

За пределами ОЧЕВИДНОГО:

Как многомерные модели
раскрывают скрытые
закономерности



ИИ

Арсений Грозовой

Арсений Грозовой

**За пределами очевидного: Как
многомерные модели раскрывают
скрытые закономерности**

«Автор»

2026

Грозовой А.

За пределами очевидного: Как многомерные модели раскрывают скрытые закономерности / А. Грозовой — «Автор», 2026

Почему люди начинают делать не то, чего от них ждут, а то, что выгодно системе? И почему данные порой уверенно ведут к неверным выводам? Часто ответ скрыт в том, как мы описываем реальность: какие признаки учитываем, а какие оставляем за пределами модели. «За пределами очевидного» объясняет, как многомерные модели помогают обнаруживать закономерности и где заканчиваются их возможности. Вы узнаете, как готовить данные, измерять сходство, сокращать размерность, искать скрытые факторы и группы, строить модели регрессии и классификации. Практические разборы научат проверять качество и устойчивость моделей, интерпретировать результаты, замечать переобучение и неопределённость, анализировать изменения во времени, отличать прогноз от причины и учитывать цену ошибки — от выбора признаков до внедрения решения. Книга адресована аналитикам, руководителям, студентам и всем, кто принимает решения на основе данных.

© Грозовой А., 2026

© Автор, 2026

Содержание

За пределами очевидного: какие вопросы помогают решать многомерные модели	6
Данные как пространство: признаки, измерения и способы представления реальности	11
Геометрия сходства: расстояния, направления и близость объектов	16
Подготовка данных: пропуски, выбросы, шкалы и кодирование признаков	21
Конец ознакомительного фрагмента.	25

Арсений Грозовой
За пределами очевидного: Как
многомерные модели раскрывают
скрытые закономерности

За пределами очевидного: Как многомерные модели раскрывают скрытые закономерности

Арсений Грозовой

За пределами очевидного: какие вопросы помогают решать многомерные модели

На экране — город в огне, над ним сгущается тьма. Ваша задача, кажется, проста и велика: спасти мир. Но за каждого побеждённого противника игра начисляет опыт, за найденные сундуки выдаёт добычу, за быстрый проход открывает достижения. А за жителя, которого вы успели вывести из опасного района, — ничего. Через пару часов вы уже не столько спасаете город, сколько прокладываете самый выгодный маршрут к следующему уровню.

Это не обязательно ошибка разработчиков и не обязательно свидетельство чьего-то цинизма. Это повод задать более точный вопрос: чему на самом деле учит система — если смотреть не на её сюжет, а на доступные действия, ограничения и награды?

Такой вопрос полезен далеко за пределами игр. Многомерные модели помогают разбирать среды, где результат складывается сразу из нескольких факторов: ресурсов, времени, правил, доступных возможностей и стимулов. Они позволяют увидеть разницу между тем, что система провозглашает, и тем, что она делает выгодным. А значит, понять не только *что* происходит, но и *почему именно так*.

Сюжет — ещё не система

Возьмём знакомую ролевую игру: надвигается беда, главный герой должен защитить город. Если пересказать её сюжет, получится история о храбрости и спасении. Если описать её как систему, появятся другие вопросы:

- Какие действия доступны игроку?
- Какой ресурс ограничен?
- За что начисляются очки и награды?
- За какие действия игрок теряет время, здоровье или доступ к новым возможностям?
- Что система не разрешает сделать?
- Какой результат она считает успехом?

Вместе эти вопросы образуют не моральную оценку игры, а карту её устройства.

Представим игрока в конкретный момент. У него есть запас здоровья, несколько лечебных зелий, оружие определённой силы, ограниченное место в инвентаре, маршрут до города и время до наступления ночи. Его состояние можно описать как набор измерений: здоровье, запас ресурсов, время, доступ к локациям, очки опыта, уровень угрозы. Ни одно из них в отдельности не объясняет выбор полностью. Игрок может свернуть за сокровищем, потому что оно повышает будущую силу, или отказаться от него, потому что времени до ночи почти не осталось.

Вот почему многомерная модель — это не просто длинный список показателей. Важно увидеть связи между ними. Дополнительная добыча может дать преимущество в бою, но отнять время. Тяжёлое снаряжение может повысить защиту, но замедлить передвижение. Спасение жителя может быть важным для сюжета, но не приносить очков и требовать расхода редких предметов. Модель помогает разглядеть не отдельные числа, а компромиссы между ними.

Карта правил

Для первого разбора пригодится инструмент «Карта правил». В ней шесть колонок. Они удерживают внимание не только на официальной цели, но и на том, как устроены реальные стимулы.

Заявленная цель

Фактическая награда

Дефицитный ресурс

Запрещённое действие

Скрытая цена

Тип поведения, который система закрепляет

Защитить город до рассвета

Опыт, добыча и отметки за побеждённых противников; спасение жителей почти не учитывается

Время до рассвета, здоровье, место в инвентаре

Нельзя договориться с врагом или эвакуировать район целиком: такие действия не предусмотрены

Поиск добычи отнимает время; обходной путь может оставить часть города без защиты

Сражаться, собирать ресурсы и выбирать маршрут по измеримым наградам

Найти источник угрозы

Доступ к новой локации и предмет, открывающий следующий этап

Ключи и расходуемые предметы

Нельзя пройти в закрытую зону без нужного ключа

Исследование побочных комнат увеличивает риск не успеть к основному событию

Исследовать настолько глубоко, насколько это окупается

Сохранить силы к финальному бою

Более лёгкое прохождение босса, если зелья не потрачены раньше

Лечебные зелья и прочность снаряжения

Нельзя восстановить ресурсы в любой момент

Использование предмета сейчас может лишить игрока преимущества позже

Планировать запасы и откладывать расход ресурсов

Здесь видна полная цепочка. **Цель** сообщает, к какому исходу стремиться. **Ресурс** показывает, чем приходится распорядиться. **Ограничение** очерчивает границы возможного. **Награда** указывает, какие действия система замечает и оплачивает. **Санкция** обозначает цену ошибки: потерю здоровья, времени, доступа или самой попытки. Наконец, правила доступа определяют, какие действия вообще можно предпринять.

Это различие принципиально. Если действие нельзя совершить, оно не конкурирует с другими стратегиями. Если оно возможно, но дорого, система не запрещает его буквально — она делает его менее привлекательным. А если за него ещё и не дают награды, оно становится почти невидимым при быстром, целевом прохождении.

Полезно также смотреть на одну и ту же систему с трёх позиций. **Пользователь** замечает, что можно сделать сейчас и чем придётся пожертвовать. **Администратор** видит настройки доступа, правила начисления очков и последствия нарушений. **Аналитик** сопоставляет правила с наблюдаемыми результатами: кто выбирает какие действия, в каких условиях и с какими последствиями. Эти роли могут обнаружить разное. Пользователь чувствует дефицит времени; администратор видит, что таймер общий для всех; аналитик замечает, что игроки с определённым снаряжением чаще успевают на побочные задания.

