по слову внутри файла. Утверждение об **отсутствии** чего-то модель проверить не может: длинный файл она нередко видит фрагментами, а не целиком (глава 7).
3. Спросите источник. Если его нет — утверждение ничего не стоит. Если есть — открывайте: это правая часть шкалы.
4. Повтор в чистом разговоре. Расхождение между двумя попытками показывает то место, где модель сочиняет. Затем одна ссылка руками.
5. Посчитайте сами — в табличном редакторе или на калькуляторе. Ни одна из четырёх проверок здесь не поможет: это не вопрос доверия, а вопрос арифметики.

Блок 4. Обезличьте документ

Для каждого случая выпишите три списка: что убрать совсем, что заменить условным обозначением, что оставить как есть. Проверьте себя вопросом: не потеряла ли задача смысл?

1. **Претензия поставщику.** Есть: ваше ФИО, адрес, телефон, номер договора, дата, сумма 148 300 рублей, название компании-поставщика, ФИО менеджера, описание брака.
2. **Ваша медицинская выписка**, которую вы хотите разобрать перед приёмом. Есть: ФИО, дата рождения, номер полиса, диагноз, назначения, фамилия врача, название клиники.

3. **Переписка с клиентом**, по которой нужно составить вежливый ответ. Есть: фамилия клиента, его телефон, номер заказа, суть претензии, внутренние комментарии ваших коллег.

Разбор блока 4

1. **Убрать:** телефон, адрес. **Заменить:** ФИО → «Заказчик», компанию → «Поставщик А», менеджера → «представитель», номер договора → «договор №…», сумму — округлить или изменить пропорционально. **Оставить:** описание брака, даты в относительном виде («через две недели после поставки»). Смысл задачи (составить претензию) сохраняется полностью.
2. **Убрать:** ФИО, дату рождения, номер полиса, фамилию врача, название клиники. **Заменить:** ничего заменять не нужно. **Оставить:** диагноз и назначения — именно их вы и разбираете. Это ваши данные, и решение ваше. С выпиской родственника так нельзя без его согласия: данные его, и решение его (глава 12). Нет согласия — опишите ситуацию словами.
3. **Убрать:** телефон, внутренние комментарии коллег — они не нужны для ответа и могут попасть в текст. **Заменить:** фамилию → «клиент», номер заказа → «заказ №…». **Оставить:** суть претензии. Внутренние комментарии — самая частая утечка в этом случае: их забывают вырезать вместе с перепиской.

Блок 5. Выберите инструмент под задачу

Назовите **класс** инструмента, а не конкретный продукт. И объясните, почему не подходят соседние классы.

1. Каждую пятницу собирать данные из трёх систем, сводить в таблицу и рассылать четырём людям.
2. Понять, что написано в присланном PDF на сорок страниц, и найти там два условия.

3. Узнать, изменились ли за последний месяц тарифы у трёх поставщиков.
4. Помощник, который ведёт длинную переписку неделями, помнит контекст и сам решает, что делать дальше.
5. Разбирать документы, которые по договору не могут покидать контур организации.

Разбор блока 5

1. **Конструктор процессов.** Маршрут известен заранее и не меняется — агент здесь добавит стоимость и непредсказуемость, не добавив пользы (глава 29).
2. **Чат-продукт, умеющий файлы.** Ни поиск, ни агент не нужны: всё лежит перед моделью. Просите ссылки на пункты (глава 7).
3. **Чат с поиском или отвечающая система.** Ответ зависит от сегодняшней даты, значит, память не годится (глава 8). Обязательны ссылки с датами.
4. **Агентная среда** — и только на четвёртой ступени: с закрытым списком инструментов, лимитами и подтверждением человека перед каждым действием, меняющим мир (главы 21 и 32).
5. **Свой проигрыватель моделей** внутри контура. Качество будет ниже, а забот больше; зато снимается передача данных внешнему поставщику (глава 23).

Что делать с результатом

Если в каком-то блоке вы ошиблись больше чем в трети пунктов, вернитесь к главам, которые этот блок тренирует, и сделайте их практику руками. Блоки 1 и 3 — главы 3, 5, 7, 8 и 9. Блок 2 — главы 6, 10, 12 и 15. Блок 4 — глава 12. Блок 5 — главы 14, 20, 21, 23, 29 и 32.

Тяжелее всего почти всем даётся блок 4. Не потому, что он сложный, а потому, что в нём нет подсказки в виде неправильного ответа — обезличивание всегда выглядит достаточным, пока кто-то не покажет, что вы забыли.

Отдельное упражнение: соберите презентацию

Единственное упражнение тренажёра, которое делается не карандашом, а за компьютером. Оно стоит здесь отдельно, потому что презентацию у ИИ просят чаще всего, а получают обычно воду.

Задача. Десять слайдов правлению: итоги трёхмесячного пилота. Решение, которое вы хотите получить, — продолжать или закрывать. Двадцать минут, считая проверку.

Шаг 1. Не просите «сделай презентацию». Просьба без материала даёт только правдоподобные общие слова (глава 3). Сначала соберите то, что у вас уже есть: таблицу с цифрами пилота, две-три переписки, где видны возражения, и одну страницу своих заметок. Это и есть материал — второе место просьбы из главы 6.

Шаг 2. Сначала структура, потом слайды. Попросите костяк презентации — и получите возможность спорить о логике, пока это ещё дёшево.

> *Вот цифры пилота, переписка и мои заметки. Предложи структуру презентации на десять слайдов для правления. Формат каждого слайда: заголовок-утверждение (не тема, а вывод) и три пункта под ним. Цель выступления — решение «продолжаем или закрываем». Аудитория знает предметную область, но не знает деталей пилота. Не придумывай цифр, которых нет в материалах; там, где данных не хватает, так и напиши.*

Последнее предложение — пятое место просьбы. Без него пустые места заполнятся правдоподобными числами, и вы заметите это на правлении.

Шаг 3. Заголовки-утверждения. Это приём, на котором держатся консультантские презентации: слайд называется не «Результаты пилота», а «Пилот окупился на втором месяце, но только в одном из трёх отделов». Если заголовки слайдов прочитать подряд, должна получиться связная история. Проверьте так:

> *Выпиши подряд только заголовки всех десяти слайдов. Прочитай их как один текст. Где логика рвётся?*

Шаг 4. Стиль задаётся тремя способами. По возрастанию надёжности.

Словами. «Тон сдержанный, без восклицаний и маркетинга. Предложения короткие. Никаких слов „революционный“, „уникальный“, „в современном мире“».

Образцом. Приложите свою прошлую хорошую презентацию: «держись этой манеры: длина фраз, обращение к аудитории, уровень детализации». Образец работает лучше описаний, потому что вы сами не знаете всех правил своего стиля.

