

According to a document from **2025-12-18**, TRQuant 已经具备：Cursor 扩展前端 + Python Bridge + MCP 六服务器 + 核心模块 + MongoDB/Redis 的完整骨架，并以 **9 步投资工作流** 串联“数据→市场→主线→候选池→因子→策略→回测→优化→报告”。

你要的“十倍股早期识别体系”应当以 **“Strategy Pack (策略包)”** 形式落地：它是平台可插拔的一种策略族群；平台内核要支撑未来新增其它策略包（红利、择时、事件套利、CTA 等），而不是把十倍股逻辑写死在核心模块里。

下面给出一份**详细、可行、精确**的“开发总方案”（Master Plan）。后续你要细化哪个环节，我们再逐段拆解成接口/表结构/任务分解/验收用例。

1. 现状基线（必须承认并复用的既有资产）

1.1 架构与交互形态

- 前端：Cursor IDE 扩展（多面板），通过 Python Bridge 调用 MCP 工具。
- 服务：6 个 MCP 服务器（core / workflow / project / trading / dev / filesystem）。
- 核心能力：trquant-core 已聚合数据、市场、因子、策略模板、回测、优化、报告、缓存、性能统计等工具域。

1.2 工作流与模块

- 9 步工作流已定义步骤、工具映射与状态机（pending/running/completed/failed）。
- 核心模块层已具备：
data/factors/backtest/templates/optimization/reporting/plugin/trading/qmt/bullettrade 等结构化模块分层。

1.3 存储与缓存

- MongoDB：jqquant 库与 trquant 库分别管理候选池、市场缓存、因子值、工作流日志、策略库、回测结果等集合；Redis 用作 L2 缓存，内存为 L1，且已有 TTL 策略。

1.4 已识别的下一步问题（文档已明确）

待打磨聚焦四点：步骤内优化、步骤间互联、用户体验、一键化与生产化监控告警。
这四点不是“锦上添花”，是你要做十倍股早识别+多策略平台的前置条件。

2. 目标产品定义（必须把边界写清楚）

2.1 平台级目标 (TRQuant Core 2.x)

1. **多策略容器**: 策略以 Strategy Pack 插件化装载；包之间隔离依赖、共享平台能力（数据/因子/回测/报告）。
2. **可复现研究**: 数据版本、因子版本、策略版本、回测版本四位一体可追溯（Reproducibility）。
3. **端到端闭环**: 9 步工作流变成可编排的 DAG，支持“增量刷新”（只重跑受影响步骤）。
4. **生产化标准**: 监控、日志、告警、权限与密钥管理、稳定性基线。

2.2 十倍股早期识别 Strategy Pack (Tenbagger Pack)

- 以“产业位置→兑现路径→财务拐点→组织信号→估值错配”的证据链为核心；输出：
 - 分层候选池 (U1/U2/U3)
 - 证据链状态机 (Stage S0–S5)
 - 评分卡 (7 维或可扩展维度)
 - 论点对象 (Thesis) 与证伪条件 (Falsification)
 - 专属评估指标: Recall@K、Time-to-detection、误杀率等 (不是只看收益曲线)

2.3 非目标 (明确不做/后做)

- 不从华泰专业版 III 客户端直连抓取 (合规风险)。
- 不把十倍股“选股”当成纯量价预测问题：主信号来自文本/事件/产业链证据链，K 线只是执行与风控叠加层。
- 不在第一版强依赖 L2 (可作为后续增强数据源插件)。

3. 总体技术方案 (Target Architecture)

原则：在不推倒现有 MCP + 9 步工作流骨架的前提下，新增“策略包层 + 证据链对象层 + 实验复现层”，并把“步骤互联”做成一等公民。

3.1 统一数据与对象模型（平台内核新增 Domain Layer）

新增平台级对象（逻辑模型）：

- SecurityMaster：证券主数据（代码/简称/交易所/变更映射）
- DataSnapshot：数据快照（数据源/时间/范围/校验 hash）
- RawDoc：原始文本（公告/年报/互动易/新闻/研报摘要/环评/招投标/招聘）
- Event：结构化事件（送样/验证/量产/扩产/高管变更/股权激励/认证等）
- Stage：状态机（S0-S5）+置信度+证伪条件
- FactorDef / FactorValue：因子定义/因子值（你们已有 factor_values，但需要补版本与口径元数据）
- ScoreCard：评分卡（维度贡献可解释）
- Thesis：论点对象（假设-证据-证伪-更新日志）
- Experiment：实验记录（策略配置+数据快照+指标+产物链接）
- StrategyPack：策略包注册表（manifest + capabilities）
- WorkflowRun：工作流运行（DAG + step outputs）