Вопросы, которые открывает модель

Такая карта помогает перейти от расплывчатого «игроки ведут себя эгоистично» к проверяемым вопросам.

Как система производит результат? Не просто сколько игроков дошли до финала, а какой маршрут к нему оказался наиболее выгодным и кому он был доступен.

Какие факторы действуют вместе? Возможно, ограничение времени само по себе не мешает спасти жителей. Но в сочетании с малым запасом здоровья и наградой за скорость оно делает это решение менее привлекательным.

Одинаково ли устроен путь для всех? Игрок с запасом зелий может выбрать рискованный маршрут, а игрок без них — нет. Средний показатель прохождения скроет эту разницу.

Где возникает компромисс? Сбор ресурсов сегодня может повысить шансы завтра, но отнять время, необходимое для выполнения основной задачи.

Что случится, если поменять одно правило? Этот вопрос ведёт от описания системы к проверке возможных причин.

Многомерная модель ценна именно тем, что не сводит сложный результат к одной удобной цифре. Высокий счёт может быть следствием мастерства, удачного маршрута, мощного снаряжения или повторного прохождения уже известного участка. Одинаковый итог не означает одинакового пути. Если измерять только счёт, различия между стратегиями исчезнут.

Есть и ещё одна опасность: измеримый показатель легко принять за саму цель. Очки опыта — не спасённый город. Число выполненных заданий — не обязательно качество прохождения. Когда показатель становится целью, участники могут оптимизировать именно его, даже если он лишь приблизительно отражает то, что было важно изначально. Это часто связывают с законом Гудхарта: хорошая мера может перестать быть хорошей мерой, когда её начинают использовать как цель. В игровом примере проблема особенно наглядна: если очки дают за победы, система хорошо измеряет победы, но плохо — заботу о жителях.

Мысленный эксперимент: меняем одно правило

Вернёмся к игре. Не меняя сюжет, карту, противников, время до рассвета и доступное снаряжение, изменим только правило начисления очков. Теперь дополнительные баллы дают не за побеждённых врагов, а за жителей, которых удалось вывести из опасной зоны.

Что произойдёт?

Игроки, вероятно, начнут внимательнее осматривать районы, искать короткие маршруты эвакуации, охранять проходы и планировать, в какой последовательности помогать людям. Некоторым станет выгодно отложить сражение. Другие постараются спасти тех, кого проще всего найти, и проигнорируют дальние или плохо обозначенные зоны. Один новый критерий вознаграждения изменил привлекательность действий, но не устранил остальные ограничения.

Это изменение не делает стратегию автоматически «хорошей» или «плохой». Оно показывает, как система перераспределяет внимание. Более того, новая награда может создать собственную лазейку: если очки начисляют за количество спасённых людей, но не за то, в каком состоянии они добрались до безопасного места, игроки будут оптимизировать количество. Это не доказательство дурных намерений. Это напоминание: поведение зависит не только от ценностей участников, но и от того, как устроен критерий успеха.

Мысленный эксперимент полезен, однако его не следует путать с доказательством. В настоящей игре на выбор влияют привычки, уровень сложности, навыки и множество других факторов. Чтобы понять реальный эффект изменения правила, нужно сравнить наблюдения при сопоставимых условиях. Одна модель формулирует гипотезу; данные помогают проверить, выдерживает ли она столкновение с реальностью.

От игры — к рабочему месту

Та же логика встречается в обычных организациях. Она не означает, что любая система тайно преследует противоположную своим словам цель. Скорее, заявленные ценности и практические стимулы нередко не полностью совпадают.

Что заявлено

Что вознаграждается

Какое поведение может возникать

«Главное — решить проблему клиента»

Количество закрытых обращений и среднее время обработки

Сотрудники быстрее закрывают простые обращения; сложные могут переходить между подразделениями. Это гипотеза, которую нужно проверять по данным, а не готовый приговор

«Мы оцениваем понимание материала»

Результаты тестов и скорость выполнения заданий

Учащиеся уделяют больше времени формату проверочной работы и типовым задачам, особенно если другие виды понимания почти не учитываются

«Нам важны содержательные рекомендации»

Время просмотра, число реакций и переходов

Авторы могут чаще выбирать материалы, способные удержать внимание; это не доказывает, что каждый автор намеренно жертвует содержанием ради кликов

В каждой строке важны слова «может возникать». Таблица не устанавливает причину и не описывает всех участников. Она подсказывает, где искать закономерность. Возможно, сотрудники не стремятся уклоняться от трудных обращений; возможно, именно так устроены их показатели, нагрузка и доступное время. Возможно, тест действительно хорошо измеряет нужные знания. Но если поведение участников устойчиво расходится с декларацией, архитектуру стимулов стоит проверить.

Важно и не делать слишком сильный вывод о влиянии игр или организаций на человека. Повторяющиеся награды способны менять выбор внутри конкретной среды: люди чаще пробуют то, что доступно, заметно и окупается. Но из этого не следует, что система напрямую формирует личность или отменяет свободу выбора. На поведение влияют опыт, убеждения, навыки, альтернативы и контекст. Анализ правил помогает увидеть давление среды, а не объявить человека её механическим продуктом.

Короткий аудит собственной среды

Попробуйте провести разбор на примере знакомой ситуации: учебного курса, приложения, отдела, семейного расписания или процесса записи к специалисту. Не начинайте с оценки людей. Начните с правил.

1. Назовите заявленную цель. Какой результат считается важным на словах?

2. Найдите фактическую награду. За что получают баллы, премию, доступ, похвалу, внимание или право закончить задачу?

3. Определите дефицит. Чего здесь не хватает: времени, денег, информации, полномочий, сил, места или доверия?

4. Запишите границы действия. Что запрещено формально, а что возможно только за слишком высокую цену? Какие варианты вообще не предусмотрены?

5. Найдите скрытую цену. Какое последствие возникает не сразу: усталость, потеря качества, перенос работы на других, упущенная возможность?

6. Опишите наблюдаемое поведение. Что люди делают регулярно, даже если официально говорится о другом?

7. Поменяйте мысленно одно правило. Как изменится наиболее выгодная стратегия? Какие новые лазейки или издержки могут появиться?

Если есть данные, разделите их по условиям и группам. Например, одинаково ли часто сложные обращения закрываются с первого раза у сотрудников с разной нагрузкой? Если данных нет, так и отметьте: это предположение, а не установленный факт. Такая аккуратность делает анализ сильнее, а не слабее.

Видеть не только слова

Скрытый кодекс системы записан не мелким шрифтом. Он проявляется в повторяющихся выборах: за что дают ресурс, что отнимает время, какой путь открыт, а какой закрыт, какую ошибку замечают и какую оставляют без внимания. Сюжет может говорить о спасении мира, но карта правил покажет, как устроен путь к победе.