Шаблоном. Если в компании есть фирменный шаблон — скажите прямо: «оформление не придумывай, я вставлю текст в наш шаблон». Тогда машина не потратит усилия на то, что вы всё равно выбросите.

Шаг 5. Довести, а не переписать. Правка с указанием, что не трогать (глава 10):

> *Слайды 4 и 7 слишком общие — замени формулировки на конкретные из таблицы. Остальные слайды и порядок не меняй.*

Шаг 6. Проверить руками. Обязательно: каждое число сверить с таблицей; каждое утверждение вида «мы стали быстрее» — с источником; каждую фамилию и дату. Это правая часть шкалы из главы 9. Отдельно спросите:

> *Что в этой презентации ты знаешь по моим материалам, а что достроила сама?*

За второй список отвечаете вы.

Слово, которое вы услышите

Заголовок-утверждение называют **action title**. Правило «заголовки подряд читаются как история» — часть подхода, который в консалтинге называют **пирамидой Минто** (*Minto Pyramid*): сначала вывод, потом обоснования.

Знать эти слова не обязательно. А вот презентация, собранная так, выдерживает вопрос «а что вы хотите, чтобы мы решили?».

Чем отличаются два продукта

Claude удобнее там, где много материала: длинные документы и таблицы он держит целиком и реже теряет детали. Результат складывается в отдельное рабочее окно, которое можно править по кускам, не пересобирая всё.

ChatGPT удобнее там, где нужен быстрый визуальный черновик и картинки к слайдам. А про встройку в почту и хранилища не верьте книге на слово: список подключений у обоих продуктов меняется от обновления к обновлению, смотрите в настройках своего.

Готовый файл в фирменном оформлении иногда уже получается: некоторые продукты собирают слайды по шаблону компании. Но порядок от этого не меняется: сначала структура и текст, потом про-

Чем отличаются два продукта (продолжение)

верка и только затем слайды — в своей программе или прямо в продукте. Что именно умеет выгружать ваш продукт сегодня, проверьте сами.

Частая ошибка

Показать правлению презентацию, которую вы не прочитали целиком вслух. Машина пишет гладко, и гладкость маскирует дыры: слайд может выглядеть уверенно и не содержать ни одного проверяемого утверждения. Прочитайте вслух — пустота слышна сразу.

Что должно получиться. Десять заголовков, которые читаются как связная история и заканчиваются вопросом о решении. Ни одного числа, которого нет в ваших материалах. Отдельный список того, чего в материалах не хватило, — с ним вы идёте не оправдываться, а просить.

Отрывные листы

Шесть страниц, которые имеет смысл вырезать или скопировать и повесить. Всё, что на них есть, разобрано в книге; здесь оно собрано так, чтобы читаться с расстояния в два шага.

Каждый лист — на своей странице: распечатайте нужные.

Где что в окне

Подписи у всех продуктов свои, а места — почти одни и те же (рисунок 1.1). Ищите по смыслу, а не по слову. Сверено в ноябре 2026 года по справке ChatGPT, Claude и Gemini; в Алисе AI и GigaChat подписи и места могут отличаться.

Что нужно	Как подписано	Где искать
Приложить файл	скрепка, плюс	в строке ввода, слева
Приложить много страниц	там же	перетащите выделенные файлы в окно; для порядка назовите «01», «02». Если продукт берёт по одному файлу — соберите страницы в один PDF
Включить поиск	глобус, «Поиск», «Web»	в меню «+» у строки ввода; во многих продуктах поиск включается сам
Надиктовать	микрофон	справа в строке ввода
Новый разговор	«Новый чат», плюс	левый верхний угол
Выключить память	«Память», «Персонализация»	настройки: раздел про память или «Персонализация»
Запретить обучение	«Улучшение модели»; в Gemini — «Действия»	настройки: раздел про данные или конфиденциальность
Удалить историю	«Удалить все чаты», «История»	настройки или список чатов: отметьте и удалите
Завести проект	«Проекты», «Навыки», «Плагины»; в Gemini — Notebooks	левая колонка или меню профиля
Сменить модель	название модели	верх окна или строка ввода, рядом с кнопкой отправки

Если не нашли. Спросите прямо в чате: «где здесь выключается память и где — обучение на моих разговорах?» Ответ подскажет, что искать, но проверьте в настройках сами: про себя машина отвечает так же правдоподобно, как про всё остальное (глава 7).

Никогда не отправляют

В личный и бесплатный чат. В корпоративный ИИ по договору — только по правилам организации (глава 12); пароли и коды — никуда.

1. **Пароли и коды**
   включая проверку надёжности
2. **Карты и счета**
   номера и полные реквизиты
3. **Документы личности**
   паспорт, права, СНИЛС, фото
4. **Медицинские данные**
   чужие — только с согласия человека
5. **Данные клиентов**
   и сотрудников: фамилии, телефоны, адреса
6. **Тайна по закону**
   банковская, врачебная, коммерческая
7. **Чужие переписки**
   и записи без согласия участников

Четыре проверки за две минуты

1. **Источник**
   «откуда это? дай ссылку или пункт»
2. **Граница знания**
   «что знаешь наверняка, а что предполагаешь»
3. **Одно место руками**
   одно число, одна фамилия, одна ссылка
4. **Заново в чистом чате**
   тот же вопрос, не показывая первый ответ

Десять слов, которые всё объясняют

Десять слов, которыми описывается всё остальное. Если держать их в голове, любое новое название встаёт на место за минуту.

01 Model · модель
Вычислительное ядро: превращает контекст в ответ.

02 Product · продукт
Приложение вокруг модели: чат, файлы, поиск, память.

03 Prompt · просьба
Инструкция для одной текущей задачи.

04 Skill · навык
Сохранённый способ делать целый класс задач.

05 Tool · инструмент
Действие, которое машине дали вызывать.

06 Workflow · процесс
Маршрут шагов нарисовал человек заранее.

07 Agent · агент
Система сама выбирает следующий шаг.

08 Orchestrator · дирижёр
Решает, кто и когда работает.

09 Runtime · среда исполнения
Крутит цикл и держит состояние.

10 RAG · MCP
RAG кладёт знания перед моделью; MCP — общий разъём к ресурсам.

Три вопроса к любому новому названию. Это ядро или оболочка вокруг ядра? Кто выбирает следующий шаг — человек, заданный маршрут или машина? Что это по роли: продукт, проигрыватель моделей, платформа, протокол или приём сборки?

Ответ или похоже на ответ

Три вопроса, которые стоит задать, и то, как звучит ответ на них. Правую колонку слышно сразу — для неё не нужен инженер. Если ответа нет, не спорьте: спросите срок, когда покажут.

«В каком месте закрытый документ отсеивается — и что будет, если эта проверка не отработает?»

