

# KLT-MATDES

## Проективно-реперная система обратного проектирования неорганических функциональных материалов

Проектно-техническая спецификация · версия 1.1

Документ для научного обсуждения, разработки прототипа и подготовки публикационного пакета

| Позиция                      | Содержание                                                                                |
|------------------------------|-------------------------------------------------------------------------------------------|
| Автор и руководитель проекта | Курпишев Иван Борисович                                                                   |
| Статус                       | Архитектурная и алгоритмическая спецификация; научная валидация — по отдельному протоколу |
| Предметная область           | Неорганические функциональные материалы                                                   |
| Контур реализации            | Python/SQLite-прототип → сетевой сервис → автономное приложение                           |

Калининград · 29 августа 2026 года

# Содержание

Структура рассчитана на прямую редактуру и экспорт. Номера научных рисунков и приложений стабильны.

|                                                      |    |
|------------------------------------------------------|----|
| Содержание                                           | 2  |
| 1. Резюме проекта                                    | 4  |
| 2. Назначение, границы и критерий законченности      | 4  |
| 2.1. Цель                                            | 4  |
| 2.2. Что входит в систему                            | 5  |
| 2.3. Что не входит без отдельной верификации         | 5  |
| 3. Научно-методический контур                        | 5  |
| 3.1. Формализация объекта                            | 5  |
| 3.2. Реперный пакет                                  | 5  |
| 3.3. Сравнение с базовой линией                      | 5  |
| 4. Концептуальная архитектура                        | 6  |
| 4.1. Принцип разделения                              | 6  |
| 5. Информационная модель KLT-RBD/ChemGraph           | 7  |
| 5.1. Основные сущности                               | 7  |
| 5.2. Минимальные графовые связи                      | 7  |
| 5.3. Контроль единиц и условий                       | 7  |
| 6. Алгоритмический конвейер                          | 7  |
| 6.1. Детерминированность                             | 8  |
| 7. Обратное проектирование и ранжирование кандидатов | 8  |
| 7.1. Целевая спецификация                            | 8  |
| 7.2. Вектор оценки                                   | 8  |
| 7.3. Устойчивость ранжирования                       | 9  |
| 8. Планирование путей получения                      | 9  |
| 8.1. Представление шага                              | 9  |
| 8.2. Поиск маршрута                                  | 9  |
| 9. Доказательный контур, неопределённость и ABSTAIN  | 10 |
| 9.1. Два независимых статуса                         | 10 |
| 9.2. Доказательные ворота                            | 10 |
| 9.3. Калибровка неопределённости                     | 10 |
| 10. Программная архитектура и режимы эксплуатации    | 11 |
| 10.1. Компоненты                                     | 11 |
| 10.2. Сетевой режим                                  | 11 |
| 10.3. Автономный режим                               | 11 |
| 10.4. Профили исполнения                             | 11 |
| 11. Пользовательские сценарии и возможности          | 11 |
| 11.1. Основные сценарии                              | 11 |
| 11.2. Представление результата                       | 11 |

|                                                          |    |
|----------------------------------------------------------|----|
| 12. Воспроизводимость, безопасность и управление данными | 12 |
| 12.1. Пакет воспроизведения                              | 12 |
| 12.2. Управление происхождением                          | 12 |
| 12.3. Тесты                                              | 12 |
| 13. Дорожная карта разработки и научной проверки         | 13 |
| 13.1. Первый проверяемый корпус                          | 13 |
| 13.2. Публикационные продукты                            | 13 |
| 14. Риски, метрики и критерии перехода                   | 14 |
| 14.1. Реестр ключевых рисков                             | 14 |
| 14.2. Метрики                                            | 14 |
| 14.3. Критерий перехода к внешнему тесту                 | 14 |
| 15. Заключение                                           | 14 |
| A. Приложение A. Термины, обозначения и статусы          | 15 |
| B. Приложение B. Контракты входа и результата            | 15 |
| B.1. Минимальный TargetSpec                              | 15 |
| B.2. Минимальный сертификат                              | 15 |
| C. Приложение C. Псевдокод ядра                          | 15 |
| C.1. Правило отказа                                      | 16 |
| D. Приложение D. Минимальная схема базы данных           | 16 |
| D.1. Инварианты схемы                                    | 16 |
| E. Приложение E. Контур API                              | 17 |
| F. Приложение F. Матрица испытаний                       | 17 |
| G. Приложение G. Реестр научных иллюстраций              | 17 |

# 1. Резюме проекта

KLT-MATDES — проект вычислительной системы обратного проектирования неорганических функциональных материалов по постановке «требуемые свойства → допустимый состав и структура → проверяемый путь получения → доказательный паспорт результата». Система объединяет классические методы химической информатики и материаловедения с авторским проективно-реперным представлением KLT-RBD. Её назначение — не заменить эксперимент, а управляемо сузить пространство поиска, показать основания каждого вычислительного шага и своевременно отказаться от вывода при недостатке данных.

### ГЛАВНЫЙ ПРОДУКТ

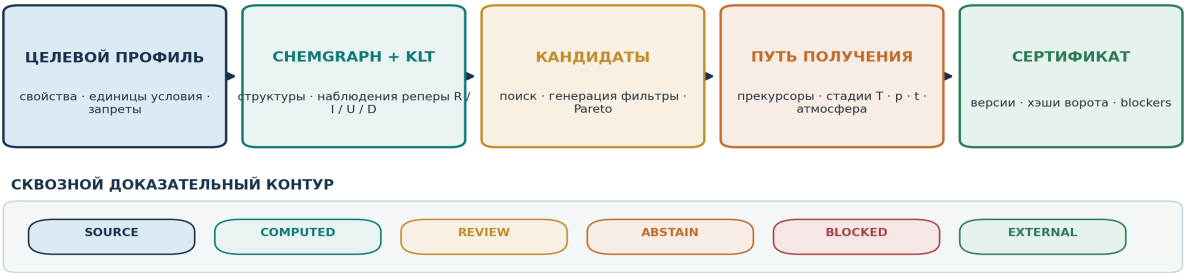
Воспроизводимый пакет кандидатов: спецификация цели, нормализованные источники, граф реперов, прогнозы свойств с неопределённостью, ограничения применимости, варианты синтеза, причины ранжирования и машиночитаемый сертификат запуска.

Первый предметный контур — неорганические функциональные материалы. Он выбран как достаточно структурированный для формализации составов, фаз, кристаллических структур, режимов обработки и измеряемых свойств, но одновременно достаточно сложный, чтобы проверить полезность реперной модели на реальных задачах многокритериального поиска.

| Компонент | Результат версии 1.1                                               | Граница утверждения                            |
|-----------|--------------------------------------------------------------------|------------------------------------------------|
| Формализм | Единая модель цели, кандидата, доказательства и отказа             | Спецификация; не экспериментальный результат   |
| Данные    | Схема SQLite и контур импорта открытых авторитетных источников     | Конкретные корпуса фиксируются манифестами     |
| Алгоритмы | Поиск, генерация, фильтрация, прогноз, Pareto-ранжирование, синтез | Каждый модуль сравнивается с базовой линией    |
| Интерфейс | Сетевой и автономный сценарии с единым контрактом                  | Функции зависят от доступных локальных моделей |
| Наука     | Протокол ретроспективной, слепой и внешней проверки                | До прохождения ворот — статус RESEARCH/ABSTAIN |

## KLT-MATDES: от целевого профиля к проверяемому решению

Обратное проектирование материала сопровождается происхождением данных, неопределённостью и правом на отказ



Выход системы — не «ответ без оговорок», а воспроизводимый пакет кандидатов с точной границей допустимого вывода.