Многомерные модели помогают задать вопросы, на которые не отвечает один показатель: какие факторы совместно производят результат, кому доступна выгодная стратегия, где возникает компромисс и что изменится при новом правиле. Их сила не в том, чтобы разоблачить любой замысел, а в том, чтобы заменить впечатление проверяемой картиной.

Поэтому при разборе любой среды полезно смотреть не только на то, что она обещает. Спросите, что в ней становится доступным, выгодным и повторяемым. Там, среди наград и ограничений, и проступает её настоящая архитектура.

Данные как пространство: признаки, измерения и способы представления реальности

До взрыва оставалось три минуты. На экране вспыхнули два предложения: немедленно установить боевой модуль и получить прибавку к урону — или отказаться и идти дальше с тем, что есть. Илья выбрал модуль. Впереди был босс, запас аптечек таял, а союзница Нера только что сказала: «Без усиления мы там не продержимся».

Решение казалось не просто разумным — единственно разумным.

— Я взял бонус, потому что иначе нас бы размазали, — сказал Илья, когда через час обсуждал прохождение с Леей, наблюдателем теста.

— Ты выбрал бонус, — уточнила Лея. — А что именно ты знал в тот момент?

Илья помолчал. Он помнил цифру: плюс тридцать процентов к урону. Помнил таймер. Помнил предупреждение о сложном бое. А вот что произойдёт с Нерой после установки модуля, не помнил. Тогда это и не выглядело частью решения.

Позже выяснилось: модуль работал на аккумуляторе, который Нера берегла для собственного снаряжения. Она отдала его, потому что доверяла Илье. Забрав аккумулятор, он не просто усилил персонажа — он закрыл ветку развития, где Нера могла стать полноценной напарницей. В следующей сцене она ушла. Вернуть её было нельзя.

Так один и тот же выбор выглядел сначала как обмен ресурса на силу, а потом — как обмен будущего на силу. Событие не изменилось. Изменилось представление о нём.

Именно здесь начинается разговор о данных. Любая модель — игровая, медицинская, финансовая, научная — не показывает реальность целиком. Она раскладывает её на признаки и размещает в пространстве, где одни различия становятся заметными, а другие исчезают.

Реальность, разложенная на координаты

Представим, что перед нами состояние игры в момент выбора. Его можно описать набором величин:

Признак

Значение для Ильи

Что он показывает

Время до боя

3 минуты

Насколько срочным кажется решение

Урон персонажа

42 единицы

Текущую боевую мощь

Бонус модуля

+30%

Немедленную выгоду

Аптечки

1

Запас на ошибку

Доверие Неры

Высокое

Состояние отношений с союзницей

Ветка развития Неры

Открыта

Возможность будущего прохождения

Каждый признак — координата. Вместе они задают точку в многомерном пространстве состояний. В одной координате Илья силён, в другой уязвим, в третьей торопится, в четвёртой ещё не знает, что находится на пороге необратимого изменения.

Число координат в таком описании может быть любым. Но важно не количество столбцов, а то, какие вопросы они позволяют задать. Если в данных есть только урон, время и аптечки, мы можем сравнить боевую готовность персонажей. Но мы не увидим цену решения для отношений и будущих маршрутов. Если добавить доверие и доступность сюжетных веток, картина изменится: бонус перестанет быть очевидно выгодным и превратится в один из нескольких компромиссов.

Признак — это не сама реальность, а выбранный способ о ней спросить. «Сколько у персонажа здоровья?» — один вопрос. «Насколько он способен выстоять в бою?» — другой: на него влияют здоровье, броня, реакция игрока, аптечки и устройство противника. В первом случае мы записываем непосредственно наблюдаемую величину. Во втором пытаемся измерить более сложное свойство через набор показателей.

Отсюда простое, но важное правило: прежде чем анализировать данные, нужно выяснить, что именно означает каждый столбец — и чего он не означает.

Почему интерфейс тоже является моделью

В момент выбора Илья видел крупную цифру «+30% к урону». Таймер горел красным. Информация о доверии Неры находилась в другом меню, а возможное влияние на её сюжетную линию не показывалось вовсе. Игра не соврала. Она лишь распределила внимание: выгоду сделала яркой, цену — почти невидимой.

— У меня не было выбора, — сказал Илья.

— Были другие варианты, — ответила Лея. — Ты мог отказаться от модуля, потратить аптечку на подготовку или обойти бой длинным маршрутом.

— Который занимал ещё двадцать минут и оставлял нас без лечения.

— Да. Они были дороже. Но не отсутствовали.

Илья был прав в одном: альтернативы выглядели тяжёлыми. Он был неправ в другом: дорогой выбор — всё ещё выбор. Разница существенна, потому что фраза «варианта не было» стирает стоимость альтернативы, а вместе с ней — устройство решения.

Это относится и к анализу данных. Модель, обученная на показателях скорости и результата, будет замечать быстрые и результативные действия. Если в ней нет признака «цена для команды», она не сможет внезапно открыть эту цену в итоговом отчёте. Отсутствующий столбец не всегда означает, что свойство не существовало. Иногда оно просто не было измерено.

Более того, сам способ сбора данных задаёт границы видимого. Если мы опрашиваем только тех игроков, которые завершили игру, то почти ничего не узнаем о тех, кто бросил её после тяжёлого решения. Если записываем только итоговый счёт, то теряем маршрут, которым этот счёт был достигнут. Данные — это не реальность в миниатюре, а след, оставленный определённым способом наблюдения.

Единицы измерения, кодировки и ловушки здравого смысла

Не все признаки устроены одинаково. Урон персонажа можно записать числом, время — в секундах, тип оружия — названием, уровень угрозы — словами «низкий», «средний», «высокий». При подготовке данных эти различия нельзя просто стереть.

Если закодировать оружие числами: меч — 1, винтовка — 2, арбалет — 3, модель может решить, будто арбалет «в полтора раза больше» меча или занимает промежуточное место между винтовкой и мечом. Но это выдуманный порядок: типы оружия не являются ступенями одной шкалы. Категории иногда кодируют отдельными признаками — например, «меч: да/нет», «винтовка: да/нет», — чтобы не навязывать им ложную числовую структуру.

Уровень угрозы уже имеет порядок, но промежутки между уровнями не обязательно равны. Переход от низкой угрозы к средней может означать одно, а от средней к высокой — совсем другое. Если считать эти категории обычными числами, это допущение должно быть осознанным, а не спрятанным в коде.

Не менее важен масштаб. В таблице Ильи время можно измерять минутами, урон — единицами, доверие — баллами от нуля до ста. Если при обучении модели один признак измеряется тысячами, а другой — долями единицы, большой числовой диапазон способен чрезмерно повлиять на некоторые алгоритмы. Масштабирование может помочь, но оно не исправляет неверную семантику. Привести доверие и урон к сопоставимому диапазону — не значит доказать, что одно и другое одинаково важны.