Ответ	Не ответ
называет этап: «на этапе поиска, до того, как текст попадёт модели»	называет намерение: «модель обучена не показывать закрытое»
при сбое — «ничего не найдётся»	«такого не может произойти»
«система ходит от имени спрашивающего»	«у неё доступ только к разрешённым папкам»

«Уберите из инструкции запрет — система сможет выполнить действие?»

Ответ	Не ответ
«не сможет: этого действия нет в списке разрешённых»	«сможет, но она так не делает»
при сбое проверки — «действие не выполнится»	«при сбое пропускаем, чтобы не останавливать работу»
«подтверждение стоит до действия, на таком-то шаге»	«подтверждение есть» — и это отчёт после
«не больше стольких шагов и стольких рублей»	«лимиты настроены»

«Сколько в проверочном наборе случаев, где правильный ответ — отказ?»

Ответ	Не ответ
называет число	«набор большой, мы всё покрыли»
называет порог и дату, когда он записан	называет число после того, как увидел результат
«вот случай, который мы не прошли, и что с ним сделали»	показывает только то, что получилось

Десять пар, которые путают

Если вы понимаете эти десять пар, вы понимаете большую часть разговоров об ИИ. Полная карта инструментов — в книге (приложение 4) и в разделе «Карта инструментов» этих материалов.

Путают	Разница в одной фразе
Модель и продукт	модель порождает текст; продукт — окно, файлы, поиск и кнопки вокруг неё
Чат и агент	в чате следующий шаг выбираете вы; агент выбирает его сам
Агент и процесс	в процессе маршрут нарисовал человек; у агента маршрута нет
Агент и дирижёр	агент делает работу; дирижёр решает, кто и когда работает
Среда исполнения и модель	среда крутит цикл и хранит состояние; думает модель
Проигрыватель и модель	проигрыватель — программа, модель — пластинка
MCP и права доступа	разъём говорит, **как** подключаться; **кому что можно**, решаете вы
Вход и права	вход (пропуск на проходной) отвечает, **кто пришёл**; права — **что этому человеку разрешено**. Проходная не открывает кабинеты
Подбор знаний и дообучение	первое кладёт документы перед моделью, второе меняет саму модель
Открытые веса и открытый код	выложены числа модели — не то же самое, что выложены код и права

Бланки для печати

Пустые бланки из приложения 3 книги — каждый на своей странице. Распечатайте нужный и заполняйте карандашом; как ими пользоваться, сказано в главах, на которые ссылается каждый бланк.

Конструктор просьбы

Заполните пять строк — и прочитайте их подряд. Получится готовый текст просьбы. Глава 6.

Место	Вопрос к себе	Ваша строка
1. Цель	Что должно получиться на выходе? Не тема, а результат.	
2. Материал	На основании чего? Файл, мой текст, общие знания.	
3. Ограничения	Длина, тон, для кого, чего не делать.	
4. Критерий	Что сделает этот ответ хорошим именно для меня?	
5. Неизвестное	Что делать, если данных не хватает?	

Если ответ вышел не тот, вернитесь к таблице и найдите пустую строку. В девяти случаях из десяти это четвёртая.

Четыре проверки за две минуты

Глава 9. Плакат на отдельном листе — в приложении 8.

1. «Откуда это? Дай ссылку или пункт».
2. «Что здесь ты знаешь наверняка, а что предполагаешь?»
3. Проверьте руками одно место — самое важное.
4. Задайте тот же вопрос в новом разговоре, не показывая первый ответ.

И приём, который работает лучше самопроверки: *«Найди в своём ответе три самых сомнительных утверждения и объясни почему»*.

Что не отправляют никогда

Глава 12. Плакат на отдельном листе — в приложении 8.

В личный и бесплатный чат. Пароли и коды · номера карт и счетов · документы, удостоверяющие личность · медицинские данные других людей · персональные данные клиентов и сотрудников без разрешения · сведения под тайной или соглашением о неразглашении · чужие переписки и записи разговоров без согласия участников.

Приём: замените названия, фамилии и суммы — смысл задачи сохранится, чувствительность исчезнет.

Три настройки, которые стоит открыть один раз

Глава 12.

1. Использование ваших разговоров для обучения — включено или нет.
2. История разговоров — хранится ли, как удалить.
3. Память о вас — что записано, как почистить.

Семейный контрольный вопрос

Глава 15. Заполните и расскажите тем, кого это касается.

Строка	Заполните
Вопрос, который мы задаём	
Ответ, который знаем только мы	
Кто в курсе	
Что делаем при неверном ответе	кладём трубку и перезваниваем сами

Дневник экспериментов

Тридцать шестое дело из главы 19 и единственное, которое стоит начать сегодня. Через месяц это ваша личная книга приёмов — точнее любой купленной, потому что она про ваши задачи.

Дата	Задача	Что попросил	Что получилось	Что менять в следующий раз

Одна строка в день. Не больше.

Сравнение двух продуктов

Глава 14. Критерии заполняются **до** того, как вы увидите ответы.

Критерий	Продукт А	Продукт Б
Выдумал ли факты		
Соблюдены ли ограничения		
Спросил ли о недостающем		
Понятность ответа		
Время до результата		
Вывод одной строкой		

Одностраничник для вашего руководителя

Часть седьмая на одной странице — для случая, когда книгу дать нельзя, а объяснить нужно. Замените то, что в скобках, а каждое утверждение без скобок проверьте: оставьте только то, что у вас уже так.

> ***ИИ в нашем отделе: что, зачем и что решаем***
>
> ***Что это.*** *Инструмент, который сокращает время на разбор документов, подготовку текстов и поиск. Не заменяет ни одного специалиста и не принимает решений.*
>
> ***Что мы уже делаем.*** *(Одна-две конкретные задачи и сколько времени они занимали раньше.)*
>
> ***Что меняется в ответственности.*** *За результат отвечает человек, который его подписал. Чужие данные во внешние сервисы не уходят; правила на одной странице у нас есть, ответственный — (имя).*
>
> ***Чего мы не делаем.*** *Не даём системе выполнять действия без подтверждения человека. Не начинаем внедрение, пока нет измеримого результата и владельца отката — человека, который при неудаче вернёт прежний порядок.*
>
> ***Что нужно от вас.*** *(Одно решение: доступ, бюджет, разрешение на пилот — с конкретной суммой и сроком.)*

Двадцать кандидатов

Глава 34. Сначала стоп-факторы: в графе пишут, какой сработал, или «нет»; сработал хоть один — кандидат выбывает. Остальных оценивайте одним словом: хорошо, средне, плохо. «Хорошо» везде значит «в пользу запуска»: большой объём и заметная выгода, дешёвая и быстро замечаемая ошибка, готовые данные с понятными правами, есть с чем сравнить, есть владелец.

Таблица не умещается в полосу, поэтому разделена надвое: первая колонка повторяется.

