

WorkSpace·Order Operations Agent

WorkSpace · 订单运营 AI Agent

AI Agent Solution Design/ AI Agent 解决方案设计

文档名称 Document Title	AI Agent 解决方案设计 / AI Agent Solution Design
版本 Version	V1.0
文档状态 Status	作品集/面试版本/Portfolio / Interview Version
适用场景 Scope	企业订单运营 / Enterprise Order Operations
项目性质 Project Type	模拟业务案例 / Simulated business case
发布日期 Date	2026-09-26

01 方案概述

01 Solution Overview

本方案面向企业订单运营场景，构建由多个专业 Agent 协同工作的企业级 AI Agent 系统。系统以订单为核心业务对象，由 Order Agent / 订单 Agent 作为主要任务入口，并根据业务任务协调 PMC、Purchase、Warehouse、Sample、Finance 等专业 Agent。

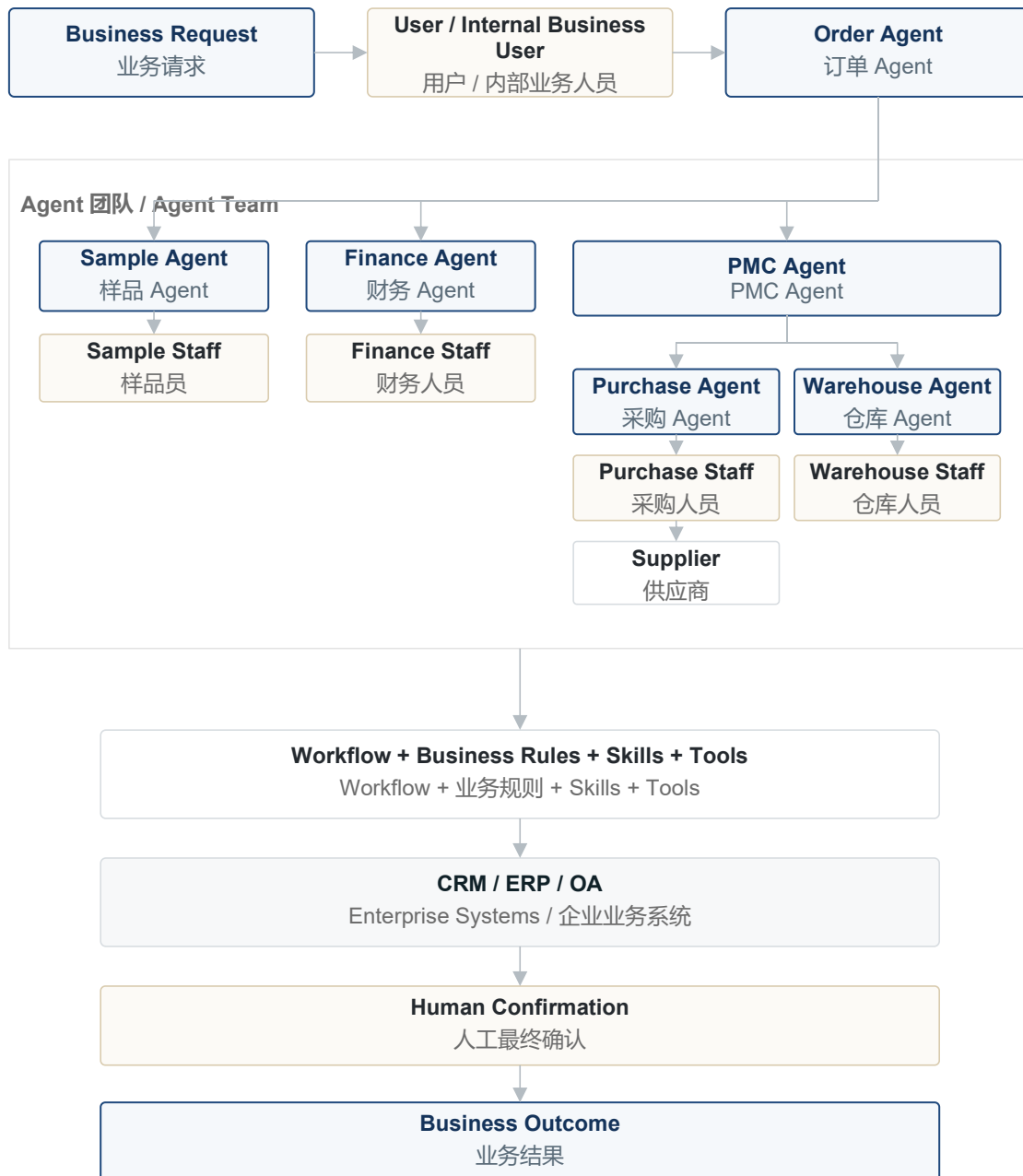
This solution is an enterprise AI Agent system for order operations. The system is centered on the order as the core business object, with the Order Agent serving as the primary task entry point and coordinating specialized Agents such as PMC, Purchase, Warehouse, Sample, and Finance.