А теперь представим, что доверие Неры в журнале не записано. Пустое поле нельзя автоматически заменить нулём. «Неизвестно» и «доверие отсутствует» — разные состояния. Пропуск может означать сбой наблюдения, недоступный экран, неактивную сюжетную линию или то, что событие ещё не произошло. Иногда сам факт пропуска сообщает о процессе сбора данных: например, оценки чаще отсутствуют именно для тех пациентов, которые не смогли прийти на повторный приём.

Любая кодировка вносит предположения. Хорошая практика — не прятать их за техническими настройками, а записывать отдельно: что измеряли, в каких единицах, как обрабатывали категории и пропуски, какие преобразования применяли.

Видимое свойство и скрытая способность

Илья видел один понятный показатель силы — урон. Но способность пройти босса этим показателем не исчерпывалась. Важны были скорость реакции, выбранная тактика, знание атак противника, наличие аптечек и поведение союзницы. Некоторые из этих величин можно измерить напрямую. Другие приходится оценивать по косвенным признакам.

В данных такие косвенные показатели называют прокси. Например, количество сообщений в рабочем чате могут использовать как приблизительную меру вовлечённости сотрудников. Но частые сообщения не равны высокой вовлечённости: человек может писать много, потому что работа плохо организована. Высокий балл за тест может быть прокси знаний, однако он также зависит от формата экзамена, опыта сдачи и условий подготовки.

Доверие Неры — свойство, которое игра напрямую измеряла скрыто, по прошлым действиям Ильи: выполнил ли обещание, поделился ли ресурсом, прислушался ли к совету. Игрок мог видеть отдельные поступки, но не обязательно знал, как именно игра сводит их в показатель. Это различие между наблюдаемыми событиями и скрытой переменной. События можно перечислить; доверие приходится выводить из них.

Когда аналитик строит такую переменную, он должен спросить: какие именно наблюдения считаются свидетельством? Не перепутали ли мы причину со следствием? Не подменили ли сложное качество удобным показателем? Один высокий балл не превращает абстрактное свойство в объективную истину.

Есть и ловушка цели. Если разработчик решит, что «успешное прохождение» — это только победа над боссом, игра может поощрять действия, которые повышают урон любой ценой. Именно так Илья и прочитал ситуацию: победа рядом, сила измерима, значит, нужно

усилиться. Но успех мог включать ещё и сохранение союзников, доверия, ресурсов или возможности пройти историю иначе. В модели то, что объявлено целью, становится особенно заметным; то, что не включено в цель, рискует превратиться в невидимую потерю.

Состояние меняется — и вместе с ним меняется карта

Многомерное описание полезно не только для сравнения точек, но и для понимания переходов. Можно обозначить состояние игры в момент времени как $\backslash(xt\backslash)$, действие игрока — как $\backslash(at\backslash)$, а следующее состояние — как результат их взаимодействия: $\backslash(x\{t+1\}=F(xt,at)\backslash)$. *Это не обещание, что мы знаем все внутренние правила. Это напоминание: решение происходит не в пустоте. Оно меняет мир, в котором будут приниматься следующие решения.*

После установки модуля у Ильи вырос урон. Одновременно снизилось доверие Неры, исчезло её снаряжение и закрылась длинная ветка развития. Позднее ему открылась новая информация: аккумулятор был не просто общим запасом, а обещанным ресурсом союзницы. В этот момент изменилось не прошлое, а карта будущего.

— Значит, я ошибся? — спросил Илья.

— Ты ошибся в чём именно? — сказала Лея. — В оценке цены? В оценке шансов на бой? Или в том, что называл этот выбор единственным?

Вопрос неприятен, потому что требует различать разные ошибки. Решение может оказаться неудачным по результату и всё же быть разумным с учётом доступной информации. И наоборот: удачный исход не доказывает, что ход был хорошо обоснован. Нельзя оценивать решение только по тому, как сложилась история. Нужно восстанавливать данные, которые были доступны до действия.

Илья победил босса с модулем. В следующей сцене, увидев, что Нера уходит, он сначала хотел продолжать ту же стратегию: если сила помогла один раз, значит, надо усиливаться дальше. Но теперь цель для него изменилась. Победа над следующим противником уже не казалась достаточной, если ради неё предстояло потерять оставшихся союзников. Он выбрал защитный маршрут, отдал часть запасов на лечение команды и перестал вкладывать ресурсы только в собственный урон.

Это не отменило прежнего решения. Оно изменило его следующее.

Практика: разберите выбор до того, как перепишете его задним числом

Когда решение уже принято, память охотно подгоняет прошлое под нынешнее знание. «Я всё просчитал» защищает образ компетентного человека. «У меня не было выбора» снижает необходимость исследовать альтернативы. Чтобы не превращать разбор в суд над собой, полезно отделить сведения, доступные тогда, от сведений, появившихся позже.

Попробуйте сделать это на одном собственном решении — игровом, рабочем или бытовом.

1. Зафиксируйте момент. Запишите, что происходило непосредственно перед выбором. Не добавляйте информацию, полученную позже.

2. Назовите цель. Чего вы пытались добиться тогда, а не чего вам хотелось бы добиться теперь?

3. Перечислите доступные признаки. Что вы знали, что видели, что оценивали лишь приблизительно? Какие важные параметры вообще не попали в поле зрения?

4. Запишите альтернативы и их цену. Не спрашивайте только «что я мог сделать?». Спросите также: «чем каждый другой путь казался хуже, рискованнее или дороже?»

5. Отделите итог от качества решения. Удача и неудача — сведения об исходе, но не готовый приговор о том, насколько разумным был выбор при тогдашних данных.

Эта практика помогает не оправдываться нехваткой информации и не обвинять себя за то, чего нельзя было знать. Она возвращает решению его координаты: цель, время, ресурсы, риск, доступные знания и ограничения.

Протокол необратимого решения

Для решений, которые закрывают возможности, полезно вести короткую запись. Не для того, чтобы предсказать всё, а чтобы после не переписывать условия, в которых был сделан выбор.

Поле

Что зафиксировать

Краткосрочная выгода

Что решение давало сразу и насколько ценным это казалось

Долгий след

Какие отношения, ресурсы, ограничения или обязательства могли измениться

Упущенная альтернатива

Какой путь был доступен, но казался слишком дорогим, медленным или рискованным

Новая информация

Что выяснилось после решения и действительно изменило понимание ситуации

Исход и цена

Что произошло и чем пришлось заплатить за результат

Изменившаяся цель

Чего теперь нужно добиваться с учётом новых обстоятельств

Новая точка выбора

Какое действие остаётся доступным сейчас

Для Ильи запись могла бы выглядеть так. Краткосрочная выгода: усиление перед сложным боем. Долгий след: расход общего аккумулятора и снижение доверия союзницы. Упущенная альтернатива: пройти обходным маршрутом, потратив время и аптечку. Новая информация: аккумулятор был обещан Нере, а модуль закроет её ветку развития. Исход и цена: босс побеждён, союзница ушла. Изменившаяся цель: не просто наращивать урон, а сохранить оставшуюся команду. Новая точка выбора: идти дальше в одиночку или вложить ресурсы в защиту других персонажей.