Кандидат	Стоп-фактор	Объём и выгода	Цена ошибки

Продолжение той же таблицы

Кандидат	Данные и права	Измеримость	Владелец

Первыми идут строки без единого «плохо»; из них — с «хорошо» в цене ошибки и измеримости; при равенстве — сначала те, что строят нужный остальным фундамент, потом — с большей выгодой. Строки без владельца в очередь пока не ставят. На листе десять строк; для двадцати кандидатов нужны два листа.

Карта инструментов

Обновляемая версия приложения 4 книги. **Последняя сверка — ноябрь 2026 года.** Когда продукты меняются, карта на этой странице обновляется и дата сверки ставится заново; в книге остаётся снимок на момент печати.

Здесь названы семейства, производители и классы — они держатся годами; номера версий сознательно не указаны: они устаревают за месяцы. Способ разбираться не устаревает вообще.

Как этим пользоваться. Не заучивайте. Для любого незнакомого названия задайте три вопроса — они собраны в конце карты, а разобраны в главах 20 и 21.

Модели — ядро, которое порождает и анализирует текст

Номера версий смотрите на сайте производителя.

Название	Производитель	Что это
GPT	OpenAI	флагманская линейка; работает внутри ChatGPT и через программный вход
Claude	Anthropic	сильна в длинных текстах, документах и работе с кодом
Gemini	Google	мультимодальная линейка; встроена в продукты Google
DeepSeek	DeepSeek	китайская линейка, сильная при низкой цене; часть моделей с открытыми весами

ПРОДОЛЖЕНИЕ

Название	Производитель	Что это
Qwen	Alibaba	большое семейство, много открытых по весам вариантов
GLM	Z.ai (Zhipu AI)	китайская линейка общего назначения
Kimi	Moonshot AI	китайская линейка, крупные модели с открытыми весами под агентные задачи
Grok	xAI	линейка с доступом к потоку публичных сообщений
GigaChat	Сбер	российская модель; где именно обрабатываются ваши данные — вопрос к договору, а не к описанию
Alice AI, YandexGPT	Яндекс	российские модели: Alice AI — новое поколение, работает внутри Алисы AI; обе линейки есть в облаке Яндекса
Llama	Meta*	открытые веса; основа множества чужих решений; новый флагман Meta с 2026 года — закрытый
Mistral	Mistral AI	европейская линейка, часть моделей с открытыми весами
Gemma	Google	открытые по весам модели для запуска у себя

* Meta Platforms Inc. признана в России экстремистской организацией, её деятельность запрещена.

Продукты — то, чем пользуется человек

Название	Производитель	Что это
ChatGPT	OpenAI	чат, файлы, поиск, изображения, голос, агентные режимы
Claude	Anthropic	чат с упором на документы, анализ и работу с текстом
Gemini (приложение)	Google	чат рядом с почтой, документами и поиском
DeepSeek	DeepSeek	простой чат к моделям DeepSeek
Алиса AI	Яндекс	голосовой и текстовый помощник; внутри — модели Alice AI
GigaChat	Сбер	чат и приложение; встроен в банковское приложение
Qwen Studio (прежде Qwen Chat)	Alibaba	чат к моделям Qwen
Copilot	Microsoft	семейство продуктов внутри Windows и офисных программ
Perplexity	Perplexity	отвечающая система: сначала ищет источники, потом отвечает

Изображения, видео, звук

Название	Производитель	Что это
Veo, Imagen	Google	видео и изображения
Midjourney	Midjourney	генерация изображений, сильна художественно
Stable Diffusion	Stability AI	открытые по весам модели изображений
Kandinsky	Сбер	российская генерация изображений; доступна в GigaChat
Шедеврум	Яндекс	российская генерация изображений
ElevenLabs	ElevenLabs	синтез и клонирование голоса

Агенты для работы с кодом

Название	Производитель	Что это
Claude Code	Anthropic	агент в папке проекта: читает, правит, запускает, проверяет
Codex	OpenAI	то же самое в экосистеме OpenAI
GitHub Copilot	GitHub / Microsoft	помощник в среде разработки; есть агентные режимы
Cursor	Anysphere	редактор кода со встроенным агентом

Агентные среды

Название	Производитель	Что это
OpenClaw	открытый проект	среда для долгоживущих агентов: сессии, память, инструменты, каналы общения
Управляемые среды агентов	OpenAI, Anthropic, Google, Microsoft	среда для агентов без сборки своей обвязки; названия меняются чаще всего в этой строке
Агентные режимы чатов	OpenAI, Anthropic, Google	привычный чат, который умеет работать над задачей минутами

Платформы процессов

Название	Производитель	Что это
n8n	n8n	визуальный конструктор процессов; можно поставить у себя
Make	Make	облачный конструктор процессов
Zapier	Zapier	связывает приложения между собой, самый простой вход
Power Automate, Copilot Studio	Microsoft	процессы и агенты внутри корпоративной экосистемы
Albato	Albato	российский конструктор интеграций

Проигрыватели моделей

Название	Производитель	Что это
Ollama	Ollama	самый простой способ запустить модель на своём компьютере
LM Studio	LM Studio	то же самое с графическим интерфейсом
llama.cpp	открытый проект	движок, на котором работает половина остальных; нужен, если хочется управлять всем самому
vLLM	открытый проект	сервер выдачи моделей под нагрузкой: ставят, когда к модели обращаются многие

Предсказательные системы — ИИ, который ставит оценки и принимает решения

Всё предыдущее на этой карте — порождающий ИИ: он создаёт текст или картинку. Но в организации вы чаще встретите другой класс — системы, которые ничего не пишут, а выставляют оценку или принимают решение.

Здесь нет громких имён: такие системы почти всегда встроены в банковские, кадровые или страховые программы и называются по имени задачи, а не модели.

Задача	Где встретится	Главный вопрос к ней
Оценка кредитоспособности	банки, микрофинансы, рассрочка	может ли человек оспорить решение?
Обнаружение мошенничества	платёжные системы, страхование	что стоит дороже — пропуск или ложная блокировка?

ПРОДОЛЖЕНИЕ

Задача	Где встретится	Главный вопрос к ней
Отбор резюме и оценка кандидатов	кадровые системы	на каких данных обучалась и кого отсеяла раньше?
Прогноз оттока и спроса	торговля, связь, подписки	с чем сравниваем — с человеком или с идеалом?
Предсказание поломки	производство, транспорт	сколько стоит ложная тревога?
Распознавание лиц и номеров	доступ, безопасность, городская среда	кто отвечает за ошибочное совпадение?

Три вещи, которые отличают этот класс от всего остального в книге.

Часто решение принимается о человеке, и он обычно даже не знает, что его оценила система.

Ошибка не выглядит ошибкой. Отказ в рассрочке выглядит как обычный отказ: предъявить нечего, значит, и жаловаться не на что.

Приёмы главы 9 здесь не работают. Система не отвечает текстом и не даёт ссылок — спросить её «откуда это?» нельзя.