Рисунок 1. Карта задачи KLT-MATDES.

# 2. Назначение, границы и критерий законченности

## 2.1. Цель

Цель проекта — построить проверяемую программно-математическую цепочку, которая принимает формализованный профиль свойств и ограничений, генерирует либо извлекает кандидаты, прогнозирует их характеристики, предлагает пути получения и выпускает результат только вместе с трассировкой происхождения данных и оценкой применимости.

## 2.2. Что входит в систему

- редактор целевого профиля: свойства, единицы, диапазоны, условия измерения, приоритеты и запреты;
- нормализованный ChemGraph для веществ, фаз, структур, процессов, измерений и источников;
- слой KLT-RBD для реперов R, I, U, D, связей между ними и сравнительных пакетов;
- поиск аналогов, генерация кандидатов и многоступенчатая химическая фильтрация;
- ансамбль моделей свойств с оценкой неопределённости и области применимости;
- планировщик путей получения с учётом прекурсоров, режимов, рисков и доказательности;
- многокритериальное ранжирование, Pareto-фронт, объяснение и режим ABSTAIN;
- машиночитаемый сертификат запуска, отчёт для исследователя и журнал версий.

## 2.3. Что не входит без отдельной верификации

- утверждение о существовании нового материала только на основании вычисления;
- утверждение о синтезируемости без маршрута, источников и экспериментальной проверки;
- перенос  $\lambda$ -истинности или иных математических координат в статус химической истинности;
- медицинские, экологические или промышленные гарантии безопасности;
- автоматическое повышение доказательного статуса после успешного программного теста.

### КРИТЕРИЙ ЗАКОНЧЕННОСТИ

Версия считается законченной на своём уровне, когда для каждого результата доступны вход, версия данных и кода, детерминированные параметры, ограничения применимости, неопределённость, трассировка решений, тесты и корректный статус PASS / REVIEW / ABSTAIN / BLOCKED.

# 3. Научно-методический контур

## 3.1. Формализация объекта

Материал представляется не одной формулой, а типизированным объектом  $x = (\text{composition, phase, structure, defects, morphology, process, environment})$ . Измерение свойства записывается вместе с методом, температурой, давлением, частотой, направлением, размерностью образца, неопределённостью и ссылкой на источник. Это предотвращает ложное объединение несопоставимых значений.

## 3.2. Реперный пакет

В рабочей модели KLT-пакет кандидата задаётся четверкой  $\text{Rep}(x) = (R, I, U, D)$ : R — опорные объекты и наблюдения; I — устойчивые отношения и инварианты внутри заданного класса преобразований; U — границы, неопределённости и условия применимости; D — структурная координата различия. Конкретный смысл каждого поля задаётся схемой предметного домена и не переносится между доменами по аналогии без явного морфизма.

### СТАТУС $\lambda$

$\lambda$ -профиль используется как внутренняя проективно-реперная координата согласования пакета. Он не равен вероятности, химической истинности или экспериментальному подтверждению. Любой эмпирический вывод проходит отдельный доказательный контур.

## 3.3. Сравнение с базовой линией

Научная ценность слоя KLT проверяется не демонстрацией отдельных удачных примеров, а заранее определённым сравнением с базовыми методами: поиском ближайших аналогов, регрессионными/графовыми моделями, физическими дескрипторами и тем же конвейером без KLT-признаков. Добавочная ценность фиксируется только при улучшении заранее выбранных метрик на отложенных и внешних данных.

| Уровень         | Допустимый вывод                                      | Недопустимое расширение                |
|-----------------|-------------------------------------------------------|----------------------------------------|
| Математический  | Свойство формальной конструкции при заданных аксиомах | Автоматический вывод о физическом мире |
| Вычислительный  | Код воспроизводит объявленный расчёт и тесты          | Научная валидность модели              |
| Ретроспективный | Метрика на исторически разделённых данных             | Проспективная способность              |
| Слепой          | Результат на скрытом канале по предрегистрации        | Универсальность вне области            |

| Уровень           | Допустимый вывод                                 | Недопустимое расширение              |
|-------------------|--------------------------------------------------|--------------------------------------|
| Внешний           | Независимое воспроизведение в заявленной области | Гарантия для любого материала        |
| Экспериментальный | Подтверждение конкретного образца и протокола    | Подтверждение всех кандидатов класса |

4. Концептуальная архитектура

Архитектура построена как набор изолированных слоёв с версионизируемыми контрактами. Это позволяет заменять источник данных или модель прогноза, не изменяя смысл сертификата и не смешивая научный статус с техническим состоянием сервиса.

| Слой                | Основные модули                                     | Выход                                 |
|---------------------|-----------------------------------------------------|---------------------------------------|
| 1. Пользовательский | TargetSpec Editor, сценарии, сравнение кандидатов   | Целевая спецификация и запрос запуска |
| 2. Оркестрация      | Run Manager, очередь, политика режимов              | Детерминированный план вычисления     |
| 3. Предметное ядро  | Normalizer, ChemGraph, KLT Mapper, Candidate Engine | Типизированные кандидаты и признаки   |
| 4. Модели           | Property Ensemble, Applicability, Uncertainty       | Прогнозы и интервалы                  |
| 5. Синтез           | Route Planner, precursor/process constraints        | Ранжированные маршруты                |
| 6. Доказательства   | Evidence Gate, ABSTAIN Policy, Certifier            | Статус и сертификат                   |
| 7. Данные           | SQLite/PostgreSQL, файловое хранилище, манифесты    | Версионированный корпус               |

Семислойная архитектура KLT-MATDES

Версионизируемые контракты позволяют заменять модели и источники без подмены смысла результата