Такой разбор не отвечает на вопрос, был ли Илья «хорошим игроком». Он отвечает на более полезный вопрос: в каком пространстве он выбирал? Какие координаты заметил, какие считал важными, какие остались за краем интерфейса?

Данные всегда уже кем-то отобраны и представлены. Признаки превращают мир в описание; шкалы и кодировки придают этому описанию форму; пропуски и скрытые переменные напоминают, что за формой остаётся невидимая часть. Многомерная модель не отменяет реальность — она предлагает конкретный способ её увидеть. Поэтому зрелый анализ начинается не с уверенного вывода, а с внимательного вопроса: какие измерения сделали этот вывод возможным, и какие важные различия они оставили за пределами картины?

Свобода не означает, что человеку доступны все сведения или что любой путь ничего не стоит. Но даже при ограниченной информации важно видеть разницу между отсутствующей альтернативой и альтернативой, которую трудно оплатить. После необратимого решения прошлое не меняется. Меняется пространство следующих шагов — и умение замечать его начинается с точного описания того, что произошло.

Геометрия сходства: расстояния, направления и близость объектов

На последнем уровне игры Илья стоит перед запертой станцией. За дверью — система управления эвакуацией, на подходе — отряд противника. Чтобы добраться до пульта, нужно отвлечь охрану. Лев, единственный из команды, способен включить радиомаяк и увести врагов за собой. Вероятность успеха — восемьдесят восемь процентов.

Игрок уже давно считает Илью человеком, который не бросает своих. Он всегда выбирал реплики о верности, носил на куртке знак отряда и однажды ради спасения пленного свернул с маршрута. Но он также привык считать шансы. Прежде чем нажать кнопку, он ещё раз сверяет цифры. Потом отдаёт приказ.

Станция открывается. Эвакуация проходит успешно. Лев возвращается живым, но больше не смотрит на Илью так, как прежде.

— Я спас тебя, — говорит Илья.

— Ты спас ресурс, — отвечает Лев. — А я оказался его условием.

Игрок помнит своё намерение: он хотел спасти город. Но игра показывает след его решений — последовательность поступков, из которой постепенно сложился герой. Между этими двумя версиями и возникает геометрия сходства: способ измерить, насколько один объект похож на другой, и увидеть, в какую сторону он меняется.

Карта поступков

В предыдущей главе данные предстали не зеркалом реальности, а координатной сеткой: одни свойства на ней видны, другие остаются за краем. Геометрия начинается там, где мы задаём следующий вопрос: какие объекты в этой сетке близки друг к другу?

Объектом может быть человек, город, покупка, медицинский случай или игровой персонаж. Чтобы сравнивать их, каждый описывают набором признаков. Например, Илью можно представить вектором из пяти чисел:

(делится ресурсами, держит обещания, рискует собой, ищет новую информацию, использует союзников как средство).

Пусть каждое значение лежит между нулём и единицей. Ноль означает, что признак почти не проявлялся, единица — что проявлялся часто. После восьми игровых эпизодов у Илья получился такой профиль:

(0,45; 0,25; 0,35; 0,70; 0,85).

Это не полное описание героя. Это краткая запись того, что система заметила в его действиях. И в этой оговорке заключена важная честность: числа не сообщают, добрый ли Илья человек. Они сообщают, как часто он делился, выполнял обещания, шёл на риск, искал сведения и назначал союзников на опасные роли.

Повторяющиеся поступки можно прочесть по-разному. Если герой часто отдаёт дефицитные припасы нуждающимся и сам принимает опасность на себя, он окажется близок к образу спасителя. Если он снова и снова ищет неизвестные сигналы и отклоняется от безопасного маршрута ради загадки — к исследователю. Если добывает сведения, обещая одно, а делая другое, может напоминать манипулятора. Если выбирает действие по его ожидаемой пользе для всей системы — прагматика.

Это не четыре тайные сущности, запертые внутри одного персонажа. Это четыре возможных способа сгруппировать наблюдаемое поведение. Тот же набор решений может приблизить героя к разным образам — в зависимости от того, какие признаки мы выбрали, насколько их ценим и с кем сравниваем.

Что значит «рядом»

Самое привычное расстояние — евклидово. Представьте, что у каждого героя есть точка в пространстве признаков. Евклидово расстояние — длина кратчайшего отрезка между двумя такими точками:

$$\sqrt{d(x,y)=\sqrt{\sum_j (x_j-y_j)^2}}.$$

В двух измерениях его легко увидеть как диагональ прямоугольника. В десяти или ста измерениях вообразить его сложнее, но формула остаётся той же: она суммирует различия по признакам и учитывает большие расхождения особенно заметно.

Предположим, мы сравниваем Илью с условными образцами ролей. Чтобы не путать их с диагнозами, будем считать их всего лишь примерными профилями: защитник чаще делится и держит обещания; исследователь чаще ищет новую информацию; прагматик соотносит риск с задачей; манипулятор чаще использует обещания и союзников как инструменты. При одинаковом весе пяти признаков Илья может оказаться ближе всего к прагматику и манипулятору, а заметно дальше — от защитника.

Это не приговор герою. Это ответ на более узкий вопрос: «Какой из заданных образцов ближе к зафиксированному профилю?» Если нам не нравится ответ, нужно не объявлять формулу неправильной, а проверить, как именно мы составили карту: какие признаки внесли, как их измерили и все ли одинаково важны.

Евклидово расстояние подходит, когда признаки числовые, сопоставимы по масштабу и большие отклонения действительно должны сильнее влиять на итог. Но если один признак измеряется в днях, другой — в долларах, а третий — по шкале от нуля до пяти, расстояние может оказаться почти полностью продиктовано самой крупной шкалой. Сравнение будет похоже на измерение маршрута линейкой, у которой одна сторона растянута.

Другой вариант — расстояние Манхэттена: оно складывает абсолютные различия по каждой координате. Его можно представить как путь по улицам города, где нельзя срезать кварталы по диагонали:

$$\sqrt{d(x,y)=\sum_j |x_j-y_j|}.$$

Оно полезно, когда важно учитывать отдельные изменения по признакам, а не позволять одному крупному расхождению доминировать за счёт возведения в квадрат. Однако само по себе оно не спасает от несопоставимых шкал: если значения не приведены к понятной основе, проблема остаётся.