Работают с ней другими инструментами книги: раздельным счётом двух типов ошибок и стоп-факторами из главы 34, проверочным набором из главы 33 и вопросами про данные из главы 31.

Не продукты: стандарты и приёмы

Название	Кто за этим стоит	Что это
MCP	открытый стандарт	общий разъём между ИИ-приложением и вашими системами
RAG	приём, не продукт	подбор кусков ваших документов в контекст перед ответом
Контрольный слой	слой в архитектуре	проверки прав, лимитов и подтверждений вокруг агента

Три строки выше — самая частая ошибка классификации в переговорных. Ни одну из них нельзя «купить и внедрить»: разъём реализуют у себя, подбор знаний собирают под свои документы, контрольный слой проектируют вместе с системой.

Что с чем путают чаще всего

Если вы понимаете эти десять пар, вы понимаете большую часть разговоров об ИИ.

Путают	Разница в одной фразе
Модель и продукт	модель порождает текст; продукт — окно, файлы, поиск и кнопки вокруг неё
Чат и агент	в чате следующий шаг выбираете вы; агент выбирает его сам
Агент и процесс	в процессе маршрут нарисовал человек, у агента маршрута нет; процесс остаётся процессом, даже если один из его шагов — агент

ПРОДОЛЖЕНИЕ

Путают	Разница в одной фразе
Агент и дирижёр	агент делает работу; дирижёр решает, кто и когда работает
Среда исполнения и модель	среда крутит цикл и хранит состояние; думает модель
Проигрыватель и модель	проигрыватель — программа, модель — пластинка
MCP и права доступа	разъём говорит, **как** подключаться; **кому что можно**, решаете вы
Вход и права	вход (пропуск на проходной) отвечает, **кто пришёл**; права — **что этому человеку разрешено**. Проходная не открывает кабинеты
Подбор знаний и дообучение	первое кладёт документы перед моделью, второе меняет саму модель
Открытые веса и открытый код	выложены числа модели — не то же самое, что выложены код и права

Три вопроса к любому новому названию

1. **Это ядро или оболочка?** Модель или что-то, построенное вокруг модели.
2. **Кто выбирает следующий шаг?** Человек, заранее заданный маршрут или сама машина.
3. **Что это по роли?** Продукт, проигрыватель моделей, платформа, протокол или приём сборки.

Три ответа — и вы знаете, с чем имеете дело, даже если название слышите впервые.

Своя модель за полчаса

Кроме «Установки для продвинутых», это единственный материал к книге, где нужно открыть чёрное окно и набрать команду. Читать не обязательно: всё, чему учит книга, работает и без этого.

Но если после главы 23 захотелось потрогать руками — вот честный минимум. Полчаса, ни одной покупки, ничего не ломается.

Что понадобится

Обычный компьютер — ноутбук последних лет подойдёт. Чем больше оперативной памяти, тем крупнее модель поместится: при восьми гигабайтах работают небольшие модели, при шестнадцати — заметно лучше. Меньше восьми — не стоит и начинать: модель либо не запустится, либо будет отвечать по слову в секунду.

Свободного места на диске нужно от трёх до двадцати гигабайт: столько весит сама модель. Выбирать её надо по памяти, а не по диску: ориентир — файл модели не больше половины оперативной памяти. Возьмёте крупнее — скорее всего, запустится, но часть придётся считать на процессоре, и ответ пойдёт по слову в секунду. На «Маках» память общая с видеокартой, там запас нужен меньше.

Интернет нужен один раз, чтобы скачать модель. Дальше — нет.

Четыре шага

1. Скачайте проигрыватель. Наберите адрес `ollama.com` руками, а не переходите по ссылке из письма или рекламы (глава 1), и скачайте версию для своей системы. Это обычная программа: скачали, запустили установщик, согласились.

2. Откройте окно команд. В Windows это «Терминал» или «Командная строка», в macOS — «Терминал». Проще всего найти его поиском по названию. Чёрное окно с курсором — это оно.

3. Наберите одну строку. Вот она целиком, с конкретным названием — такую можно набрать не думая:

```
ollama run qwen3:4b
```

Здесь `qwen3` — семейство модели, а `4b` — её размер, четыре миллиарда параметров. Это как раз небольшая модель, которая помещается в восемь гигабайт памяти.

Названия версий живут месяцами, поэтому одно уточнение: **если строка выдаст ошибку вроде «model not found» — скорее всего, этой версии больше нет в списке или в команде опечатка.** Откройте раздел с моделями на сайте Ollama, возьмите оттуда любое имя небольшой модели и подставьте его вместо `qwen3:4b`. Только не берите имя со словом `cloud`: такая модель работает не у вас, а на серверах Ollama, и ваши вопросы уходят туда. Больше в команде менять нечего.

Первый запуск скачает модель: несколько гигабайт, от двух до двадцати минут. Дальше она запускается мгновенно.

Как понять, что получилось. Сначала побежит полоса загрузки с процентами. Потом появится строка `>>>` и мигающий курсор — это приглашение, всё готово. Если вместо этого вы видите сообщение об ошибке — вернитесь к уточнению про «model not found»: дело почти всегда в названии.

4. Пишите. Печатайте вопрос прямо после `>>>`, как в обычном чате, и нажимайте ввод. Первый ответ пойдёт медленнее обычного: модель прогревается. Выйти: наберите `/bye`.

Что вы увидите

Три вещи, ради которых это стоило делать.

Оно работает без интернета. Отключите сеть и спросите ещё раз. Ответ придёт.

Оно заметно слабее. Задайте тот же вопрос, что и привычному чату. Разница в качестве будет видна сразу — это цена того, что данные никуда не уходят (глава 23).

Никакой магии внутри нет. На вашем компьютере лежит файл с числами и программа, которая их крутит. Всё, о чём говорит глава 3, помещается в одну папку.

Чего это не даёт

Не даёт безопасности «из коробки»: кто имеет доступ к этому компьютеру, что пишется в журналы, откуда скачаны веса — всё это по-прежнему ваши вопросы (глава 23).

Не даёт корпоративного решения. То, что вы поставили за полчаса на ноутбуке, и то, чем пользуется отдел, отличаются как велосипед и автопарк: второе требует сервера, человека и порядка обновлений.

И не даёт повода спорить с теми, кто пользуется чужим сервисом: это разные способы для разных случаев.

Осторожно

Если ноутбук рабочий — скорее всего, у вас нет прав ставить на него программы, и это не поломка, а политика вашей организации.

Не обходите запрет и не ставьте ничего «просто попробовать»: спросите тех, кто отвечает за технику, или пробуйте на домашнем компьютере.

Установка для продвинутых

Это самый технический из материалов к книге. Он написан для тех, кто ставит инструменты и себе, и отделу: для продвинутых пользователей и для тех, кто отвечает за ИТ.