系统不是通过一个“大而全”的 Agent 完成所有业务，而是将业务职责、Workflow、Business Rules、Skills、Tools 和 Human Control 分层设计，使 Agent 能够在受控范围内完成信息获取、业务判断、跨角色协同和系统操作。

The system does not rely on a single general-purpose Agent to perform all business tasks. Business responsibilities, Workflows, Business Rules, Skills, Tools, and Human Control are separated so that Agents can retrieve information, make business decisions, coordinate across roles, and execute system operations within controlled boundaries.

核心运行链路如下：

The core operating model is:



02 受众与设计目标

02 Audience & Design Objectives

2.1 目标读者

2.1 Intended Audience

角色 / Role	主要关注点 / Primary Concern	本文提供的信息 / Information Provided
AI 产品经理 / AI Product Manager	业务如何映射到 Agent 能力 / How business requirements map to Agent capabilities	Business Scenario → Agent Capability → Workflow → System / 业务场景 → Agent 能力 → Workflow → 系统
解决方案架构师 / Solution Architect	Agent、系统和 Tool 如何协同 / How Agents, systems, and Tools collaborate	Runtime Architecture、Agent Boundary、Tool Gateway、State / 运行时架构、Agent 边界、Tool Gateway、状态管理
AI / Backend Engineer / AI / 后端工程师	Agent 如何真正执行任务 / How Agents execute real tasks	Workflow、Tool Contract、Agent Contract、System Integration / Workflow、Tool Contract、Agent Contract、系统集成
业务负责人 / Business Owner	Agent 如何帮助业务完成工作 / How Agents support business operations	Business Operating Model、Human Collaboration、Business Outcome / 业务运行模式、人工协同、业务结果
QA / Evaluation / QA 与评估人员	如何验证 Agent 是否正确 / How Agent behavior is evaluated	Evaluation Dataset、Business Rules、Failure Cases / 评估数据集、业务规则、异常场景
Security / IT / 安全与 IT	Agent 如何被控制和审计 / How Agents are governed and audited	Permission、Audit、Human Control、Governance / 权限、审计、人工控制、治理

2.2 设计目标

2.2 Design Objectives

目标一：业务职责清晰

Objective 1: Clear Business Ownership

每个 Agent 对应明确的业务领域和内部协同人员，不跨越未经授权的业务边界。

Each Agent owns a clearly defined business domain and works with designated internal users without crossing unauthorized business boundaries.

目标二：Agent 能够执行真实业务任务

Objective 2: Agents Can Execute Real Business Tasks

Agent 不只负责回答问题，而是能够在 Workflow 驱动下读取业务信息、调用专业 Agent、调用企业系统 Tool，并完成受控的业务操作。

Agents do not only answer questions. Through Workflows, they can retrieve business information, collaborate with specialized Agents, invoke enterprise-system Tools, and execute controlled business operations.

目标三：业务事实与模型推理分离

Objective 3: Separate Business Facts from Model Reasoning

实时订单状态、交期、物料状态、审批状态等业务事实来自企业系统，而不是由 LLM 自行生成。

Real-time order status, delivery dates, material status, approval status, and other business facts come from enterprise systems rather than being generated by the LLM.

目标四：自动化与人工控制并存

Objective 4: Combine Automation with Human Control

Agent 可以完成信息准备、查询、校验、业务协同和允许范围内的系统操作，但所有最终提交/确认动作保留人工控制。

Agents can perform information preparation, retrieval, validation, business coordination, and permitted system operations, while all final submission and confirmation actions remain under human control.

03 产品与系统边界

03 Product & System Boundary

3.1 业务范围

3.1 Business Scope

系统覆盖订单运营中的核心业务场景，并通过不同阶段逐步扩展自动化深度。

The system covers core order-operation scenarios, with automation depth expanding progressively across product phases.

当前业务范围包括：

The current business scope includes:

- 订单查询与订单识别
Order retrieval and identification
- Customer PO 与内部订单对象匹配
Customer PO to internal order-object matching
- 订单创建与订单变更
Order creation and modification
- 交期确认与交期复查
Delivery-date confirmation and revalidation
- 样品协同
Sample coordination
- 物料状态协同
Material-status coordination
- 采购协同
Purchase coordination
- 出货准备
Shipment preparation
- 财务审批协同
Finance approval coordination
- CRM 信息维护
CRM information maintenance
- OA 审批协同
OA approval coordination
- 跨部门任务协调与结果回传
Cross-functional task coordination and result reporting
- 异常识别、升级与人工介入
Exception detection, escalation, and human intervention

The Order Agent serves as the unified business entry point, while specialized Agents provide domain-specific capabilities for PMC, Finance, Sample, Purchase, and Warehouse operations.