Косинусное сходство сравнивает не длину вектора, а угол между двумя векторами. Грубо говоря, оно отвечает на вопрос: «Похожи ли направления усилий?» Это особенно удобно, когда важен рисунок распределения, а не общий объём. Например, два текста могут быть похожи по тематике, даже если один намного длиннее другого. Или две команды могут тратить ресурсы в похожих пропорциях, хотя одна располагает гораздо большим бюджетом.

Но и здесь есть важная оговорка: косинус не учитывает масштаб так, как это делает обычное расстояние, а для нулевого вектора угол не определён. Иными словами, если у объекта вообще нет ни одного зафиксированного действия, нельзя честно вывести его «направление».

Расстояние Махаланобиса учитывает ещё одну тонкость: некоторые признаки связаны между собой. Например, герой, который часто идёт на личный риск, может также часто спасать союзников. Если считать эти свойства двумя полностью независимыми основаниями для сходства, мы рискуем дважды учесть одну и ту же закономерность. Расстояние Махаланобиса корректирует сравнение с учётом разброса и связей между признаками. Оно полезно, но требует достаточно качественных данных: в небольшой выборке оценка таких связей может оказаться ненадёжной.

Ни одна метрика не является универсальной линейкой сходства. Разные вопросы требуют разных расстояний. «Насколько похожи действия по каждому признаку?» — не то же самое, что «Насколько похож общий рисунок?» Сначала нужно сформулировать вопрос, затем выбирать меру.

Не только близость, но и направление

Расстояние говорит, насколько объекты различаются. Но оно не всегда показывает, **в чём именно** состоит разница и куда объект движется.

Для этого полезно смотреть на направление изменения: на разницу между вектором до и после решения. Допустим, раньше герой редко делился запасами, а в последних эпизодах стал отдавать их чаще. Его координата «делится ресурсами» сдвинулась. Если одновременно он стал чаще выполнять обещания, профиль движется в сторону выбранного нами образа защитника. Если он лишь совершил один эффектный поступок, а остальные решения не изменились, общий сдвиг может быть небольшим.

Косинусное сравнение здесь помогает оценить, совпадает ли направление фактических перемен с желаемым. Герой может быть далеко от защитника, но постепенно двигаться к нему. А может на один эпизод оказаться близко к нему — и затем уверенно пойти в обратную сторону.

Это различие знакомо и вне игр. Человек, который начал работать в новой команде, может спросить не только «На кого я похож?», но и «В чью сторону меня ведут повторяющиеся привычки?» Отдельный поступок — точка на маршруте. Повторяющиеся решения дают возможность различить курс.

Вот как выглядели бы пять повторяющихся ходов Ильи, если рассматривать не только намерение, но и последствия.

Повторяющийся поступок Предполагаемое намерение Непосредственное последствие Отложенное последствие Какой образ складывается

Откладывает лекарства, пока не станет ясно, кто нужен группе для выполнения задачи
Сохранить запас для критического момента
Запас остаётся в команде
Люди, которым он отказал, начинают сомневаться в его бескорыстии
Бережливый или расчётливый

Обещает пленному безопасность в обмен на сведения, затем передаёт его командованию
Не допустить допроса силой и получить маршрут
Команда узнаёт путь к станции
Следующие собеседники меньше доверяют обещаниям
Дипломат или манипулятор

Сворачивает с безопасной дороги к неизвестному сигналу
Понять, с чем предстоит столкнуться
Теряется время
Позже команда обнаруживает засаду и избегает её
Исследователь

Сам входит в опасную зону, если там лежит необходимый ключ
Не перекладывать риск на других и завершить задачу
Ключ добыт, цель продвинулась
Союзники видят смелость, но замечают и избирательность его риска
Защитник или тактик

Назначает Льва приманкой, потому что тот лучше всех управляется с маяком
Успеть открыть станцию и эвакуировать людей
Система запускается
Лев жив, но доверие между ними повреждено
Прагматик, который обращается с людьми как с ресурсом

Таблица не объясняет поступки вместо самого Ильи. Она позволяет удержать рядом четыре вещи, которые обычно сливаются в одну: намерение, немедленный результат, позднее последствие и образ, прочитанный другими. Так становится видно, почему «я хотел помочь» и «мне доверяют» не всегда совпадают.

Тихий перелом

В финале игрок снова видит числа: восемьдесят восемь процентов успеха, если Лев отвлекает охрану, и сорок один — если команда идёт обходным путём. В первом случае можно запустить эвакуацию и вывести из опасной зоны сотни людей. Во втором Лев останется рядом с Ильёй, но шансы успеть к пультам резко упадут.

Оба выбора имеют цену. Отдать приказ — значит использовать союзника в роли, для которой его способности особенно полезны, и рискнуть доверием. Отказаться — значит сохранить верность ближайшему человеку, но принять больший риск для остальных. Игра не выдаёт универсального морального ответа. Она показывает цену, которую игрок готов принять, и того, кому эту цену предстоит заплатить.

Игрок нажимает кнопку. Поступок выглядит разумным в логике задачи. Возможно, он и был правильным для спасения людей. Но после победы Ильи не может убедить Льва одной лишь декларацией о верности. Реплика союзника не отменяет намерения Ильи — она добавляет к портрету ещё один ракурс.

Вот внутренний перелом: если внимательно смотреть на повторяющиеся решения, становится труднее прятаться за тем, каким героем хочется быть. Это верно не только для персонажа. Мы тоже склонны называть себя терпеливыми, надёжными, щедрыми — и замечать прежде всего те случаи, которые подтверждают желаемый образ. Но окружающие нередко считают другие координаты: как мы поступаем под давлением, что откладываем, чьи интересы принимаем в расчёт и какие последствия считаем допустимыми.

Практика: портрет по действиям

Попробуйте составить такой портрет для себя — не по убеждениям, а по повторяющимся поступкам за последние несколько недель или месяцев. Выберите пять действий, которые вы действительно совершали неоднократно. Не берите только крупные события: привычка перебивать, держать обещание, откладывать трудный разговор или соглашаться на лишнюю работу тоже оставляет след.

Для каждого поступка заполните пять колонок: **что я делаю; чего намерен добиться; что происходит сразу; что меняется позже; какой образ может складываться у других.** Полезно добавить ещё одну пометку: как часто это действие происходило и в каком контексте. Она уберёжет от слишком смелых выводов по единичному случаю.

Затем пройдите маршрут из шести шагов.

1. Назовите признаки. Выберите несколько наблюдаемых свойств, а не оценок вроде «хороший человек». Например: выполняет обещания, делится временем, ищет обратную связь.

2. Задайте шкалы. Решите, как фиксировать каждый признак: числом, частотой или отметкой «да/нет».

3. Проверьте масштаб. Убедитесь, что одна координата не начинает влиять на результат только потому, что у неё более крупные значения.