Claude и ChatGPT здесь — приложения для компьютера: модель при этом работает у поставщика, и данные уходят к нему. Своя модель на своём компьютере — в материале «Своя модель за полчаса».

Остальная книга объясняла, что происходит. Здесь — что нажимать. Порядок один и тот же: сначала команда или шаг, сразу за ним — зачем он нужен и что сломается, если его пропустить.

Осторожно

Названия пунктов меню, номера портов и имена образов меняются чаще, чем печатаются книги. Всё, что ниже, проверялось при последней сверке. Если шаг не сходится, ищите его на официальной странице продукта, а не на первом попавшемся сайте с инструкцией: подменённый образ или чужой скрипт установки — самый дешёвый способ отдать посторонним весь свой сервер.

Часть первая. Claude на своём компьютере

Что мы ставим. Настольное приложение — это то же самое, что сайт, плюс три вещи, которых на сайте нет: доступ к файлам на вашем диске, подключения к программам на вашем компьютере и вызов из любого места по горячей клавише. Подключения к облачным сервисам есть и на сайте. Если ничего из этого не нужно, браузера достаточно.

Шаг 1. Установка. Скачайте приложение со страницы claude.ai/download и установите обычным способом: на macOS перетащите в «Программы», на Windows запустите установщик. *Почему со страницы продукта, а не из поиска:* приложение популярное, и под него делают подделки.

Шаг 2. Вход и тариф. Войдите под своей учётной записью. Бесплатного режима хватит, чтобы всё настроить и попробовать; ограничения вы почувствуете на длинных документах. *Что важно для отдела:* у тарифа для команды другие условия обработки данных и другие права администратора. Это договорные условия, и выяснять их нужно по договору — до того, как отдел начнёт класть в чат рабочие файлы.

Шаг 3. Три настройки, которые открывают один раз. В настройках найдите и осознанно выберите: используются ли ваши разговоры для улучшения моделей (раздел приватности), помнит ли вас Claude между разговорами (раздел про память) и что делать с историей — её удаляют сами. *Зачем:* это те самые три вопроса из главы 12. После этого к настройкам можно не возвращаться месяцами.

Шаг 4. Проекты. Проект — это папка с постоянным содержимым: файлы, которые нужны всегда, и инструкция, как с ними работать. Заведите один проект под свою повторяющуюся работу и положите туда образцы: как у вас оформляют документы, что считается хорошим тоном в письмах, какие сокращения приняты. *Что это даёт:* вы перестаёте объяснять одно и то же в каждом новом разговоре. Это и есть один раз накрытый стол из главы 11.

Шаг 5. Навыки. Навык — это сохранённый набор указаний для одной повторяющейся работы, который подхватывается сам, когда задача под него подходит. Сначала включите в настройках создание и правку файлов (Create and edit files): без этого навыки не работают. Сами навыки заводятся в разделе Customize → Skills («Настроить» → «Навыки»); там же лежат уже готовые. *Как понять, что*

навык нужен: если вы третий раз пишете одну и ту же длинную преамбулу — её пора превратить в навык.

Шаг 6. Подключения к своим данным. Приложение умеет разговаривать с другими программами через общий разъём из главы 27. Часть подключений ставится в один щелчок из каталога расширений. Остальные описываются в файле настроек.

Файл лежит здесь:

```
macOS:
  ~/Library/Application Support/Claude/claude_desktop_config.json
Windows:
  %APPDATA%\Claude\claude_desktop_config.json
```

Одно условие до того, как править файл: `npx` — это часть Node.js, и без него подключение молча не поднимется. Проверьте в окне команд `node --version`: если версии нет, поставьте Node.js с `nodejs.org` (адрес набирайте руками) и начните заново.

Простейшее содержимое — доступ к одной папке на диске:

```
{
  "mcpServers": {
    "files": {
      "command": "npx",
      "args": ["-y", "@modelcontextprotocol/server-filesystem",
               "/Users/ваше-имя/Documents/Работа"]
    }
  }
}
```

Это путь для macOS. В Windows путь пишется с удвоенными обратными косыми — иначе файл перестаёт быть правильным JSON и приложение не прочитает его целиком:

```
"C:\\Users\\ваше-имя\\Documents\\Работа"
```

Разбор по строкам. `mcpServers` — список подключений; их может быть несколько. `files` — имя, которое увидите вы; оно произвольное. `command` и `args` — какую программу запустить, чтобы она обслуживала это подключение. Последний аргумент — **единственная папка**, к которой открывается доступ. Доступ этот не только на чтение: внутри указанной папки приложение может создавать и изменять файлы. Поэтому папка должна быть рабочей, а не единственной копией чего-то важного.

Осторожно

Указывайте конкретную рабочую папку, а не домашний каталог и не корень диска. Подключение даёт доступ ко всему, что внутри указанного пути, и этот доступ переживёт ваше хорошее настроение. Правило из главы 31 действует и на своём компьютере: права выдаются под задачу, а не «чтобы не мешало».

После правки файла приложение нужно закрыть и открыть заново — настройки читаются при запуске.

Шаг 7. Рутинные задачи. Соберите связку: проект с материалами + навык с инструкцией + подключение к папке. Дальше утренняя работа выглядит как одна фраза: «собери сводку по вчерашним файлам в папке „Входящие“ по нашему шаблону». *Что проверить в первый месяц:* открывайте результат целиком, а не по диагонали. Связка экономит время только тогда, когда вы ей доверяете обоснованно.

Шаг 8. Проверка после настройки. Откройте файл настроек и список подключений и сверьте со своим списком — это и есть проверка. Ответ самого приложения на вопрос «к каким папкам

у тебя доступ» сюда не годится: он порождён так же, как любой другой текст. Увидели лишнее — уберите и перезапустите.

Часть вторая. ChatGPT на своём компьютере

Порядок тот же, слова другие.

Шаг 1. Установка. Приложение для macOS, Windows и Linux скачивается со страницы chatgpt.com/download. Есть версия для телефона, и она синхронизирует те же разговоры.

Шаг 2. Вход и тариф. Бесплатного режима хватает для настройки. Для отдела работает то же замечание: условия обработки данных на командном тарифе другие, и читать их нужно до, а не после.

Шаг 3. Настройки данных. В настройках найдите раздел про управление данными и решите два вопроса: используются ли ваши разговоры для обучения и что делать с историей: сама она не стирается, её удаляют вручную, а для разговора без следа есть временный чат. А память, если она вам мешает, отключается рядом, в разделе «Персонализация».

Шаг 4. Пользовательские инструкции. Это короткий текст, который добавляется к каждому вашему разговору: кто вы, чем занимаетесь, как хотите получать ответы. *Что сюда писать:* не «будь полезным», а проверяемые вещи — «пиши без вступлений», «максимум три абзаца», «если чего-то не знаешь, скажи прямо». Последнее — то самое пятое место просьбы из главы 6, вынесенное в постоянные настройки.