落库建议（MongoDB / trquant 库新增集合）

- raw_docs, events, stages, theses, scorecards, experiments, strategy_packs, data_snapshots
并为 security_id + timestamp、event_type + date、pack_name + version 建复合索引。

现有 MongoDB 结构已覆盖 workflow_logs/states/reports、strategies、backtest_results 等，可直接扩展集合而不破坏既有结构。

3.2 Adapter 抽象：数据源/执行端/回测端三类适配器

TRQuant 已有 unified_data_provider 与多提供者（JQData/AKShare/Mock）。

下一步要把“数据源差异”完全隔离：

- MarketDataAdapter: 行情/估值/财务/指标
- AltDataAdapter: 文本、舆情、研报、项目、招聘、招投标等（可按授权接入）
- ExecutionAdapter: QMT/BulletTrade/券商（你们已有 qmt/bullettrade/broker/trading 模块，统一接口即可）。

关键点：策略包只能调用 Adapter Interface，不能直接调用某个 provider SDK。

3.3 9 步工作流从“线性面板”升级为“可编排 DAG”

文档已指出“步骤间互联”是待打磨重点。

需要新增：

- WorkflowContext (强类型对象)：贯穿步骤的输入输出载体
- StepArtifact: 每一步产物 (JSON + 指向 Mongo 的 artifact_id)
- DependencyGraph: 定义 Step 依赖与增量重跑规则 (例如因子变了，只重跑回测/优化/报告)

3.4 Strategy Pack 插件体系 (平台核心增强点)

TRQuant 已有 core/plugin (插件系统) 与 core/templates (策略模板)，但需要**产品级规范**。

建议定义 StrategyPack Manifest (JSON/YAML)：

- pack metadata: name/version/author/description
- required adapters: market/altdata/execution (可选)
- workflow hooks: 覆盖或扩展 9 步中的某些 step (例如 Tenbagger 会重写 Step3 主线、Step4 候选池、Step9 报告模板)
- metrics suite: 回测评价指标集
- report templates: 报告模板列表

并提供 pack.run(workflow_context) MCP 工具：一键执行“该策略包的推荐路径”。

3.5 观测性与生产化

TRQuant 已有监控面板、缓存统计与 perf 工具域。

下一步要求“生产化”：要把 日志、指标、告警、限流做成标准件。

4. 分阶段交付路线 (**Milestones + DoD**)

不给工期，只给依赖顺序与可验收标准 (Definition of Done)。

M1：平台底座加固（解决“步骤互联+复现”）

交付物

1. WorkflowContext + StepArtifact (每步产物落库、可追溯)
2. DataSnapshot (数据版本化：每次数据拉取产生 snapshot_id)
3. Experiment (回测/优化的运行登记：输入、参数、数据版本、产物)
4. 工作流 DAG：支持“仅重跑受影响步骤”
5. 统一错误模型与日志规范 (step fail 可定位)

验收 (DoD)

- 同一策略同一 snapshot_id 能 100% 复现结果
- 任一步骤产物可在 Mongo 中检索到并回放
- Step 依赖变更触发增量重跑 (不必全跑 9 步)

(直接对齐文档“步骤间互联/生产化”待打磨项。)

M2：Strategy Pack 框架产品化（多策略容器落地）

交付物

1. StrategyPack Manifest 规范 + pack registry (DB + API)
2. pack.* MCP 工具域：install/list/validate/run
3. pack 的沙箱运行与依赖隔离 (至少逻辑隔离：命名空间/数据隔离/版本隔离)
4. 报告模板可由 pack 注入 (Step9)

验收 (DoD)

- 新增一个 pack 不修改 trquant-core 主逻辑，只注册即可运行

- pack 运行的输入/输出与 9 步工作流完全兼容（通过 WorkflowContext）
-

M3: Tenbagger Pack MVP (十倍股早期识别“最小可用闭环”)

交付物 (核心)

1. IndustryGraph (产业链图谱知识库：节点/材料/替代路线/验证周期；支持人工维护入口)
2. RawDocStore (公告/年报/互动易/新闻等抓取入库)
3. EventExtractor (规则优先：送样/验证/量产/扩产/组织变更等事件抽取)
4. StageMachine (S0-S5 + 置信度 + 证伪触发)
5. ScoreCard Engine (7 维评分卡，可解释维度贡献)
6. 输出 U1/U2/U3 分层候选池 (并记录入池/出池原因)
7. Tenbagger 专属报告模板：证据链时间线 + 财务拐点 + 估值错配