Order Agent 作为统一业务入口，由 PMC、Finance、Sample、Purchase 和 Warehouse 等专业 Agent 提供领域协同能力。

All Agents operate within internal enterprise workflows. They do not directly communicate with customers, suppliers, or freight forwarders.

所有 Agent 均运行于企业内部业务流程中，不直接与客户、供应商或货代进行外部沟通。

Final business confirmation and high-risk decisions remain under human control.

最终业务确认及高风险业务决策仍由人工负责。

3.2 产品能力演进

3.2 Product Capability Evolution

The product does not attempt to automate the entire order-operation lifecycle from the beginning. Capability expansion follows a staged approach from business-value validation to end-to-end automation and then proactive operations.

产品并非从一开始就尝试自动化整个订单运营生命周期，而是按照“业务价值验证 → 端到端自动化验证 → 规模化智能”的路径逐步扩展。

P0 — 验证业务价值

P0 — Validate Business Value

P0 focuses on validating whether the Agent can reduce repetitive order-operation work and improve basic follow-up efficiency.

P0 重点验证 Agent 是否能够减少重复性的订单运营工作，并提升基础订单跟进效率。

Core capabilities include:

核心能力包括：

- 订单查询与订单识别
Order retrieval and identification
- Customer PO 与内部订单对象匹配
Customer PO to internal order-object matching
- 基础订单信息维护
Basic order information maintenance
- 基础订单创建与订单变更协同
Basic order creation and modification coordination
- CRM 信息查询与维护
CRM information retrieval and maintenance

- 基础交期查询与跟进
Basic delivery-date inquiry and follow-up
- 基础跨部门任务协调
Basic cross-functional task coordination

The primary objective is to validate business value through measurable reductions in repetitive work and follow-up effort.

核心目标是通过重复工作量和订单跟进投入的可量化下降，验证业务价值。

P1 — Validate End-to-End Automation

P1 extends the Agent from basic information retrieval and follow-up toward coordinated execution across business workflows and enterprise systems.

P1 将 Agent 从基础信息查询和跟进进一步扩展到跨业务流程和企业系统的协同执行。

Core capabilities include:

核心能力包括：

- PMC 计划协同
PMC planning coordination
- 物料状态协同
Material-status coordination
- 采购协同
Purchase coordination
- 样品协同
Sample coordination
- 财务审批协同
Finance approval coordination
- 出货准备与 Warehouse 协同
Shipment preparation and Warehouse coordination
- OA / ERP / CRM 跨系统操作
Cross-system operations across OA, ERP, and CRM
- 异常识别与升级
Exception detection and escalation
- 跨 Agent 协同
Multi-Agent coordination
- 端到端 Workflow 执行与人工确认
End-to-end Workflow execution with human confirmation

P1 focuses on validating whether the Agent can coordinate a complete business process across people, specialized Agents, and enterprise systems while preserving

human control at critical decision points.

P1 重点验证 Agent 能否在保留关键节点人工控制的前提下，协同人员、专业 Agent 与企业系统完成完整业务流程。

P2 — 规模化智能

P2 — Scale Intelligence

P2 extends the system from reactive order execution toward proactive operational intelligence.

P2 将系统从被动订单执行进一步扩展到主动运营智能。

Core capabilities include:

核心能力包括：

- 客户智能
Customer intelligence
- 客户订单行为与历史模式分析
Customer order behavior and historical pattern analysis
- 主动风险识别
Proactive risk detection
- 主动任务触发与提醒
Proactive task triggering and reminders
- 交付风险预测与提前干预
Delivery-risk prediction and early intervention
- 基于历史反馈的规则优化建议
Feedback-driven rule optimization recommendations
- 更高层次的受控自动化
Higher levels of controlled automation

P2 does not imply fully autonomous business decision-making. Automation depth continues to depend on business risk, reversibility, accuracy, human acceptance, and governance requirements.

P2 并不意味着完全自主的业务决策。自动化程度仍取决于业务风险、可逆性、准确性、人工接受度及治理要求。

3.2 系统职责边界

3.2 Enterprise System Responsibility Boundary

系统 / System	主要职责 / Primary Responsibility
CRM / CRM	订单状态查询与订单业务信息记录；维护客户要求交期、确认交期等订单运营信息 / Order-status query and order business information; maintains customer-requested and confirmed delivery dates
ERP / ERP	订单创建、订单变更、ERP 业务校验、出货单创建，以及必要的执行状态查询 / Order creation, order modification, ERP validation, shipment-document creation, and required execution-status queries
OA / OA	业务审批 / Business approval
LLM / LLM	语言理解、任务规划、信息归纳和受控推理 / Language understanding, task planning, information synthesis, and controlled reasoning
Agent Runtime / Agent 运行时	Agent 调度、Workflow 执行、上下文和任务状态管理、Tool 调用 / Agent orchestration, Workflow execution, context and task-state management, and Tool invocation

