МОДУЛЬ 3 · ТЕМА 3.1

Поиск и сравнение технических решений

Модель ищет варианты, выбор остаётся за инженером



3.1

День 3 · Виталий Пиков · архитектор информационной безопасности · pikov.expert

Сквозная задача дня – К-1

14 суток

при −25 °C

+2 кг

максимум

- Как не допустить глубокого разряда батарей ЭС-24 на стоянке до **14 суток при −25 °C**
- Выросла из замечаний 3-1 и 3-2
- Тахограф и охрана должны работать всю стоянку
- Батареи и генератор не меняем, масса **+2 кг** максимум

Документ 3.1 дополнения к кейсу

От задачи к выбранному решению

1

Задача, ограничения, критерии – инженер

2

Варианты, аналоги, доводы против – модель

3

Отсев по физике и ограничениям – инженер

4

Матрица: веса и оценки – инженер

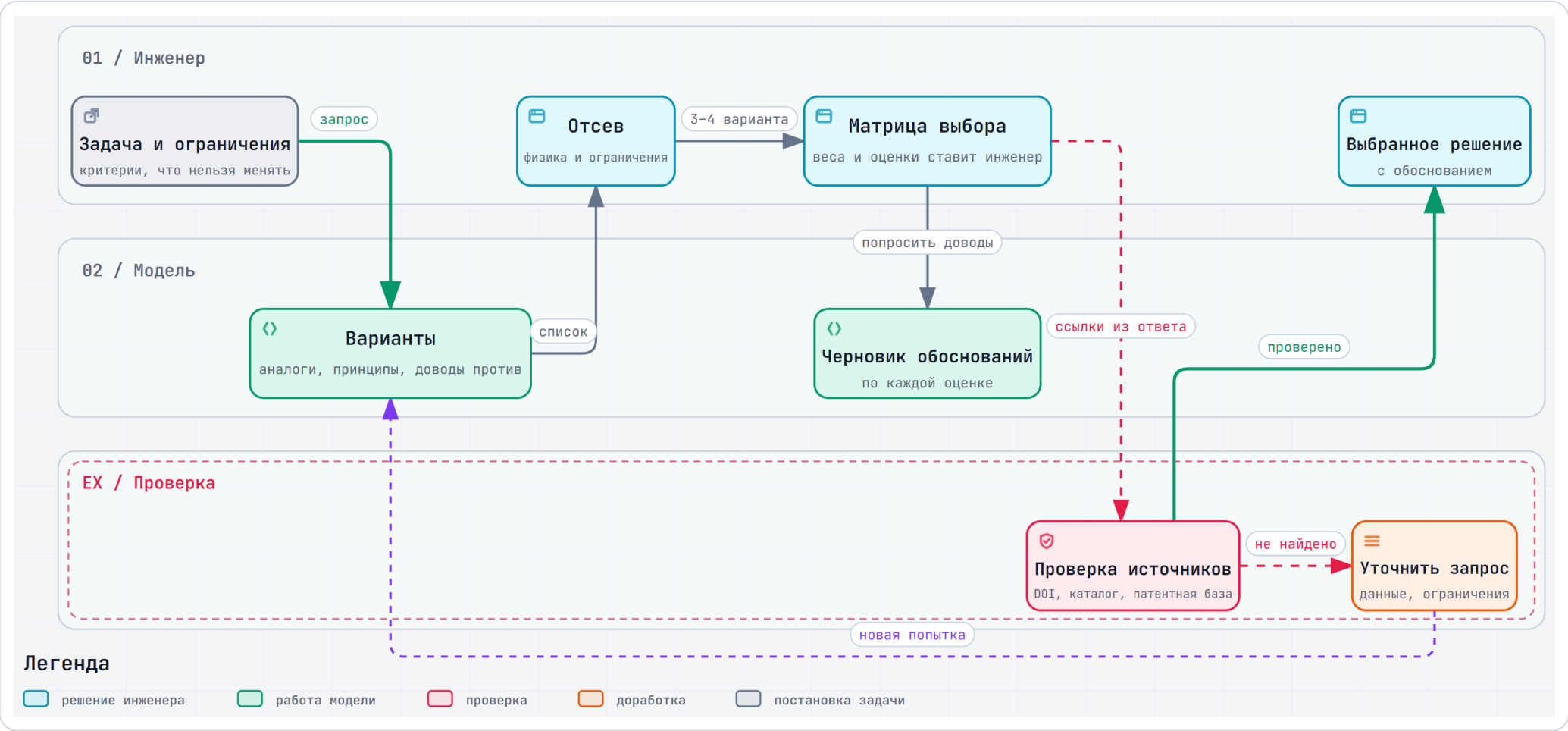
5

Проверка каждой ссылки

6

Выбор и обоснование – инженер

От задачи к выбранному решению



Интерактивная версия – 08-Схемы/d3-put-k-resheniyu.html

Модель в поиске идей

Быстро вспоминает аналоги и принципы действия

В эксперименте на 12 задачах варианты модели **осуществимее**, варианты людей **новее**

Широту дают роли по очереди и вторая модель-оппонент

Сбои: бросает идеи при трудностях, усложняет, застревает на одном решении

Добротно, но не ново

> Экспертные оценки показывают, что решения языковой модели в среднем более осуществимы и полезны, а решения, собранные у множества людей, отличаются большей новизной.

Ма и соавт., 2023, препринт arXiv, doi.org/10.48550/arXiv.2306.01779

Противоречия в задаче К-1

Обесточить всё на стоянке – но тахограф и охрана должны работать

Поставить батареи больше – но ёмкость не меняется, масса ограничена

Приём: сначала попросить модель **сформулировать противоречия**, потом искать варианты

Просить для каждого варианта **доводы против** и что измерить

Матрица выбора: три уязвимых места



- **Шкала:** что значит 5 и что значит 1
- **Веса:** сумма равна 1, ставит инженер