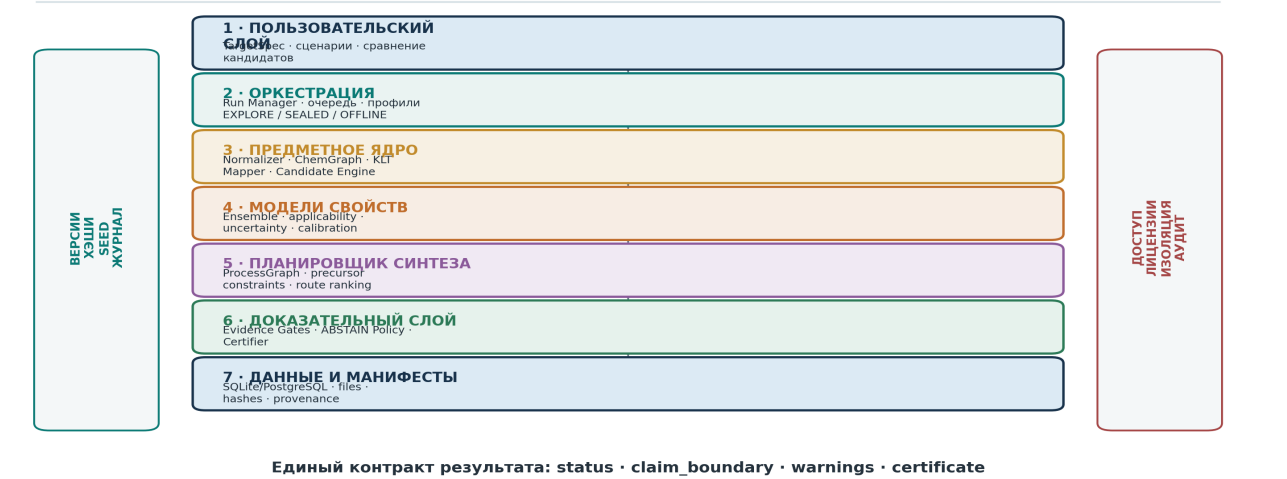


Рисунок 2. Семислойная архитектура.

4.1. Принцип разделения

- данные источника сохраняются отдельно от нормализованных представлений;
- расчётные признаки не перезаписывают исходные наблюдения;
- предсказание свойства не считается фактом базы до внешнего подтверждения;
- маршрут синтеза хранится как гипотеза с набором предпосылок;
- интерфейс отображает ограничения и причины отказа наравне с успешными результатами.

## 5. Информационная модель KLT-RBD/ChemGraph

### 5.1. Основные сущности

| Сущность            | Ключевые поля                                      | Правило качества                                 |
|---------------------|----------------------------------------------------|--------------------------------------------------|
| Material            | состав, идентификаторы, заряд, чистота             | формула нормализована; исходная строка сохранена |
| Phase/Structure     | пространственная группа, решётка, позиции, дефекты | структура связана с конкретной фазой             |
| Sample              | форма, размер, морфология, история                 | не смешивать образцы с разной обработкой         |
| PropertyObservation | значение, единица, метод, условия, $\sigma$        | единица и контекст обязательны                   |
| ProcessStep         | операция, прекурсоры, T, p, t, атмосфера           | каждый параметр имеет источник/диапазон          |
| Source              | DOI/URL/файл, лицензия, версия, страницы           | трассировка до первичного фрагмента              |
| ReperNode/Edge      | тип R/I/U/D, связь, вес, основание                 | семантика задана схемой домена                   |
| Prediction          | модель, версия, значение, интервал, applicability  | не записывать как observation                    |
| RunCertificate      | вход, hashes, код, данные, статус, blockers        | неизменяемый после выпуска                       |

### 5.2. Минимальные графовые связи

Граф содержит связи HAS\_PHASE, HAS\_STRUCTURE, PREPARED\_BY, MEASURED\_AS, UNDER\_CONDITION, DERIVED\_FROM, ANALOG\_OF, TRANSFORMS\_TO, SUPPORTS, CONTRADICTS и OUTSIDE\_DOMAIN. Каждое ребро имеет тип, направление, источник, уровень доказательности и временную метку версии. Для KLT-слоя дополнительно фиксируется пакет, класс допустимых преобразований и правило вычисления координат.

#### ChemGraph и проективно-реперный пакет

Наблюдения, процессы и источники образуют предметный граф; KLT-слой добавляется как типизированная надстройка

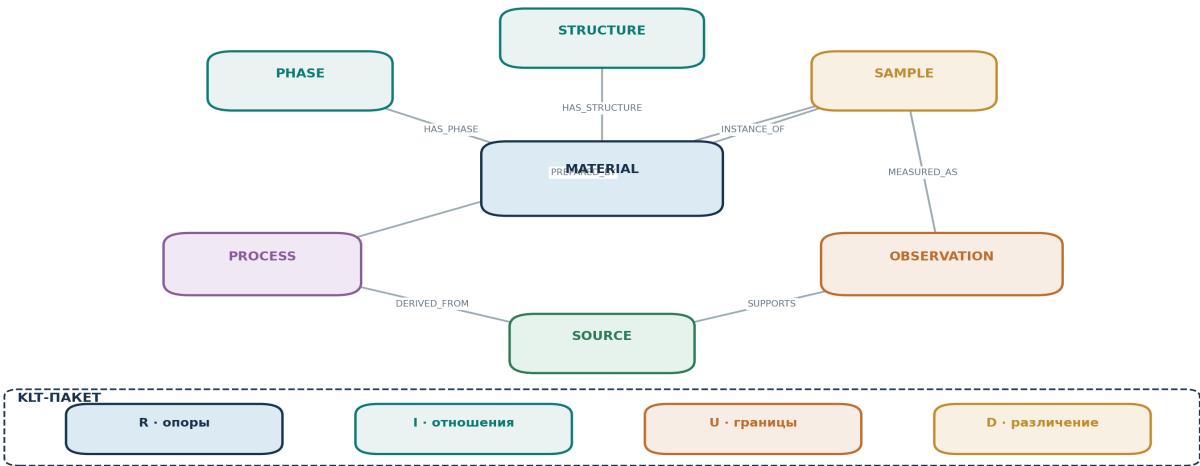


Рисунок 3. ChemGraph и реперный пакет.

### 5.3. Контроль единиц и условий

Все численные свойства хранятся в исходной и канонической единице. Преобразование сопровождается формулой, версией справочника и тестом размерности. Если условия измерения отсутствуют, запись остаётся доступной для поиска, но может быть исключена из обучения или понижена по доказательному весу.

## 6. Алгоритмический конвейер

1. Принять TargetСpec: целевые диапазоны, обязательные и мягкие ограничения, условия эксплуатации и веса целей.
2. Проверить размерности, совместимость требований и полноту обязательных полей; при противоречии вернуть BLOCKED.
3. Сформировать область  $\Omega$  поиска: классы составов, фаз, структур, допустимых элементов и процессов.
4. Найти опорные материалы и наблюдения; построить локальный ChemGraph и реперный пакет.

5. Сгенерировать кандидаты поиском по базе, заменами, допированием, прототипированием структур и подключаемыми генеративными моделями.
6. Применить жёсткие фильтры химической допустимости, состава, стоимости, токсикологического запрета и технологических ограничений.
7. Получить прогнозы свойств ансамблем моделей; вычислить неопределённость и область применимости.
8. Рассчитать KLT-координаты сравнительного пакета; не использовать их вне зафиксированной схемы домена.
9. Построить Pareto-фронт и устойчивое ранжирование при вариации весов и неопределённостей.
10. Для верхних кандидатов построить пути получения; проверить доступность прекурсоров и критические условия.
11. Применить доказательные ворота. При недостатке данных — ABSTAIN с точным реестром блокиров.
12. Выпустить отчёт и сертификат запуска; сохранить хеши входа, корпуса, моделей, кода и результатов.