04 Agent Team Architecture

04 Agent Team Architecture / Agent 团队架构

六个 Agent 使用统一的：

The six Agents follow a unified:

Agent → Internal User → External Party

Agent → 内部协同人员 → 外部参与方

协作模型。

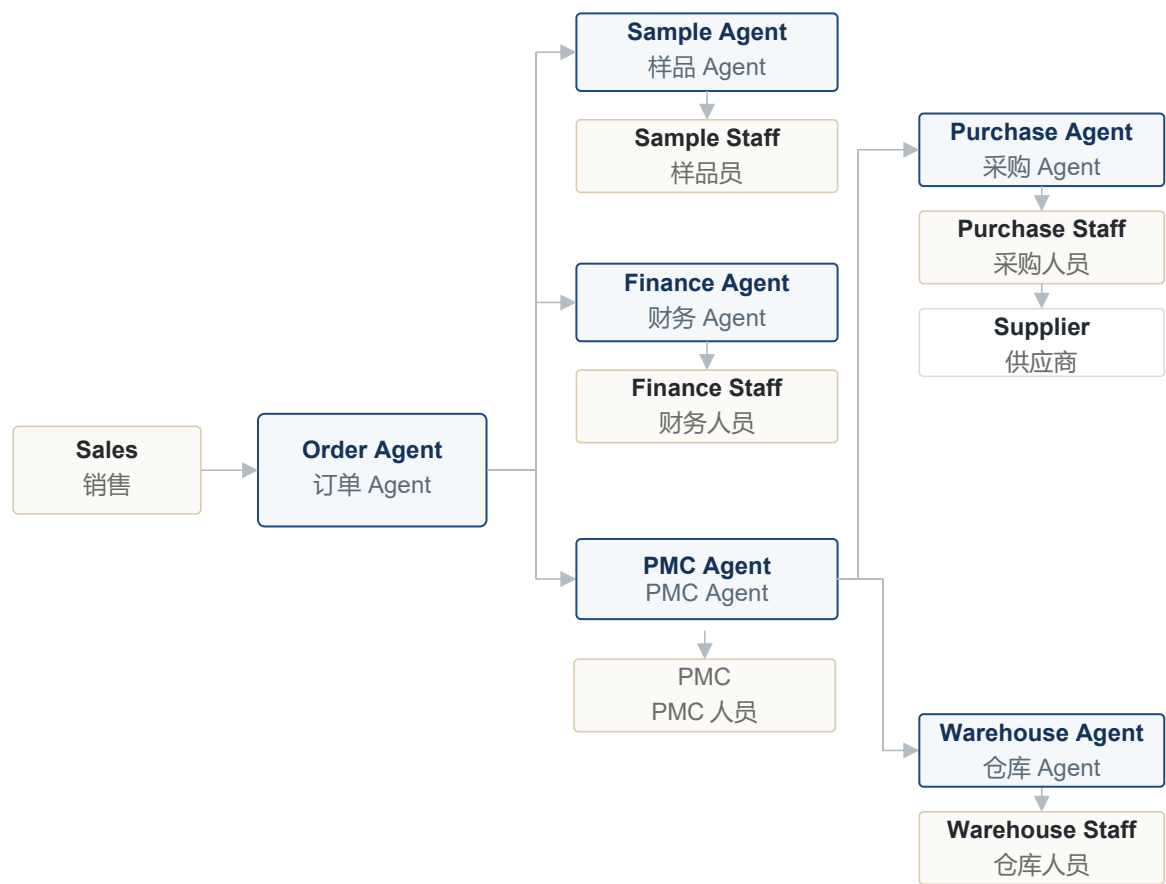
collaboration model.

Agent / Agent	业务领域 / Business Domain	内部协同人员 / Internal User	外部参与方 / External Party
订单 Agent / Order	订单运营 / Order	销售 / Sales	—

Agent / Agent	业务领域 / Business Domain	内部协同人员 / Internal User	外部参与方 / External Party
Agent	Operations		
PMC Agent / PMC Agent	计划与交付协调 / Planning & Delivery Coordination	PMC 人员 / PMC Staff	—
采购 Agent / Purchase Agent	采购协同 / Purchase Coordination	采购人员 / Purchase Staff	供应商 / Supplier
仓库 Agent / Warehouse Agent	仓储与出货协同 / Warehouse & Shipment Coordination	仓库人员 / Warehouse Staff	—
样品 Agent / Sample Agent	样品协同 / Sample Coordination	样品员 / Sample Staff	—
财务 Agent / Finance Agent	财务与审批协同 / Finance & Approval Coordination	财务人员 / Finance Staff	—

Agent 协作关系

Agent Collaboration



Purchase Agent 的业务边界明确为：Purchase Agent 只与采购人员协同，不直接联系供应商。采购人员负责实际供应商沟通，并将供应商信息反馈给 Purchase Agent。

The Purchase Agent has an explicit business boundary: it collaborates with Purchase Staff and does not directly contact suppliers. Purchase Staff is responsible for actual supplier communication and feeds supplier information back to the Purchase Agent.