4. Выберите сравнение. Спросите, что важно: близость по каждому отдельному признаку, сходство общего рисунка или движение в желаемую сторону.

5. Ищите повторение. Сравнивайте не только отдельные поступки, но и устойчивые привычки, а также ситуации, в которых они проявляются.

6. Сверьте карту с реальностью. Спросите нескольких людей, как они видят последствия ваших действий. Их ответы не будут абсолютной истиной, но помогут заметить признаки, которых вы сами не внесли в модель.

Завершите упражнение двумя фразами: «Я намеревался...» и «По моим повторяющимся поступкам другие могут заключить...». А затем попробуйте заполнить формулу без самооправдания и самоуничижения: **«Я часто делаю..., поэтому окружающие могут видеть во мне...»**

Не стремитесь немедленно исправить получившийся портрет. Сначала важно увидеть его целиком. Геометрия сходства не выносит приговоров; она показывает, какие признаки сближают объекты, какие отдаляют и в каком направлении ведут изменения. Её сила — не в том, чтобы назначить человеку единственную роль, а в том, чтобы сделать видимым рисунок, который отдельными эпизодами легко не заметить.

И всё же карта сходства надёжна лишь настолько, насколько надёжны координаты. Если один поступок не записан, другой посчитан дважды, а шкалы устроены несогласованно, расстояния будут точными только на бумаге. Прежде чем сравнивать объекты, нужно разобраться, что делать с пропусками, выбросами и несопоставимыми признаками.

Подготовка данных: пропуски, выбросы, шкалы и кодирование признаков

В понедельник в 14:12 на экране появляется сообщение: «Нужно обсудить вчерашнее решение». Плечи напрягаются, челюсть сжимается. Вы думаете: «Меня сейчас будут обвинять». Ответ откладывается до вечера. Через несколько часов вы рассказываете себе, что просто хотели подобрать точные слова.

Что здесь произошло? Раздражение? Осторожность? Усталость? А если вы вечером не помните, насколько сильно сжали челюсть, это значит, что напряжения не было? Если ответили через четыре часа, можно ли уверенно считать, что именно раздражение вызвало задержку?

Такие вопросы похожи на головоломку, но на деле это вопросы подготовки данных. Любая модель — от сложной статистической системы до нашего личного дневника — видит не жизнь целиком, а набор записей. И прежде чем искать закономерности, нужно разобраться, что означают пустые ячейки, необычные значения, разные шкалы и слова, превращённые в числа.

Эта глава — небольшая лаборатория наблюдения. На семь дней мы отложим оценку себя и попробуем точнее заметить, что происходит. Не чтобы вынести себе вердикт, а чтобы создать аккуратные данные о повторяющихся реакциях. Каждый день займёт не больше десяти минут. Если подходящий эпизод не случится, так и запишите: «не было». Это не провал наблюдения. Это тоже факт.

Сначала — не улучшать, а фиксировать

Возьмите бумагу или заметку в телефоне. Для каждого эпизода используйте пять обязательных полей:

- 1. Ситуация:** что наблюдалось снаружи — где, когда, что произошло.
- 2. Первая реакция:** что вы почувствовали в теле и какая мысль возникла первой.
- 3. Совершённое действие:** что вы действительно сделали или не сделали.
- 4. Альтернативное объяснение:** какая ещё причина могла повлиять на реакцию.
- 5. Маленькая проверка в жизни:** что можно осторожно проверить, не подвергая себя риску.

Для ясности разделяйте наблюдение и толкование. «Сообщение пришло в 14:12» — факт. «Меня хотели пристыдить» — объяснение, возможно верное, но пока не подтверждённое. «Я ответил в 18:05» — факт. «Я избегал разговора» — гипотеза о причине.

Эта граница важна не только для дневника. Аналитик, который заполняет неизвестный ответ нулём, принимает отсутствие сведений за отсутствие события. Человек, который записывает «не было раздражения», потому что ничего не вспомнил, может сделать ту же ошибку. Качество вывода начинается с качества различий.

Можно представить работу с записью как небольшой конвейер.

Роль

Задача

Пример

Пользователь

Записывает событие без самооценки

«Ответил в 18:05»

Администратор данных

Определяет поля и допустимые значения

Дата, ситуация, реакция, действие

Аналитик

Проверяет, что запись позволяет заключить

Задержка ответа не доказывает избегание

Эти роли могут принадлежать одному человеку. Важно лишь не смешивать их в одну торопливую фразу.

День 1. Раздражение: заметить момент, а не характер

Вопрос дня: в какой ситуации возникло раздражение, что вы почувствовали и что сделали?

Выберите один эпизод. В поле «ситуация» запишите только то, что могла бы зафиксировать камера: «На совещании коллега перебил меня дважды». В поле «первая реакция» разделите телесное ощущение и мысль: «жар в лице; подумал, что мою идею не уважают». В «совершённом действии» — наблюдаемое: «перестал говорить», «попросил не перебивать» или «ничего не сказал».

В альтернативном объяснении не нужно оправдывать ни себя, ни другого человека. Достаточно допустить ещё одну возможность: «Коллега мог торопиться закончить встречу». Маленькая проверка может быть совсем небольшой: «В следующий раз уточнить, правильно ли я понял, что меня перебили намеренно».

Подготовка данных часто начинается с определения самого измеряемого признака. «Раздражительность» — не простая величина вроде длины стола. Она может означать телесное напряжение, резкость ответа, частоту эпизодов или оценку по шкале. Если всё объединить в один ярлык, модель получит удобное слово, но не обязательно понятное измерение. Поэтому не спешите записывать «я раздражительный». Записывайте то, что произошло.

День 2. Отложенное решение: увидеть цену паузы

Вопрос дня: какое решение вы отложили и какую наблюдаемую цену имела пауза?

Факт может быть таким: «Не подтвердил время встречи до вечера; свободный слот занял другой человек». Объяснение — «боялся выбрать неудобное время». Цена паузы не обязательно драматична: пропущенный срок, лишнее напоминание, дополнительное усилие, а иногда — никакой заметной цены.

В данных особенно опасно подменять пропуск нулём. Если поле «цена задержки» пустое, это не то же самое, что «цены не было». Возможно, вы не заметили последствий; возможно, их нельзя оценить; возможно, вопрос неприменим. Такие причины важно различать. В аналитике пропуски часто описывают как случайные, связанные с другими известными признаками или зависящие от самого неизвестного значения. Например, человек может чаще не отмечать сильные эмоции именно тогда, когда они сильны. Тогда пустая запись сама несёт информацию.

В маленьком дневнике не нужно заполнять неизвестное догадкой. Используйте ясные пометки: «не помню», «не применимо», «не записал». В крупном исследовании иногда применяют статистическое заполнение — например, медианой или модельной оценкой, — но это не возвращает утраченное наблюдение. Заполнение создаёт предположение, а предположение должно быть обозначено.