## Алгоритмический конвейер KLT-MATDES

Двенадцать воспроизводимых стадий с явными ветвями BLOCKED и ABSTAIN

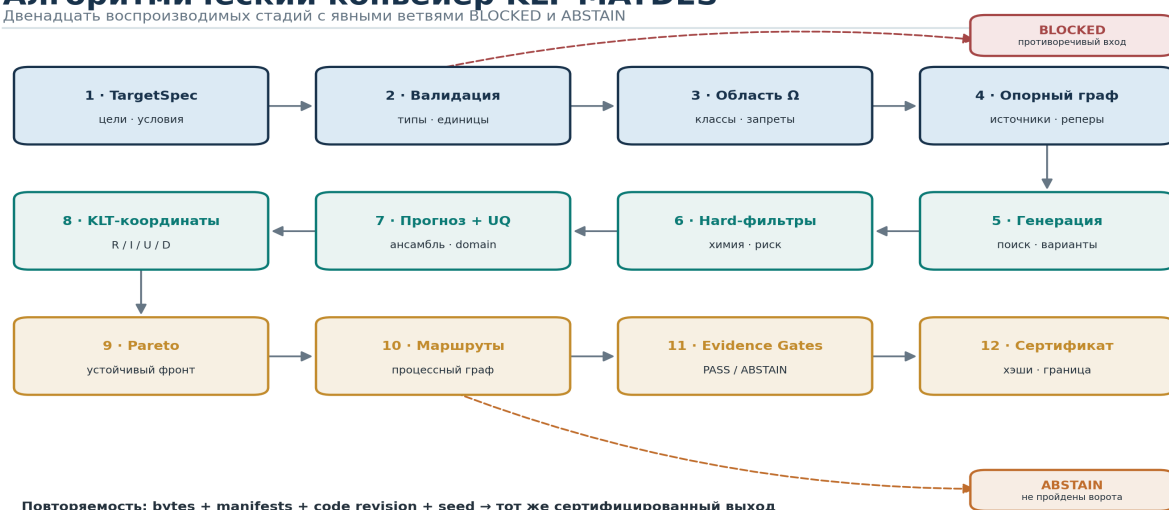


Рисунок 4. Конвейер «цель → кандидат → маршрут → сертификат».

### 6.1. Детерминированность

Любой стохастический модуль получает фиксируемый seed. Порядок фильтров, версии моделей, гиперпараметры и состояние справочников входят в RunManifest. Повтор запуска на тех же байтах должен воспроизводить список кандидатов либо явно сообщать о недетерминированной зависимости.

## 7. Обратное проектирование и ранжирование кандидатов

### 7.1. Целевая спецификация

TargetSpec = (Y\*, C\_hard, C\_soft, E, W), где Y\* — целевые свойства и интервалы; C\_hard — запреты и необходимые условия; C\_soft — предпочтения; E — условия эксплуатации и измерения; W — веса либо правило Pareto-сравнения. Если пользователь не может задать веса, система возвращает фронт не доминируемых решений и сценарный анализ.

### 7.2. Вектор оценки

Для кандидата  $x$  формируется  $q(x) = (d\_property, u\_model, a\_domain, e\_evidence, c\_synthesis, r\_risk, k\_KLT)$ . Здесь  $d\_property$  — отклонение от цели;  $u\_model$  — неопределённость;  $a\_domain$  — близость к области обучения;  $e\_evidence$  — качество источников;  $c\_synthesis$  — оценка сложности пути;  $r\_risk$  — штраф риска;  $k\_KLT$  — реперная согласованность по конкретной схеме. Скалярный рейтинг допускается только как представление выбранного сценария, а не как абсолютная мера качества.



| Механизм генерации       | Назначение                        | Обязательная защита                               |
|--------------------------|-----------------------------------|---------------------------------------------------|
| Поиск аналогов           | быстрый базовый список            | разделение наблюдаемых и прогнозных свойств       |
| Элементные замены        | локальное расширение состава      | валентность, заряд, радиусы, запрещённые элементы |
| Допирование              | тонкая настройка свойств          | диапазон концентраций и фазовая устойчивость      |
| Структурные прототипы    | поиск изоморфных семейств         | совместимость состава и прототипа                 |
| Оптимизатор по суррогату | эффективный поиск дорогой функции | uncertainty + stopping rule                       |
| Генеративный модуль      | дальний поиск                     | жёсткая валидация и маркировка гипотезы           |

7.3. Устойчивость ранжирования

Вместе с позицией кандидата выводятся интервалы ранга при вариации весов, границ свойств и модельной ошибки. Кандидат, остающийся на Pareto-фронте в большинстве допустимых сценариев, маркируется как устойчивый. Перестановка кандидатов при малом изменении предпосылок показывается пользователю как чувствительность, а не скрывается усреднением.

8. Планирование путей получения

Планировщик синтеза работает по графу процессов: узлами являются вещества, фазы и промежуточные состояния, рёбрами — операции с параметрами. Маршрут оценивается по химической правдоподобности, источникам, доступности прекурсоров, числу стадий, диапазонам температуры/давления, чувствительности к атмосфере, ожидаемой селективности, отходам и риску.

8.1. Представление шага

ProcessStep = (inputs, operation, T, p, t, atmosphere, additives, equipment, outputs, evidence, uncertainty).  
Неизвестный параметр не заполняется правдоподобным значением автоматически: он остаётся OPEN и включается в список экспериментальных факторов.

8.2. Поиск маршрута

- ретросинтетический поиск от целевой фазы к доступным прекурсорам;
- прямой поиск от библиотеки прекурсоров к целевому узлу;
- beam search/A\* с ограничениями и штрафами неопределённости;
- альтернативные маршруты с различными профилями оборудования и риска;
- остановка с ABSTAIN, если отсутствует связный путь достаточной доказательности.

Граф альтернативных маршрутов получения

Маршрут является гипотезой, пока его параметры и переходы не связаны с источниками и экспериментом

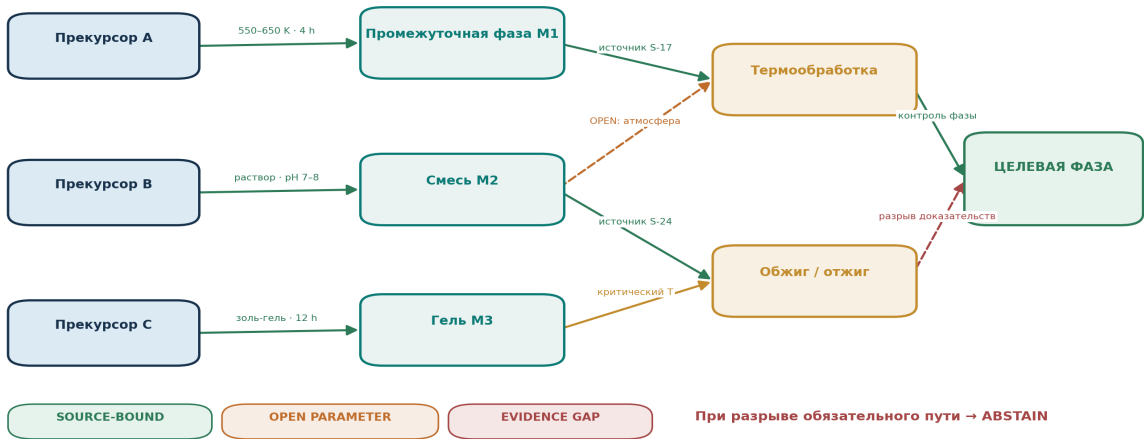


Рисунок 5. Граф маршрутов получения.