05 Business Operating Model

05 业务运行模式

企业业务不是简单的“User → Agent → System”，而是由 Agent 作为业务协同层连接业务人员、专业角色和企业系统。

Enterprise operations are not simply “User → Agent → System.” The Agent acts as a coordination layer connecting business users, specialized roles, and enterprise systems.



这种模式确保 Agent 不被设计为“替代所有业务角色”，而是承担任务理解、信息整合、流程推进和系统操作等能力，同时保留业务人员的专业判断和最终控制权。

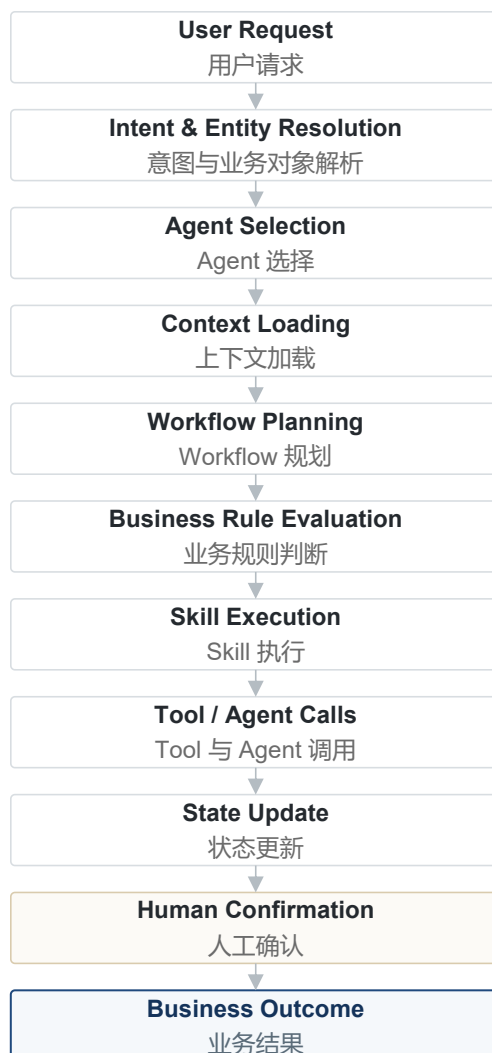
This model ensures that Agents are not designed to replace all business roles. Instead, they provide task understanding, information synthesis, workflow progression, and system operation capabilities while preserving human expertise and final control.

06 Agent Runtime Architecture

06 Agent Runtime Architecture / Agent 运行时架构

Agent Runtime 负责将用户请求转换为可执行任务，并协调 Prompt、Workflow、Business Rules、Skill、Tool 和 Human Control。

The Agent Runtime converts user requests into executable tasks and coordinates Prompts, Workflows, Business Rules, Skills, Tools, and Human Control.