- **Оценки:** у каждой есть обоснование
- До матрицы – отсев: недопустимый вариант баллы не спасают

Критерий успеха = сценарий + результат (П. Гагарин)

Матрица на примере: устойчивость выбора

Критерий	Вес	Вес после сдвига	А	Б
Влияние на пуск	0,5	0,7	4	5
Стоимость	0,3	0,1	4	2
Масса	0,2	0,2	5	4
Итог при основных весах			**4,2**	3,9
Итог после сдвига			4,2	**4,5**

А: 0,5 × 4 + 0,3 × 4 + 0,2 × 5 = 4,2. Сдвиг 0,2 веса со стоимости на пуск меняет лидера: решает влияние на пуск, его проверяют расчётом или испытанием

Кто что делает

Шаг	Модель	Инженер
Постановка	Формулирует противоречия, вопросы	Ограничения и критерии
Варианты	Варианты и доводы против	Отсев
Матрица	Проверка на противоречия, черновики обоснований	Веса и оценки
Источники	Подсказывает, где искать	Открывает и сверяет
Решение	Черновик записки	Выбор и ответственность

Почему оценкам модели нельзя доверять напрямую

Смещение по позиции: выше оценивает вариант на определённом месте

Смещение по длине: предпочитает многословное описание

Проверка: переставить варианты, выровнять описания, спросить в новом чате

Изменился итог – оценки модели не годятся даже как черновик

Выдуманные источники

47,5 %

подтверждено не больше

36–61 %

ни подтвердить, ни опровергнуть

2 настоящие из 35

проверка 2023 года

- Эксперимент 2026 года, четыре модели без доступа к источникам: подтверждено не больше **47,5 %** ссылок
- Ещё **36–61 %** ссылок автоматическая проверка не смогла ни подтвердить, ни опровергнуть
- Отдельная проверка 2023 года: **2 настоящие из 35**
- Код: выдуманные библиотеки – не меньше **5,2 %** у коммерческих и **21,7 %** у открытых моделей
- Поиск в интернете эту угрозу не устраняет

Правильное оформление не подтверждает существование работы

Из 17 443 сгенерированных ссылок ни у одной модели доля существующих работ не превышает 0,475. […] а оформление ссылок при этом остаётся правильным.