БЕЗОПАСНОСТЬ

Документ описывает вычислительное планирование. Практический синтез требует отдельной оценки опасностей, SDS, совместимости оборудования, локальных регламентов и одобрения квалифицированного специалиста.

## 9. Доказательный контур, неопределённость и ABSTAIN

### 9.1. Два независимых статуса

Каждый результат имеет технический статус выполнения и научный статус утверждения. Успешное завершение кода означает только то, что расчёт выполнен согласно контракту; научный статус определяется качеством данных, дизайном проверки и областью применимости.

| Статус  | Смысл                                                             | Типичный триггер                                           |
|---------|-------------------------------------------------------------------|------------------------------------------------------------|
| PASS    | расчёт завершён, все обязательные ворота текущего уровня пройдены | полный вход, модели применимы, сертификат собран           |
| REVIEW  | результат пригоден для экспертного рассмотрения                   | ограниченная неоднозначность или чувствительность          |
| ABSTAIN | система сознательно не ранжирует или не рекомендует               | вне области, высокая неопределённость, конфликт источников |
| BLOCKED | вычисление не должно начинаться/продолжаться                      | противоречивый вход, отсутствуют обязательные данные       |

### 9.2. Доказательные ворота

1. Source Gate: происхождение и лицензия данных известны; критические поля имеют ссылки.
2. Unit Gate: единицы и условия сопоставимы; преобразования проверены.
3. Domain Gate: кандидат находится внутри или около объявленной области применимости.
4. Uncertainty Gate: интервал прогноза совместим с целевым диапазоном и порогом решения.
5. Baseline Gate: улучшение относительно базовой линии измерено на корректном разделении.
6. Leakage Gate: отсутствует утечка между обучением, подбором и испытанием.
7. Synthesis Gate: маршрут имеет опорные источники либо явно маркирован как гипотеза.
8. Reproducibility Gate: вход, версии, хеши и seed достаточны для повторения.

### Лестница доказательности KLT-MATDES

Переход между уровнями возможен только после выполнения заранее зафиксированных ворот

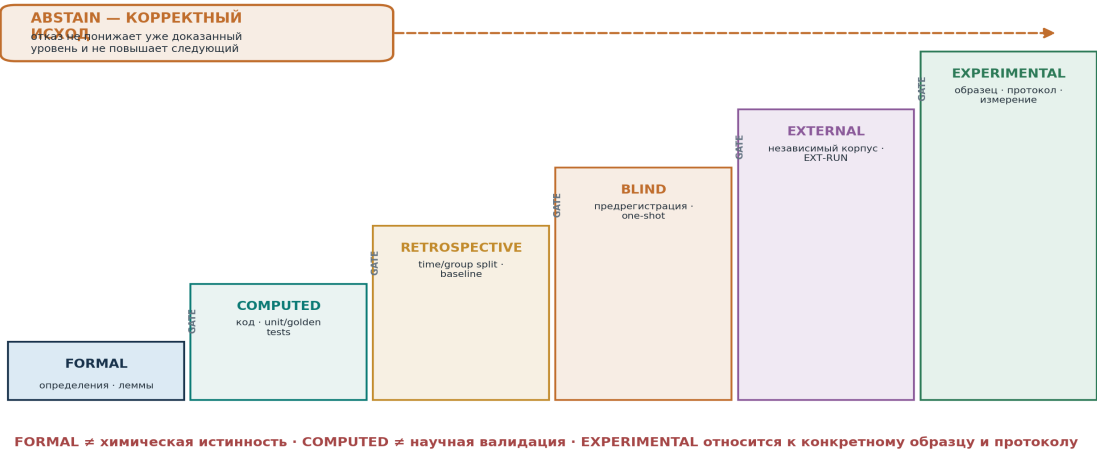


Рисунок 6. Лестница доказательности.

### 9.3. Калибровка неопределённости

Предпочтителен ансамбль источников неопределённости: разброс моделей, ошибки измерений, вариативность образцов, расстояние до области обучения и несовпадение условий. Интервалы проверяются на калибровочном наборе. Если покрытие хуже предзаданного допуска, модель не допускается к автоматическому ранжированию.

## 10. Программная архитектура и режимы эксплуатации

### 10.1. Компоненты

| Компонент      | Рекомендуемая реализация            | Ответственность                                 |
|----------------|-------------------------------------|-------------------------------------------------|
| Core           | Python package                      | типы, нормализация, граф, алгоритмы, сертификат |
| Database       | SQLite локально / PostgreSQL сервер | данные, версии, транзакции, индексы             |
| API            | REST/JSON + пакетный CLI            | единый контракт сетевого и автономного режима   |
| Workers        | очередь задач                       | долгие расчёты, отмена, повтор, лимиты          |
| Web UI         | адаптивный интерфейс                | TargetSpec, графы, сравнение, отчёты            |
| Offline UI     | Flutter/desktop shell               | локальные корпуса, приватный расчёт             |
| Model Registry | файлы + manifest                    | версии, метрики, область, подписи               |
| Report Engine  | DOCX/PDF/JSON                       | человекочитаемый отчёт и certificate.json       |

### 10.2. Сетевой режим

Сетевой сервис предоставляет актуальные модели, централизованный корпус и тяжёлые вычисления. Пользователь видит очередь, оценку ресурсов, источники и статус обработки. Загруженные данные изолируются по проектам; публикация в общий корпус требует отдельного решения и лицензионной проверки.

### 10.3. Автономный режим

Автономная сборка использует локальную SQLite-базу, компактные модели и те же JSON-контракты. Она обязана выдавать сопоставимые статусы и сертификаты, даже если глубина поиска ниже. Синхронизация выполняется пакетами с манифестом и конфликт-политикой; отсутствие сети не должно приводить к скрытой подмене версии данных.

### 10.4. Профили исполнения

| Профиль   | Назначение                               | Ограничения                              |
|-----------|------------------------------------------|------------------------------------------|
| EXPLORE   | интерактивное исследование пространства  | результаты не считаются слепой проверкой |
| REPRODUCE | повтор известного запуска                | никакого подбора параметров              |
| SEALED    | предрегистрационный скрытый канал        | одноразовое раскрытие и аудит            |
| OFFLINE   | локальная работа с фиксированным пакетом | отчёт указывает дату/версию корпуса      |

## 11. Пользовательские сценарии и возможности

### 11.1. Основные сценарии

| Сценарий                      | Вход                                | Результат                                  |
|-------------------------------|-------------------------------------|--------------------------------------------|
| Подбор материала              | целевой профиль и ограничения       | Pareto-список кандидатов с основаниями     |
| Замена дефицитного элемента   | исходный материал + запрет          | варианты замещения и ожидаемый компромисс  |
| Оптимизация процесса          | кандидат + диапазоны параметров     | план экспериментов и чувствительность      |
| Аудит литературного кандидата | структура/статья/данные             | карта доказательств, конфликтов и пробелов |
| Планирование синтеза          | целевая фаза + доступные прекурсоры | альтернативные маршруты и blockers         |
| Сравнение моделей             | фиксированный корпус и split        | метрики baseline/KLT/ensemble              |
| Слепая проверка               | предрегистрация + скрытый канал     | одноразовый сертификат результата          |

### 11.2. Представление результата

- карточка кандидата: состав, фаза, структура, свойства и условия;
- граф оснований: какие источники и аналоги повлияли на вывод;
- интервалы и область применимости для каждого предсказания;
- Pareto-диаграмма и анализ чувствительности весов;
- маршруты получения с критическими параметрами и альтернативами;
- реестр OPEN/BLOCKER и минимальный следующий эксперимент;
- экспорт DOCX/PDF для исследователя, CSV для анализа и JSON для воспроизведения.