验收 (DoD)

- 每周一键刷新候选池：分层名单 + 分数变化 + Stage 变化 + 证伪预警
 - 任意股票进入/退出 U2/U3 必须给出可追溯证据 (Event/Factor/Rule)
 - 事件抽取支持人工纠错回写 (为后续学习做准备)
-

M4: Tenbagger Pack 评估体系 (从“能用”到“可迭代”)

交付物

1. TenbaggerMetricSuite: Recall@K、Time-to-detection、误杀率、Thesis-survival 等
2. 事件时间戳一致性框架 (避免前视偏差：公告发布时间 vs 可交易时间)
3. 因子/事件权重可配置与版本化 (ScoreCard v1/v2)

验收 (DoD)

- 可以对历史窗口跑“早识别效果评估”(不是只看收益)

- 能对比两个版本评分卡/抽取器的改进幅度 (A/B)
-

M5：平台扩展到“多策略组合与执行”（可选增强，按你们商业目标排期）

TRQuant 已具备 trading/broker/qmt/bullettrade 模块基础。

在不改变你们“策略生成中枢”定位的情况下，建议把执行当成 可插拔输出端：

交付物

- SignalSchema (标准信号/意图: entry/exit/size/risk tags)
- PortfolioAllocator (多策略资金分配 + 风险预算)
- PaperTrading (仿真/回放)
- ExecutionAdapter 对接 (QMT/PTrade/BulletTrade)

验收 (DoD)

- 任意策略包能输出统一 signal，被组合层消费
 - 执行端失败不影响研究闭环 (隔离)
-

5. MCP 工具域的扩展规划 (与现有 core 工具一致风格)

现有 trquant-core 已按

data_source/market/factor/template/backtest/optimizer/report/cache/perf 分域。

新增建议 (不破坏旧工具，增量扩展)：

- snapshot.*: create/get/compare (数据快照)
- doc.*: ingest/search/dedup/stats (RawDoc)
- event.*: extract/list/validate/feedback (事件抽取与纠错)
- stage.*: compute/get/override/history (状态机)
- thesis.*: upsert/link_evidence/falsify/timeline (论点对象)
- scorecard.*: compute/explain/versioning (评分卡)
- pack.*: install/list/validate/run (策略包)

- experiment.*: register/get/compare/export (实验对比)
 - portfolio.* (可选): allocate/risk/report
-

6. 质量体系（测试、稳定性、变更管理）

TRQuant 已有 15 个测试文件与 MCP 工具测试框架。

在上述里程碑推进时，测试体系必须升级为三层：

1. **Contract Tests** (接口契约)：每个 MCP tool 输入输出 schema 固定 (破坏即失败)
2. **Golden Dataset Tests** (黄金数据集)：选一小段固定市场窗口与固定数据快照做回归测试
3. **E2E Workflow Tests** (端到端)：9 步+某策略包一键跑通 (你们已有端到端测试基础，但要覆盖 pack.run)

同时，变更记录与里程碑管理可以复用你们 MCP 项目管理工具 (task/milestone/change 等)。

7. 风险与对策（工程上必须提前写进方案）

1. **数据授权与合规**：AltData (研报/新闻/互动易/招投标/招聘) 接入必须以授权/公开可用为前提；适配器层需携带 license_tag 与使用范围约束。
 2. **事件抽取质量**：V1 规则引擎先跑通闭环；V2 才上模型/LLM；必须有人工反馈回写机制 (否则无法迭代)。
 3. **前视偏差**：事件时间戳规则必须在 M4 前做完，否则所有“早识别评估”不可信。
 4. **系统复杂度膨胀**：靠 Strategy Pack 隔离复杂度；平台内核只做通用能力。
-

8. 你接下来要我先细化哪一块（我建议从这两块之一开刀）

为了让后续细化“最省力且不返工”，我建议你在下面二选一作为下一轮细化入口（你选哪块，我就把接口/DB 表/任务拆分/验收用例写出来）：

1. **M1: WorkflowContext + DataSnapshot + Experiment** (平台复现与步骤互联)
——这是所有策略的地基。
2. **M3: Tenbagger Pack MVP (RawDoc→Event→Stage→ScoreCard→候选池→报告)**
——这是十倍股闭环最核心路径。

(你回复“先细化 M1”或“先细化 M3”，我就按对应里程碑输出“工程级细化方案：模块拆分、接口、数据表、工具清单、DoD、测试用例”。)

引用：TRQuant 系统总结文档（生成时间 2025-12-18）。