День 3. Обещание: отделить намерение от события

Вопрос дня: какое обещание вы дали и было ли оно выполнено в оговорённый срок?

Запишите обещание конкретно: «сегодня отправлю документ», а не «буду внимательнее». Затем укажите срок и результат: отправлен, не отправлен, перенесён по взаимной договорённости. Не превращайте намерение в действие: «я собирался написать» не равно «я написал».

Именно здесь часто появляется привлекательная, но слабая переменная: самоописание. «Я ответственный» или «я всё откладываю» — крупные ярлыки, в которые сливаются разные ситуации. Для анализа полезнее определить, что считается выполнением, на каком промежутке и при каких условиях. Это похоже на настройку прибора: прежде чем сравнивать показания, нужно понять, что он измеряет.

Если всё же используете числовую шкалу, задайте ей устойчивые якоря. Например, оценка напряжения от 0 до 4: 0 — не замечал напряжения, 1 — слабое, 2 — заметное, 3 — сильное, 4 — мешало действовать. Не меняйте смысл цифр от дня ко дню. И помните: такая шкала порядковая. Мы знаем, что 4 означает больше напряжения, чем 2, но не можем автоматически утверждать, что разница между 2 и 4 ровно вдвое больше разницы между 1 и 2.

День 4. Реальное действие: увидеть, что последовало за обещанием

Вопрос дня: что вы фактически сделали после данного себе или другому обещания?

Сопоставьте запись с третьего дня. Не выносите оценку «сдержал слово» или «подвёл». Отметьте последовательность: обещание, срок, действие, результат. Например: «Планировал отправить отчёт к 16:00; открыл файл в 15:40; отправил в 16:25». Тут уже виден более точный рисунок, чем в общем впечатлении «я всегда тяну до последнего».

В многомерных данных важна не только отдельная цифра, но и сочетание признаков. Время до действия может быть связано с типом задачи, усталостью, ясностью следующего шага или ожиданием ответа. Один задержанный отчёт не выявит закономерности. Он даст материал для дальнейшего наблюдения.

Но даже хорошо записанные признаки бывают в разных масштабах: минуты, количество напоминаний, оценка напряжения. Для моделей, основанных на расстояниях или градиентных вычислениях, масштаб может сильно влиять на результат. Если время измерено в минутах, а число напоминаний — единицами, первая переменная способна численно «заглушить» вторую просто из-за своей шкалы.

Распространённое решение — стандартизация: из значения вычитают среднее и делят на стандартное отклонение. Тогда признаки выражаются в единицах типичного разброса. Для сильно перекошенных данных бывает полезнее масштабирование через медиану и межквартильный размах — расстояние между центральными половинами наблюдений. Выбор зависит от данных и метода. Масштабирование помогает сопоставлять величины, но не делает плохой признак хорошим и не превращает условную оценку в объективный факт.

День 5. Оправдание: услышать объяснение, не принять его за доказательство

Вопрос дня: какое объяснение вы дали своему действию и какие факты его поддерживают?

Например: «Не ответил на письмо, потому что день был загружен». Факты: в календаре было пять встреч; письмо открыто в 11:20; ответа нет. Это совместимо с объяснением, но не доказывает его. Возможна и другая версия: ответ казался трудным, а встречи стали удобной причиной отложить его.

Не требуется разоблачать себя. Требуется держать рядом несколько возможных объяснений и не путать их с записью. Эта дисциплина особенно нужна при обработке выбросов — необычных значений. Если однажды вы ответили через два дня, это ещё не ошибка. Возможно, телефон был выключен, задача была сложной или вы уехали. А если в журнале вдруг появилось «ответил через 20 000 минут», причиной может быть ошибка единиц измерения.

Выброс не следует удалять только потому, что он портит красивую картину. Сначала выясните, что произошло. Ошибка ввода — основание исправить значение по первоисточнику. Редкий, но реальный эпизод — повод сохранить его и отметить контекст. Если экстремальные значения сильно влияют на расчёты, применяют устойчивые методы, преобразования или ограничение крайних значений, но такие решения нужно обосновать и задокументировать. Иначе чистота данных окажется просто аккуратным способом спрятать неудобный факт.

День 6. Повтор: найти рисунок, не придумать его

Вопрос дня: повторилась ли за неделю ситуация, первая реакция или действие, похожие на прежние?

Сравнивайте записи не по ощущению «это всегда так», а по конкретным признакам. Например: в четырёх случаях неопределённый запрос сопровождался паузой перед ответом. Это уже наблюдаемый рисунок — но пока не объяснение его причины. Возможно, влияет неопределённость; возможно, все четыре запроса пришли в конце рабочего дня.

Чтобы анализировать категории, их тоже нужно кодировать разумно. Если вы отмечаете тип ситуации словами «срок», «критика», «неясная просьба», «изменение плана», не присваивайте им числа 1, 2, 3, 4 так, будто четвёртый тип больше первого. Для категорий без естественного порядка часто используют отдельный индикатор для каждого значения: например, «был срок — да/нет», «была критика — да/нет». Это называется one-hot-кодированием. Порядковый код подходит только там, где порядок действительно имеет смысл: например, «слабое — среднее — сильное»; и даже тогда расстояния между уровнями могут быть неравны.

Редкие категории требуют осторожности. Можно объединить похожие значения, если это оправдано смыслом, но не ради того, чтобы таблица выглядела проще. Слова «другое» и «не знаю» также не должны превращаться в одну категорию: за ними могут стоять разные состояния данных.

Для семидневного дневника всё это — способ мыслить аккуратнее, а не приглашение запускать сложную модель на семи наблюдениях. Такой объём не позволяет надёжно оценить устойчивую закономерность. Ценность недели — в проверке качества вопросов и в первых гипотезах, а не в статистическом приговоре.

День 7. Перенос: выбрать один безопасный эксперимент

Вопрос дня: какой один повторяющийся паттерн стоит проверить в ближайшую неделю?

Перечитайте записи и выберите не самый драматичный эпизод, а самый ясный и переносимый. Например: «Когда запрос неясен, я часто откладываю ответ». Сформулируйте эксперимент так, чтобы он был безопасным, конкретным и не требовал менять всю жизнь: «В следующий раз, когда получу неясный запрос, в течение рабочего дня задам один уточняющий вопрос».

Конец ознакомительного фрагмента.

Текст предоставлен ООО «Литрес».

Прочитайте эту книгу целиком, [купив полную легальную версию](#) на Литрес.

Безопасно оплатить книгу можно банковской картой Visa, MasterCard, Maestro, со счета мобильного телефона, с платежного терминала, в салоне МТС или Связной, через PayPal, WebMoney, Яндекс.Деньги, QIWI Кошелек, бонусными картами или другим удобным Вам способом.