Шаг 5. Проекты. Как и у Claude, это папка с постоянными файлами и своей инструкцией. Один проект на одну повторяющуюся работу — иначе материалы начнут мешать друг другу.

Шаг 6. Свой помощник. Настраиваемый помощник собирается из трёх вещей: описания задачи, файлов со знаниями, разрешённых действий. В ChatGPT таких помощников, по объявлению OpenAI, к концу 2026 года заменят плагинами — подключаемыми наборами, где описание задачи становится навыком, а файлы — справочными материалами. Порядок перехода смотрите в справке OpenAI; ту же работу уже сейчас делает проект из шага 5. *Чего ожидать:* первая версия почти всегда слишком общая. Доводится она не переписыванием описания заново, а точечными правками — ровно как в главе 10.

Шаг 7. Подключение к рабочим программам. Готовые подключения к почте, календарю и хранилищам ставятся из настроек приложения. *Правило то же:* подключайте по одному и проверяйте, что именно стало видно. «Подключили всё сразу» — это описание аварии, а не настройки.

Часть третья. Свой сервер: n8n

Дальше считаем, что у вас есть виртуальная машина с Ubuntu, доступ по SSH и права администратора. Минимум для спокойной работы: два вычислительных ядра и четыре гигабайта памяти.

Шаг 1. Docker. Ставим то, в чём всё будет работать:

```
sudo apt update && sudo apt upgrade -y
curl -fsSL https://get.docker.com | sudo sh
sudo usermod -aG docker $USER
```

Зачем: Docker — это способ запускать программу вместе со всем, что ей нужно, не засоряя систему. Третья строка разрешает вам работать с ним без `sudo`; чтобы она подействовала, выйдите из сессии и зайдите снова. Знайте, на что идёте: право работать с Docker без

`sudo` равно праву администратора на этой машине. Если так нельзя — не выполняйте третью строку и работайте через `sudo`.

Вторая строка — тот самый случай, когда скрипт из сети запускают с правами администратора. Здесь это официальный скрипт Docker с его собственного адреса, и риск небольшой. Но сам Docker называет этот скрипт непригодным для промышленных установок. Поэтому для сервера, на котором будет работать отдел, правильнее поставить Docker из его официального репозитория для Ubuntu — по инструкции на docs.docker.com. Пакеты `docker.io` и `docker-compose-v2` из репозитория дистрибутива Docker считает неофициальными: берите их, только если внутренние правила запрещают сторонние репозитории. Получится то же самое, только версии будут старше.

Шаг 2. Папка и файл настроек.

```
mkdir ~/n8n && cd ~/n8n
```

Создайте файл `compose.yml`:

```
services:
  n8n:
    image: n8nio/n8n
    restart: always
    ports:
      - "127.0.0.1:5678:5678"
    environment:
      - N8N_HOST=avtomatizacia.vashdomen.ru
      - N8N_PROTOCOL=https
      - N8N_WEBHOOK_URL=https://avtomatizacia.vashdomen.ru/
      - N8N_PROXY_HOPS=1
      - GENERIC_TIMEZONE=Europe/Moscow
      - N8N_ENCRYPTION_KEY=длинная-случайная-строка   # см. ниже
    volumes:
      - n8n_data:/home/node/.n8n
volumes:
  n8n_data:
```

Разбор. `image` — откуда берётся программа; берём с официального адреса. `restart: always` — поднимется сама после перезагрузки сервера. В строке портов важна приставка `127.0.0.1`: она означает, что снаружи напрямую никто не достучится; войти можно будет только через веб-сервер, который мы поставим на шаге 4. `N8N_ENCRYPTION_KEY` — ключ, которым шифруются все пароли и токены, которые вы сохраните в подключениях. **Придумайте длинную случайную строку и сохраните её отдельно**: потеряете ключ — потеряете все сохранённые подключения. `N8N_WEBHOOK_URL` нужен, чтобы внешние программы знали, куда стучаться. `N8N_PROXY_HOPS=1` сообщает n8n, что перед ним стоит один веб-сервер — тот, что появится на шаге 4. Если вы закрепили версию n8n старше 2.30, эта переменная называется `WEBHOOK_URL`; новые версии старое имя ещё понимают, но предупреждают о нём при запуске. `volumes` —

место, где хранятся ваши сценарии; без него они исчезнут при первом же обновлении.

Две вещи делают сразу, а не потом. Первая — закрыть файл с ключом от посторонних:

```
chmod 600 compose.yml
```

Вторая — заменить в строке `image` плавающий тег на конкретный номер версии; номер возьмите на странице релизов проекта. С плавающим тегом очередное обновление может перевести вас через мажорную версию и сломать готовые сценарии — а книга в главе 30 сама требует версионировать даже текст инструкции.

И не запустите файл с ключом из примера: `длинная-случайная-строка` — это место, куда вы подставляете свою строку, а не значение.

Шаг 3. Запуск.

```
docker compose up -d
docker compose ps
```

Вторая команда должна показать в графе состояния `Up` и время работы.

Шаг 4. Домен и шифрование. До этого шага направьте свой домен A-записью на адрес сервера и убедитесь, что снаружи открыты порты 80 и 443: сертификат выдаётся по проверке домена, и без этого шаг не пройдёт. Дальше ставим веб-сервер, который получит сертификат сам. В репозитории Ubuntu он есть, но версия там старая; берём официальный репозиторий, чтобы получить свежую:

```
sudo apt install -y debian-keyring debian-archive-keyring \
  apt-transport-https curl
curl -1sLf 'https://dl.cloudsmith.io/public/caddy/stable/gpg.key' \
  | sudo gpg --dearmor \
    -o /usr/share/keyrings/caddy-stable-archive-keyring.gpg
curl -1sLf \
  'https://dl.cloudsmith.io/public/caddy/stable/debian.deb.txt' \
  | sudo tee /etc/apt/sources.list.d/caddy-stable.list
sudo apt update && sudo apt install -y caddy
```

Дальше в файл `/etc/caddy/Caddyfile` кладём три строки:

```
avtomatizacia.vashdomen.ru {
    reverse_proxy 127.0.0.1:5678
}
```

Перезапускаем: `sudo systemctl reload caddy`. *Что произошло:* Caddy сам получил сертификат и теперь шифрует соединение. Без этого шага вы будете вводить пароли и ключи в открытом виде.

Пока вы здесь, закройте всё остальное — две команды:

```
sudo ufw allow OpenSSH && sudo ufw allow 80,443/tcp \
  && sudo ufw --force enable
sudo ss -tlnp | grep -vE '127\.0\.0\.|\[::1\]'
```

Первая оставляет снаружи вход по SSH и веб-сервер. Одной её мало: порт, который Docker публикует наружу, ufw не закрывает — Docker пишет свои правила в обход него. Поэтому в строке портов и стоит `127.0.0.1`. Вторая команда это проверяет: она показывает, что слушает наружу, и в выводе должны остаться только 22, 80 и 443. Если видите там 5678 — приставка `127.0.0.1` в строке портов потерялась, вернитесь к шагу 2.