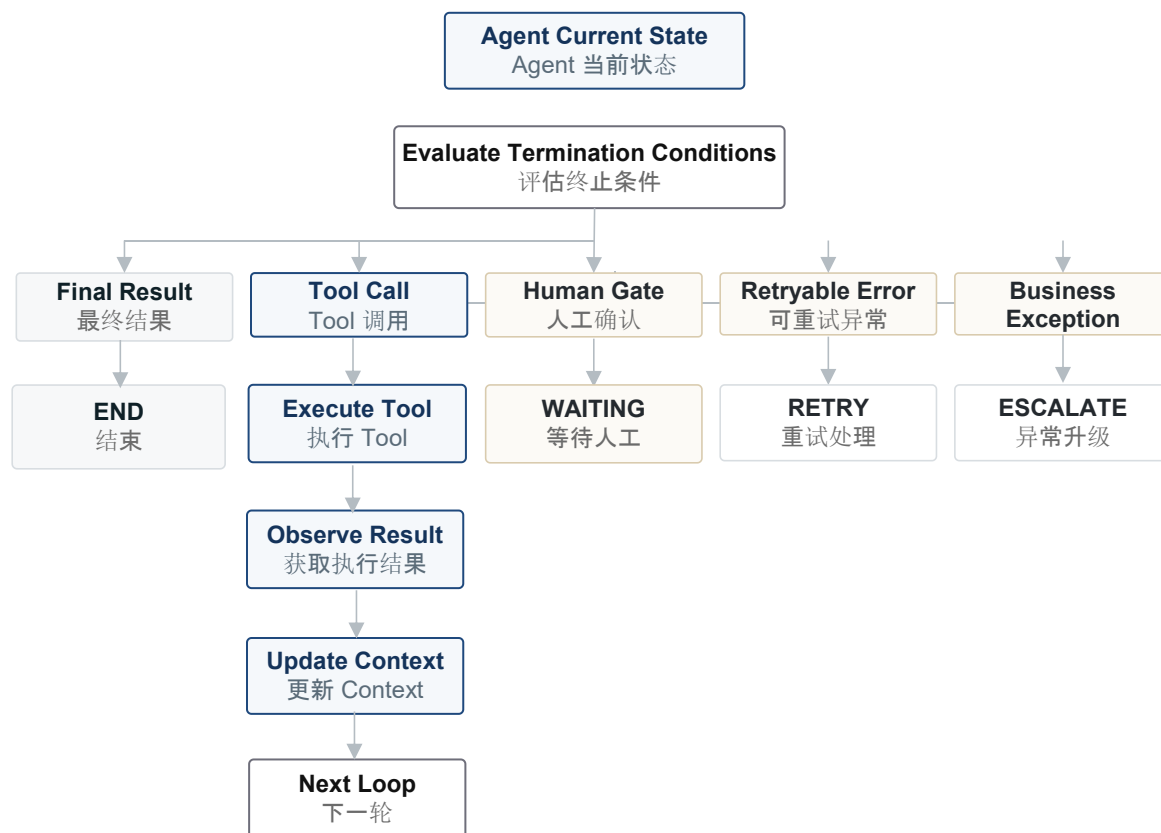
6.1 Agent Loop Guard / Agent 循环控制

执行必须具备明确的运行时终止条件，以防止无限循环、过度 Tool 调用和不必要的模型消耗。Agent execution must have explicit runtime termination conditions to prevent uncontrolled loops, excessive tool calls, and unnecessary model usage.

Agent

在每个执行步骤后评估当前 Agent 状态，并决定继续执行、等待、重试、完成或升级处理。The Runtime evaluates the current Agent state after each execution step and determines whether to continue, wait, retry, complete, or escalate.

Runtime



Runtime Guards include:

运行时保护机制包括：

Max Iterations / 最大循环次数

Execution Timeout / 执行超时

Tool Call Limit / Tool 调用次数限制

Token & Cost Budget / Token 与成本预算

Cancellation / 任务取消

Termination Conditions / 终止条件

When a runtime guard is triggered, the Agent must stop autonomous execution and enter a controlled exception or human-escalation state.

当运行时保护条件被触发时，Agent 必须停止自主执行，并进入受控异常状态或人工升级状态。

The LLM determines the next action within the available capabilities, while the Runtime remains responsible for execution control, safety limits, and termination.

LLM 在可用能力范围内决定下一步动作，而 Runtime 负责执行控制、安全限制与任务终止。

6.2 Production Runtime Considerations / 生产环境运行时考虑

当 Agent 从原型阶段进入生产环境并承担大规模任务后，Runtime 需要从简单任务执行模式演进为具备持久化、可观测性和并发控制能力的执行模式。

As Agent workloads scale from prototype usage to production workloads, the Runtime must evolve from simple task execution toward a persistent, observable, and concurrency-controlled execution model.

Key production considerations include:

生产环境需要重点考虑：

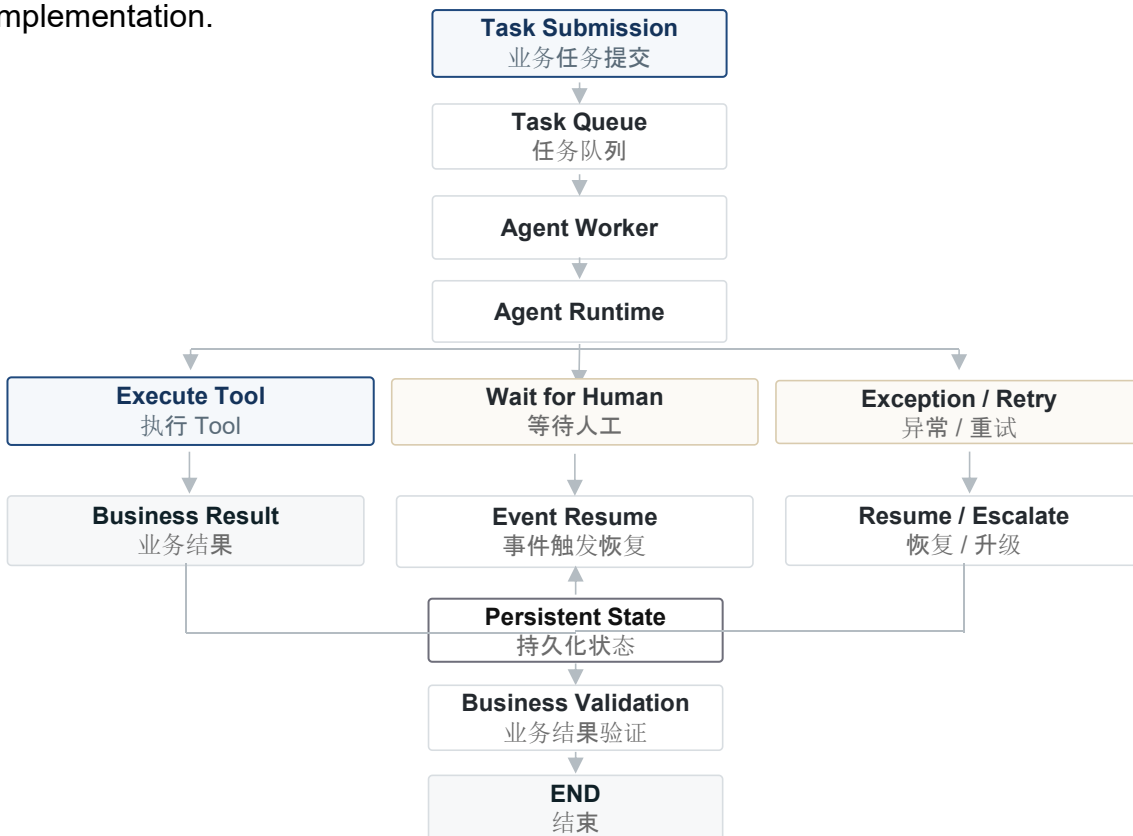
Area / 领域	Production Consideration / 生产环境考虑
Concurrency Control / 并发控制	Limit concurrent Agent and Tool execution to protect enterprise systems and model services. / 控制 Agent 与 Tool 的并发执行，避免企业系统和模型服务过载。
Task Queue & Workers / 任务队列与 Worker	Decouple task submission from execution and support asynchronous, long-running workflows. / 将任务提交与执行解耦，支持异步及长流程 Workflow。
Persistent State / 状态持久化	Persist execution state so tasks can survive service restarts or worker failures. / 持久化执行状态，使任务能够跨越服务重启或 Worker 故障继续执行。
Idempotency & Deduplication / 幂等与去重	Prevent duplicate business operations when Tool calls time out or are retried. / 防止 Tool 超时或重试导致重复业务操作。
Checkpoint & Resume / 检查点与恢复	Resume long-running tasks from a validated checkpoint instead of replaying completed actions. / 从经过验证的检查点恢复长流程任务，避免重复执行已经完成的操作。
Observability / 可观测性	Track Agent state, Tool calls, failures, latency, cost, human intervention and business outcomes. / 追踪

Area / 领域	Production Consideration / 生产环境考虑
	Agent 状态、Tool 调用、异常、延迟、成本、人工介入和业务结果。
Rate & Cost Control / 速率与成本控制	Control model calls, Tool calls and execution budgets under high workload. / 在高负载情况下控制模型调用、Tool 调用以及执行预算。

生产环境 Runtime 还需要支持可能等待数小时甚至数天的长流程任务，例如审批或人工确认。Production Runtime should also support long-running tasks that may remain in a waiting state for hours or days, such as approval or human confirmation.

处于等待状态的任务应释放执行资源，并仅在收到所需事件或人工操作后恢复执行。Waiting tasks should release execution resources and resume only when the required event or human action is received.

对于当前作品集范围，这些能力作为 Runtime 设计考虑进行定义，而不展开为完整的分布式基础设施实施方案。For the current portfolio scope, these capabilities are treated as Runtime design considerations rather than a detailed distributed infrastructure implementation.



7.1 核心业务对象

7.1 Core Business Objects

系统以订单及订单相关业务对象作为 Agent 的核心上下文。

The system uses the order and its related business objects as the primary Agent context.

业务对象 / Business Object	说明 / Description
Customer PO / 客户 PO	用户最常使用的订单识别信息 / Primary user-facing order identifier
ERP Internal Order Number / ERP 内部订单号	ERP 生成的内部订单识别信息，用户也可以直接提供 / ERP-generated internal order identifier that users may also provide
Internal Order Object / 内部订单对象	Agent 通过企业系统解析得到的实际业务对象 / Actual business object resolved through enterprise systems
Customer Requested Delivery Date / 客户要求交期	CRM 中记录的客户要求交付日期 / Customer-requested delivery date recorded in CRM
Confirmed Delivery Date / 确认交期	PMC 确认后写入 CRM 的业务交期 / Business delivery date confirmed by PMC and recorded in CRM
Order Status / 订单状态	CRM 中维护的订单业务状态 / Order business status maintained in CRM
Material Status / 物料状态	ERP 中与执行相关的物料信息 / ERP material information related to execution
Shipment Status / 出货状态	订单执行过程中需要查询的出货信息 / Shipment execution information queried when required
Approval Status / 审批状态	OA / Finance 流程中的审批信息 / Approval information from OA / Finance workflows