Zhao, Tang, Qian, 2026, препринт arXiv, doi.org/10.48550/arXiv.2603.07287

17 443

сгенерированных ссылок

не превышает 0,475

ни у одной модели доля существующих работ

Как проверить ссылку

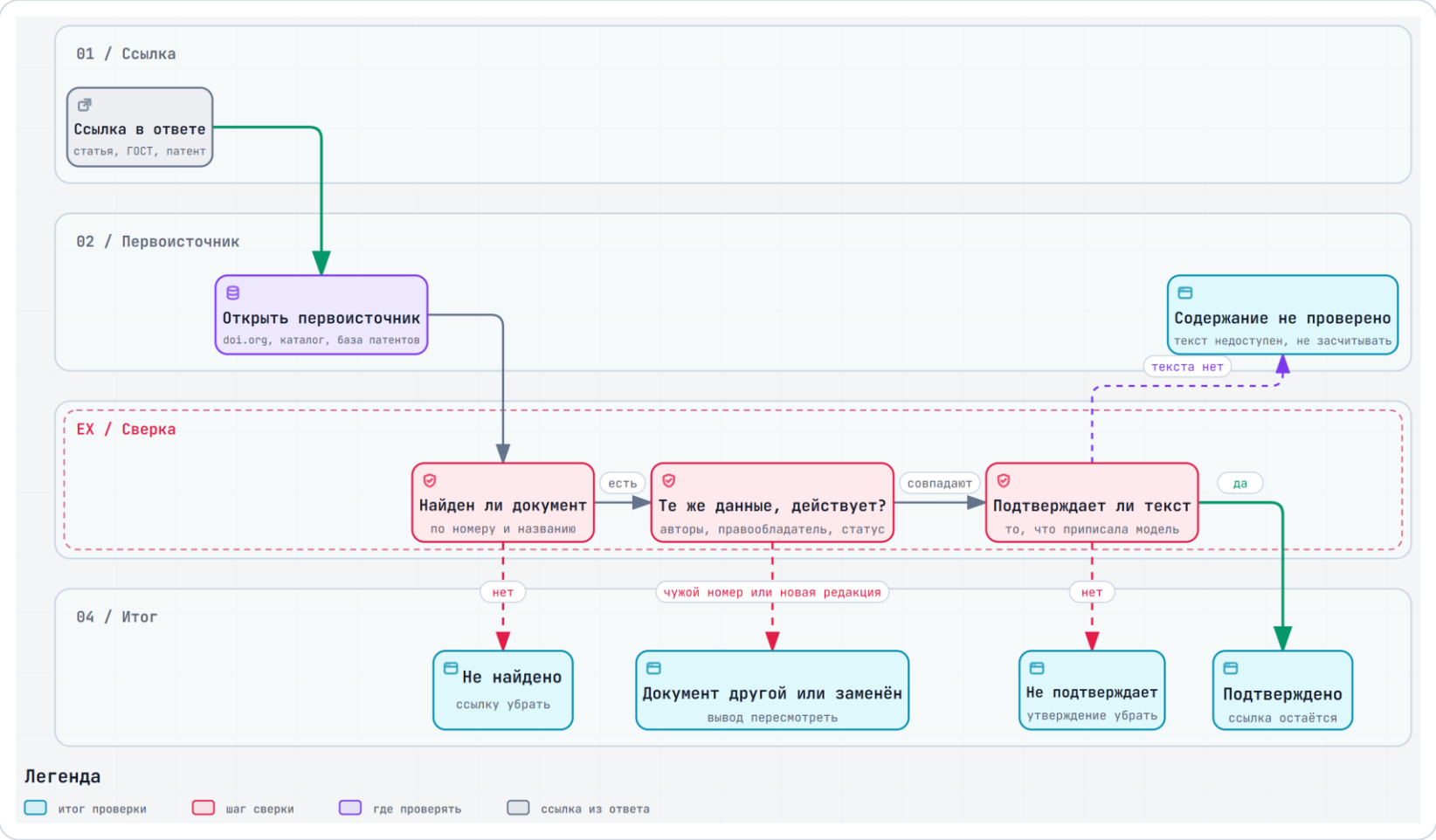
Что в ответе	Где проверять	Что сверить
Статья с DOI	https://doi.org/…	Название, авторы, год, журнал
ГОСТ	Каталог стандартов, сайт Росстандарта	Название, статус, чем заменён
Патент	Реестр ФИПС, Google Patents	Название, правообладатель, МПК
Цифра «по данным работы»	Текст работы	Есть ли цифра и в каком контексте