## Рабочий экран исследователя

Макет интерфейса: постановка задачи, устойчивый Pareto-список и доказательная карточка кандидата

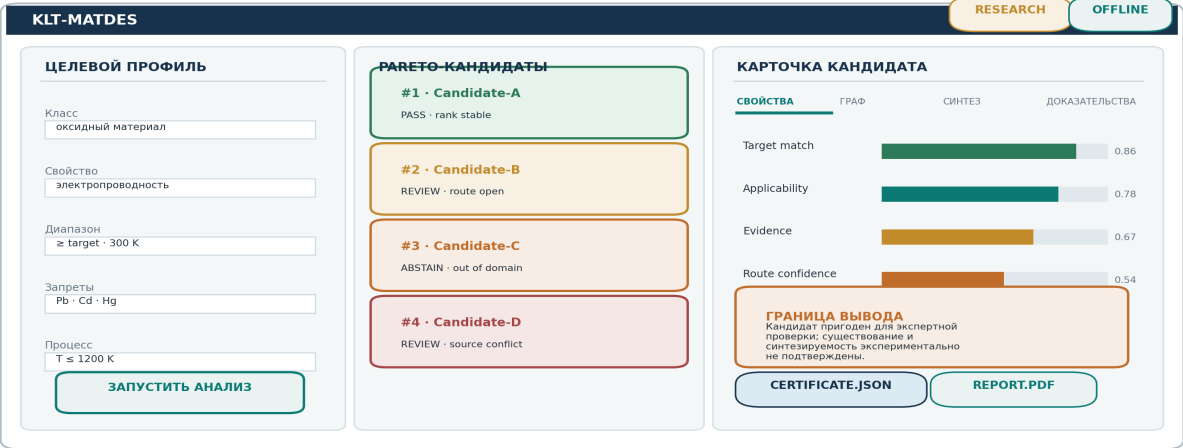


Рисунок 7. Макет рабочего экрана.

## 12. Воспроизводимость, безопасность и управление данными

### 12.1. Пакет воспроизведения

Каждый существенный запуск создаёт каталог с target\_spec.json, run\_manifest.json, corpus\_manifest.json, model\_manifest.json, predictions.parquet/csv, routes.json, certificate.json, report.docx/pdf и tests.log. Манифесты содержат SHA-256 контрольные суммы, версии схем и программного окружения.

### 12.2. Управление происхождением

- источник и лицензия фиксируются до включения записи в обучающий корпус;
- исправления выполняются add-only либо через версионированную запись замены;
- удаление недостоверной записи не стирает факт её прежнего использования в запуске;
- ручные решения эксперта подписываются ролью, датой и причиной;
- секреты, персональные и коммерческие данные не входят в отчёт и хешируются/шифруются по политике проекта.

### 12.3. Тесты

| Класс тестов | Что проверяется                     | Критерий                             |
|--------------|-------------------------------------|--------------------------------------|
| Unit         | парсеры, единицы, графовые операции | точные ожидаемые значения            |
| Property     | инварианты схемы и преобразований   | нет нарушения типов и размерностей   |
| Golden       | эталонные маленькие корпуса         | совпадение сертифицированного выхода |
| Integration  | API, DB, модели, отчёт              | полный воспроизводимый запуск        |
| Scientific   | split, baseline, калибровка         | предзаданные метрики и пороги        |
| Adversarial  | неполный/противоречивый вход        | корректный ABSTAIN/BLOCKED           |
| Security     | доступ, загрузки, архивы            | изоляция и отсутствие утечек         |

### 13. Дорожная карта разработки и научной проверки

| Этап              | Горизонт    | Основной результат                              | Ворота выхода                    |
|-------------------|-------------|-------------------------------------------------|----------------------------------|
| M0. Контракт      | 0–2 мес.    | схемы TargetSpec/ChemGraph/Certificate, эталоны | schema tests; status firewall    |
| M1. Прототип      | 2–4 мес.    | Python/SQLite, CLI, импорт, аналоги             | воспроизводимый golden run       |
| M2. Модели        | 4–7 мес.    | baseline, uncertainty, Pareto                   | временной split; калибровка      |
| M3. KLT-сравнение | 7–10 мес.   | абляция KLT-признаков                           | предзаданный эффект vs baseline  |
| M4. Слепой контур | 10–14 мес.  | sealed hidden-channel run                       | предрегистрация; one-shot reveal |
| M5. EXT-RUN       | 14–18 мес.  | независимая сборка и расчёт                     | внешний отчёт                    |
| M6. Эксперимент   | 18–24 мес.  | малый пилот синтеза                             | протокол, образцы, измерения     |
| M7. Продукт       | после M5/M6 | сетевой и автономный релиз                      | безопасность и SLA               |

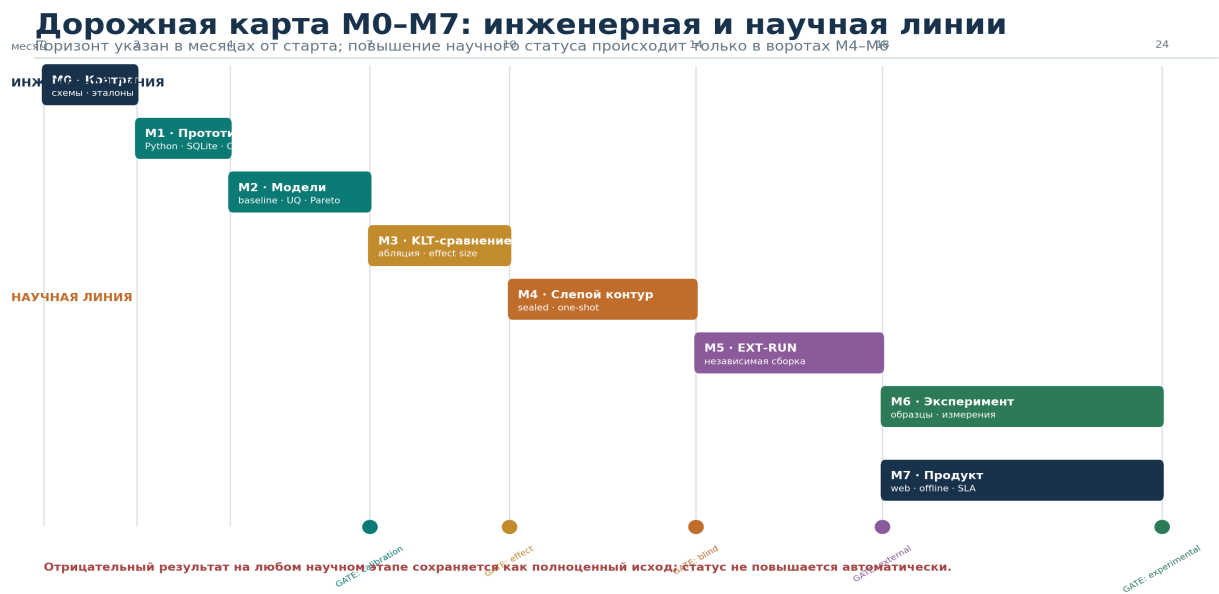


Рисунок 8. Дорожная карта M0–M7.