И два дела на сервере, которые стоят пяти минут один раз: вход по SSH только по ключу (пароль отключается строкой `PasswordAuthentication no` в `/etc/ssh/sshd_config`) и второй фактор у учётной записи владельца. В n8n он есть и включается в её настройках.

Шаг 5. Первый вход. Откройте свой адрес в браузере и заведите учётную запись владельца. Она создаётся один раз и на первом же открытии — не откладывайте, иначе это сделает за вас первый, кто найдёт адрес.

Шаг 6. Подключение моделей. В сценарии добавьте узел нужного поставщика и вставьте ключ.

Поставщик	Где взять ключ	Что указать дополнительно
Anthropic (Claude)	platform.claude.com, раздел ключей	ничего
OpenAI (GPT)	platform.openai.com, раздел ключей	ничего
DeepSeek	platform.deepseek.com	ничего: берите узел DeepSeek, он есть в списке

Почему это вообще стоит знать: многие поставщики отвечают по одному и тому же соглашению, поэтому к ним годится один и тот же узел — меняются только адрес и ключ. Если своего узла в списке нет, берите узел OpenAI и впишите адрес поставщика в поле базового адреса точно так, как он указан в его документации (часто он кончается на `/v1`). Эта же механика объясняет, почему сменить поставщика в готовом сценарии часто — дело двух полей.

Осторожно

Ключ доступа — это не пароль от вашего личного кабинета, а платёжное поручение: кто его получил, тот тратит ваши деньги. Заведите отдельный ключ на каждую систему, поставьте лимит расходов в личном кабинете поставщика и никогда не кладите ключ в текст сценария — только в хранилище учётных данных.

Шаг 7. Что закрыть до первого рабочего сценария. Резервная копия папки с данными по расписанию. Порядок обновлений — кто и когда обновляет образ. Список тех, кому заведены учётные записи. И ответ на вопрос из главы 30: как вы узнаете, что сценарий сломался, если это случится ночью.

Часть четвёртая. Свой сервер: OpenClaw

Чем отличается от n8n. В n8n вы сами рисуете последовательность шагов. В OpenClaw вы даёте задачу, а порядок шагов он выбирает сам. Отсюда и разница в осторожности: маршрут делает то, что нарисовано, а агент — то, что счёл нужным.

Шаг 1. Получаем исходники и запускаем. Нужно то же, что в третьей части: сервер с Ubuntu, доступ по SSH, права администратора, установленный Docker и `git`. Брандмауэр — как в шаге 4 третьей части. Установка идёт скриптом из репозитория проекта, поэтому сначала его нужно склонировать:

```
git clone https://github.com/openclaw/openclaw.git
cd openclaw
export OPENCLAW_IMAGE="ghcr.io/openclaw/openclaw:latest"
./scripts/docker/setup.sh
```

Третья строка говорит скрипту брать готовый образ из реестра GitHub, а не собирать его у вас. Для сервера отдела замените `latest` на номер версии со страницы релизов проекта — по той же причине, что в третьей части: плавающий тег может молча перевести вас на новую мажорную версию. Дальше скрипт задаст вопросы, соберёт настройки и поднимет шлюз.

Почему не «просто docker run»: у системы несколько частей, и скрипт связывает их между собой; ручная сборка имеет смысл, только если вы понимаете, что именно меняете.

Шаг 2. Ключи поставщиков. Отвечайте на вопросы скрипта — он спросит ключи сам. Не вводите их строкой в терминале: всё, что набрано в командной строке, остаётся в истории оболочки, а если ключ попал в аргументы запускаемой программы — его видно и в списке процессов. Если установка идёт без вас, положите их в файл `.env` с правами `600` и передайте его при запуске (`docker compose --env-file .env ...` или строкой `env_file` в `compose.yml`). Сами переменные такие:

```
OPENAI_API_KEY=...
ANTHROPIC_API_KEY=...
OPENCLAW_GATEWAY_TOKEN=...
```

Последняя — пароль к самому шлюзу. Обычно его генерирует сам установочный скрипт и кладёт в `.env`: тогда вам нужно не придумать его, а прочитать оттуда и хранить как пароль от сервера, потому что это он и есть.

Шаг 3. Доступ. Панель управления открывается по адресу `http://127.0.0.1:18789/`. Наружу этот порт не открывают, а ufw, как вы помните из третьей части, порт Docker не закроет. Поэтому откройте файл compose, который собрал скрипт, найдите у шлюза строку портов и допишите в её начало `127.0.0.1:`, затем выполните

`docker compose up -d`. Проверьте той же командой `ss` из третьей части: порта 18789 в выводе быть не должно. Если работать нужно удалённо, пробрасывайте порт через SSH:

```
ssh -L 18789:127.0.0.1:18789 вы@ваш-сервер
```

После этого панель доступна у вас в браузере, а снаружи порт закрыт.

Шаг 4. Состояние. Всё, что агент помнит и умеет, лежит в каталоге `/home/node/.openclaw` внутри контейнера. Сам контейнер пересоздаётся при каждом обновлении, поэтому резервировать надо не его, а то место на сервере, которое в этот каталог подключено: посмотрите в собранном скриптом `compose.yml`, какой том или какая папка стоит в строке `volumes`, и копируйте её. Проверьте копию восстановлением хотя бы раз — иначе это не копия, а надежда.

Шаг 5. Ворота перед действием. Начните с того, что у агента уже есть: в установке по умолчанию он работает с оболочкой и файлами внутри своего контейнера, и всё, что вы подключите к нему с сервера, становится ему доступно. Поэтому первый вопрос здесь тот же, что вы задаёте разработчикам в главе 28: какую папку он видит и что может изменить. Дальше — четверо ворот из главы 32: есть ли у него право, ограничен ли список действий, стоит ли лимит, кто подтверждает необратимое. У поставщика есть кому сказать «мы это не разрешаем» — на своём сервере такого человека нет.

Что проверить у любой такой установки

Шесть вопросов, одинаковых для всех четырёх частей этой инструкции. Если хоть на один нет ответа, установка не закончена.

1. Кто имеет доступ и как он отзывается, когда человек уходит из компании.

2. Где лежат ключи и кто может их прочитать.
3. Что и куда пишется в журналы — и не попадают ли туда те самые данные из главы 12.
4. Кто обновляет это через полгода и что произойдёт, когда он уволится.
5. Есть ли резервная копия и восстанавливали ли вы из неё хоть раз.
6. Как вы узнаете о поломке раньше, чем о ней расскажет пользователь.