Шаги проверки и итоги



- Шаги: найден ли документ, совпадают ли данные, действует ли, подтверждает ли текст
- Итоги: **подтверждено**, **не найдено**, **документ другой**, **заменён**
- **Не подтверждает:** документ настоящий, утверждения в нём нет
- **Содержание не проверено:** текст недоступен, ссылку не засчитывать
- Самые коварные – «документ другой» и «не подтверждает»

Как проверить ссылку из ответа модели



Интерактивная версия – 08-Схемы/d3-proverka-istochnika.html

Информированный скептицизм



- «Доверяй, но проверяй» уже мало: ошибки модели хорошо маскируются
- Понимать, как ошибаются модели
- Оценивать риск по тому, где будет использован результат
- Выбирать способ проверки под этот риск

Faker, Toker, 2026 – работа о проверке программного кода

Задание 1. Варианты решения и отсев



45 мин

1. Получить от модели 6–8 вариантов задачи К-1 с доводами против
2. Таблица отсева: физика, ограничения 1–6 (ТЗ – редакция 1)
3. Посчитать энергобаланс хотя бы одного варианта: допустимо 13,3 А·ч за 336 ч
4. Отметить, что отсеяли бы в учебном ответе модели (документ 3.3)

```
Предложи 6–8 вариантов
решения с разными
принципами действия.
Для каждого: как работает,
какие ограничения может
нарушить,
два довода против, что
измерить или посчитать.
Отдельно –
неработоспособные варианты
с объяснением. Источники
не приводи.
```

Задание 2. Матрица выбора



45 мин

1. 2а. Проверить матрицу модели: сумма весов, итоги, состав, обоснования
2. 2б. Своя матрица на 3–4 варианта: веса, оценки, обоснования
3. 2в. Модель как оппонент в новом чате: дать задачу, ограничения, описания вариантов и матрицу; веса не менять

Задание 3. Проверка источников и готовые аналоги



35 мин

1. За, 20 мин: статья, патент и ГОСТ из документа 3.3
2. Статья – doi.org; патент – ФИПС или Google Patents; ГОСТ – статус в каталоге
3. Итог: подтверждено, не найдено, документ другой, заменён, не подтверждает
4. Зб, 15 мин: два серийных реле отключения нагрузки для варианта В1
5. Характеристики – со страницы производителя, а не из ответа модели

Задание повышенной сложности. Требования к покупному модулю

30–45 мин

1. Для тех, кто справился с основными заданиями. Можно сделать и после занятия
2. Сами считаем допустимый ток модуля по ТЗ
3. Черновик требований: числа только с источником
4. Ищем числа без источника и проверки без критерия
5. Результат: таблица требований и вопросы поставщику

Что запомнить

- Модель – для широты поиска, решение за инженером
- Сначала отсев, потом матрица
- Веса и оценки ставит инженер, у каждой оценки – обоснование
- Каждая ссылка – гипотеза, пока документ не открыт
- Режим поиска не гарантирует настоящих ссылок

Автор курса – Виталий Пиков



Регалии

Архитектор информационной безопасности · РБПО · DevSecOps · AppSec

Лектор и автор образовательных программ по ИБ и РБПО

Телефон

+7 (963) 782-92-78

E-mail

vitaly@pikov.expert

Сайт

https://pikov.expert

Telegram

@UnderLineSecurity