#### 13.1. Первый проверяемый корпус

Для первого цикла рекомендуется выбрать один однородный класс неорганических функциональных материалов с достаточным числом сопоставимых структур, единообразными условиями свойства и доступными отрицательными примерами. Конкретный класс закрепляется отдельной предрегистрацией после аудита покрытия, лицензий и риска утечки.

#### 13.2. Публикационные продукты

- методическая статья: формализм, доказательные статусы и архитектура;
- data descriptor: корпус, схема, качество и ограничения;
- benchmark paper: baseline, KLT-абляция, временной и внешний тест;
- software paper: воспроизводимый прототип и сертификаты;
- experimental report: результаты пилота независимо от знака эффекта.

## 14. Риски, метрики и критерии перехода

### 14.1. Реестр ключевых рисков

| Риск                      | Ранний индикатор                  | Снижение риска                                |
|---------------------------|-----------------------------------|-----------------------------------------------|
| Неоднородность данных     | большая межисточниковая дисперсия | узкий домен, условия как часть объекта        |
| Утечка                    | аномально высокая метрика         | time/group split, аудит дублей                |
| Слабая добавка KLT        | абляция не ухудшает результат     | публиковать отрицательный вывод; не усложнять |
| Переуверенность           | плохое покрытие интервалов        | калибровка и ABSTAIN                          |
| Несинтезируемые кандидаты | нет связанных маршрутов           | route gate до финального ранга                |
| Смещение источников       | доминирование одной лаборатории   | source-aware split и внешние данные           |
| Невоспроизводимость       | запуски расходятся                | контейнер, seeds, hashes, golden tests        |
| Лицензионные ограничения  | неясный статус данных             | реестр лицензий и исключение до импорта       |

### 14.2. Метрики

| Контур         | Метрики                                              | Правило интерпретации                   |
|----------------|------------------------------------------------------|-----------------------------------------|
| Свойства       | MAE/RMSE, rank correlation, coverage                 | по материалам/семействам, не по строкам |
| Поиск          | recall@k, enrichment factor                          | на временно скрытых открытиях           |
| Ранжирование   | hypervolume, regret, rank stability                  | с учётом неопределённости               |
| Синтез         | route validity, precursor coverage, expert agreement | не равно экспериментальному выходу      |
| Доказательства | source coverage, blocker rate, abstention quality    | штраф за уверенную ошибку выше отказа   |
| Инженерия      | reproducibility, latency, failure rate               | порог зависит от профиля исполнения     |

### 14.3. Критерий перехода к внешнему тесту

Переход допускается, если зафиксированы данные и split; baseline реализован без KLT; метрики и минимальный эффект предзаданы; uncertainty откалибрована; утечки проверены; код и контейнер воспроизводимы; есть процедура one-shot reveal; отрицательный результат принят как допустимый исход.

## 15. Заключение

KLT-MATDES спроектирован как исследовательская система с сильной дисциплиной происхождения данных и доказательных границ. Его отличительный замысел — связать обратный дизайн материалов с реперным представлением KLT-RBD, но проверять полезность этого слоя наравне с любой другой моделью: через базовые линии, абляцию, слепое разделение и независимое воспроизведение.

Ближайший законченный результат — Python/SQLite-прототип узкого класса материалов, который строит ChemGraph, принимает TargetSpec, формирует кандидатов, оценивает свойства и неопределённость, планирует маршрут, умеет отказать и выпускает сертификат. Только после прохождения ретроспективных, слепых, внешних и экспериментальных ворот допустимо повышать научный статус конкретных утверждений.

#### РЕШЕНИЕ ДЛЯ СТАРТА

Заморозить схемы TargetSpec v1, ChemGraph v1 и RunCertificate v1; собрать малый source-bound корпус; реализовать golden run; затем выбирать первый доменный benchmark на основании покрытия и риска утечки.

## А. Приложение А. Термины, обозначения и статусы

| Термин        | Рабочее определение                                                                    |
|---------------|----------------------------------------------------------------------------------------|
| TargetSpec    | машиночитаемая постановка целевых свойств, условий, ограничений и приоритетов          |
| ChemGraph     | типизированный граф материалов, структур, наблюдений, процессов и источников           |
| KLT-RBD       | реперная база данных и операции над пакетами R/I/U/D в заданной предметной схеме       |
| R             | опорные объекты, наблюдения и источники                                                |
| I             | отношения/инварианты при объявленном классе преобразований                             |
| U             | неопределённости, границы и условия применимости                                       |
| D             | структурная координата различения, определённая в конкретной схеме                     |
| Ω             | нормированная область поиска и применимости                                            |
| λ-профиль     | внутренняя проективно-реперная координата; не вероятность и не эмпирическая истинность |
| Pareto-фронт  | кандидаты, не ухудшаемые по одной цели без ухудшения по другой                         |
| Applicability | оценка того, насколько запрос и кандидат представлены обучающими данными               |
| ABSTAIN       | контролируемый отказ от содержательного вывода                                         |
| BLOCKED       | остановка из-за некорректного/неполного обязательного входа                            |
| EXT-RUN       | независимая повторная сборка корпуса и расчёт по фиксированному протоколу              |

## В. Приложение В. Контракты входа и результата

### В.1. Минимальный TargetSpec

```
{
  "schema": "klt-matdes.target.v1",
  "material_class": "<фиксируемый класс>",
  "target_properties": [
    {
      "name": "<property>",
      "min": null,
      "max": null,
      "unit": "<SI/unit>",
      "conditions": {"temperature_K": null},
      "priority": "hard|soft",
      "weight": 1.0
    }
  ],
  "allowed_elements": [],
  "forbidden_elements": [],
  "process_limits": {},
  "evidence_policy": "research|sealed|external",
  "seed": 20260829
}
```

### В.2. Минимальный сертификат

```
{
  "schema": "klt-matdes.certificate.v1",
  "run_id": "<uuid>",
  "status": "PASS|REVIEW|ABSTAIN|BLOCKED",
  "created_utc": "<ISO-8601>",
  "input_sha256": "<hash>",
  "corpus_manifest_sha256": "<hash>",
  "model_manifest_sha256": "<hash>",
  "code_revision": "<immutable revision>",
  "seed": 20260829,
  "domain": {"inside": false, "score": null, "reason": ""},
  "gates": [{"name": "Source Gate", "pass": false, "evidence": []}],
  "candidates": [],
  "blockers": [],
  "claim_boundary": "<точная граница допустимого вывода>"
}
```

## С. Приложение С. Псевдокод ядра

```
function design_material(target_spec, manifests):
    validate_schema(target_spec)
    if contradictory(target_spec):
        return certificate(BLOCKED, blockers)

    omega = define_search_domain(target_spec, manifests.corpus)
    anchors = retrieve_evidence_bound_anchors(omega, target_spec)
    chem_graph = build_typed_chemgraph(anchors)
    reper_pack = map_klt_representation(chem_graph, omega)

    candidates = generate_candidates(target_spec, chem_graph)
    candidates = apply_hard_chemical_filters(candidates, target_spec)
    if candidates is empty:
        return certificate(ABSTAIN, ["NO_ADMISSIBLE_CANDIDATES"])

    for x in candidates:
        x.predictions = predict_properties_with_uncertainty(x)
        x.applicability = assess_domain(x)
        x.klt_features = compute_klt_features(x, reper_pack)
        x.routes = plan_synthesis_routes(x, target_spec.process_limits)
        x.evidence = trace_sources(x)

    gates = evaluate_evidence_gates(candidates, target_spec)
    if gates.require_abstention:
        return certificate(ABSTAIN, gates.blockers)

    front = robust_pareto_rank(candidates, uncertainty=true)
    return certificate(PASS, front, gates, manifests)
```

C.1. Правило отказа

```
ABSTAIN if any mandatory condition holds:
    missing_unit_or_conditions
    OR out_of_applicability_domain
    OR uncertainty_crosses_decision_boundary
    OR unresolved_source_conflict
    OR no_evidence_bound_synthesis_route
    OR reproducibility_manifest_incomplete
```

D. Приложение D. Минимальная схема базы данных

| Таблица       | Назначение                            | Ключевые связи                    |
|---------------|---------------------------------------|-----------------------------------|
| materials     | канонические составы и идентификаторы | material_id                       |
| phases        | фазы и диапазоны существования        | material_id → materials           |
| structures    | кристаллические структуры             | phase_id → phases                 |
| samples       | конкретные образцы                    | phase_id, process_id              |
| properties    | словарь свойств и размерностей        | property_id                       |
| observations  | измерения с условиями и σ             | sample_id, property_id, source_id |
| processes     | маршруты и этапы получения            | process_id                        |
| process_steps | операции, параметры, порядок          | process_id                        |
| sources       | происхождение, лицензия, версия       | source_id                         |
| reper_nodes   | R/I/U/D узлы                          | reper_id, entity_ref              |
| reper_edges   | типизированные связи                  | from_id, to_id, source_id         |
| models        | версии моделей и области              | model_id                          |
| predictions   | расчёты, не наблюдения                | run_id, model_id, candidate_id    |
| runs          | параметры запуска и статус            | run_id                            |
| certificates  | неизменяемый паспорт                  | run_id, hash                      |

D.1. Инварианты схемы

- prediction никогда не обновляет observation;
- observation без unit/conditions получает признак incomplete;
- каждый производный объект имеет DERIVED\_FROM;
- сертификат ссылается на неизменяемые манифесты;
- удаление source не разрушает историческую трассировку запуска.



## Е. Приложение Е. Контур API

| Метод | Маршрут                    | Назначение                       |
|-------|----------------------------|----------------------------------|
| POST  | /v1/targets/validate       | валидация TargetSpec без запуска |
| POST  | /v1/runs                   | создать запуск                   |
| GET   | /v1/runs/{id}              | статус и прогресс                |
| POST  | /v1/runs/{id}/cancel       | контролируемая отмена            |
| GET   | /v1/runs/{id}/candidates   | список и Pareto-фронт            |
| GET   | /v1/candidates/{id}        | карточка кандидата               |
| GET   | /v1/candidates/{id}/routes | маршруты получения               |
| GET   | /v1/runs/{id}/certificate  | certificate.json                 |
| GET   | /v1/runs/{id}/report       | DOCX/PDF отчёт                   |
| POST  | /v1/corpora/import         | контролируемый импорт корпуса    |
| GET   | /v1/manifests/{hash}       | получение версии/манифеста       |

### ПРАВИЛО API

Любой ответ с кандидатом включает status, claim\_boundary, warnings и ссылку на certificate. HTTP 200 не тождествен научному PASS.

## Ф. Приложение Ф. Матрица испытаний

| ID   | Испытание                       | Вход                              | Ожидаемый исход                    |
|------|---------------------------------|-----------------------------------|------------------------------------|
| T-01 | валидная узкая цель             | полные свойства/условия           | PASS/REVIEW + candidates           |
| T-02 | противоречивые hard-ограничения | несовместимые диапазоны           | BLOCKED + причина                  |
| T-03 | нет единицы                     | число без dimension               | BLOCKED                            |
| T-04 | кандидат вне домена             | дальний состав/структура          | ABSTAIN                            |
| T-05 | конфликт источников             | несовместимые observations        | REVIEW/ABSTAIN + карта             |
| T-06 | нет маршрута                    | изолированный candidate           | ABSTAIN synthesis                  |
| T-07 | повтор запуска                  | те же bytes/seed                  | идентичные hashes/output           |
| T-08 | изменён корпус                  | новый manifest                    | новый certificate; старый сохранён |
| T-09 | baseline vs KLT                 | фиксированный split               | абляционная таблица                |
| T-10 | скрытый канал                   | sealed test                       | one-shot certificate               |
| T-11 | опасная рекомендация            | неполные risk data                | ABSTAIN + safety warning           |
| T-12 | офлайн пакет                    | фиксированная SQLite/model bundle | тот же контракт статусов           |

## Г. Приложение Г. Реестр научных иллюстраций

| № | Название                   | Финальное техническое задание                                |
|---|----------------------------|--------------------------------------------------------------|
| 1 | Карта задачи KLT-MATDES    | объединить постановку, граф, кандидаты, маршрут и сертификат |
| 2 | Семислойная архитектура    | показать слои и сквозные контуры версий/безопасности         |
| 3 | ChemGraph и реперный пакет | типизированные узлы, R/I/U/D и доказательные рёбра           |
| 4 | Алгоритмический конвейер   | 12 этапов, ветви BLOCKED и ABSTAIN                           |
| 5 | Граф маршрутов получения   | альтернативы, условия, источники, разрыв                     |
| 6 | Лестница доказательности   | формальное → вычислительное → внешнее → экспериментальное    |
| 7 | Макет рабочего экрана      | TargetSpec, Pareto-список, карточка кандидата                |
| 8 | Дорожная карта M0-M7       | инженерная и научная дорожки с воротами                      |

Все иллюстрации имеют стабильные номера и подписи. Исходные SVG и высококачественные PNG входят в графический пакет версии 1.1.

# KLT-MATDES

## Проектно-техническая спецификация · v1.1

© Курпишев Иван Борисович, 2026. Авторская концепция и компоновка проекта.