7.2 Context 获取原则

7.2 Context Loading Principles

用户可以通过 Customer PO 或 ERP Internal Order Number 识别订单。

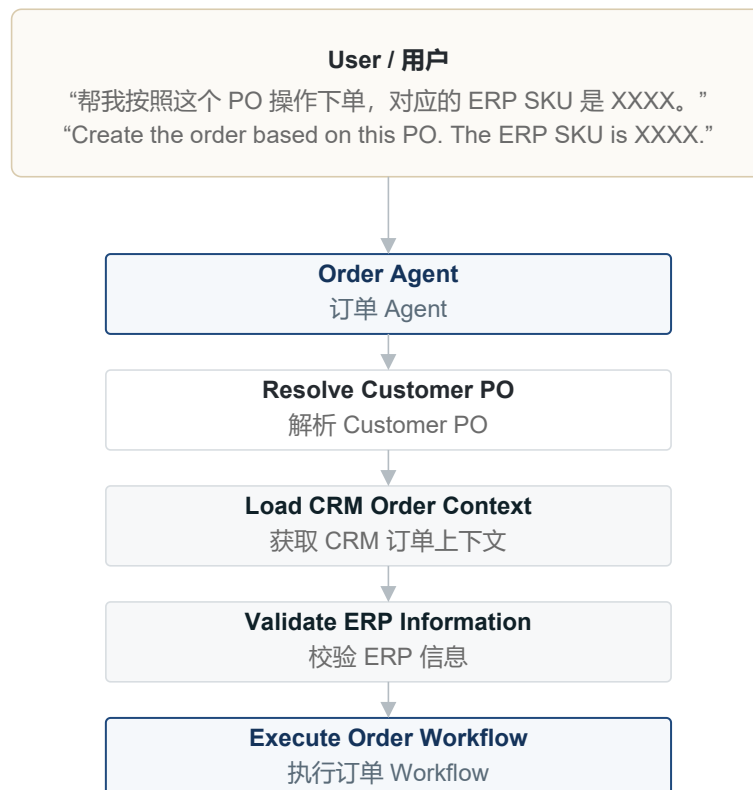
Users can identify an order using either a Customer PO or an ERP Internal Order Number.

Agent 不要求用户提供内部 Business ID，而是负责通过企业系统完成业务对象解析。

The Agent does not require users to provide an internal Business ID. It resolves the relevant business object through enterprise systems.

例如：

For example:



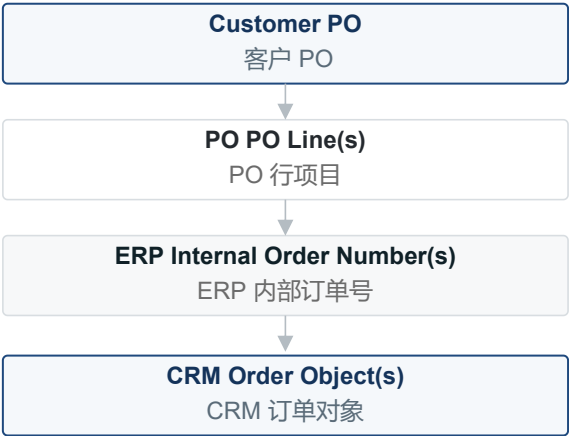
如果用户已经提供 ERP SKU，则 Agent 使用该信息并通过 ERP 进行校验；如果用户没有提供，则在需要 ERP SKU 的阶段向用户请求补充。

If the user has already provided an ERP SKU, the Agent uses it and validates it through ERP. If the SKU has not been provided, the Agent requests it when the ERP stage requires it.

7.3 复杂 ID 映射

7.3Complex ID Mapping

企业订单对象并不总是一一对应。Customer PO、PO Line、ERP Internal Order Number 和 CRM Order Object 可能形成一对一、一对多、多对一或多对多关系。Enterprise order objects are not always mapped one-to-one. Customer PO, PO Line, ERP Internal Order Number, and CRM Order Object may form one-to-one, one-to-many, many-to-one, or many-to-many relationships.



7.4 映射解析规则

7.4 Mapping Resolution Rules

情况 Situation	Agent 行为 Agent Behavior
唯一匹配	继续执行 Workflow / Continue the Workflow
多个匹配	不进行猜测，要求用户确认 / Do not guess; request human confirmation
部分匹配	标识未匹配的订单或行项目，并请求澄清 / Identify unmatched orders or line items and request clarification
无匹配	阻断 Workflow，并要求提供有效标识 / Stop the Workflow and request a valid identifier
人工确认映射	保存确认后的映射关系，并继续执行 / Persist the confirmed mapping and continue the Workflow

08 Business Rules & Policy

08 业务规则与策略

Business Rules 作为独立的业务决策层存在， 并与 Prompt、Workflow 分离管理。

Business Rules are implemented as an independent business decision layer and managed separately from Prompts and Workflows.

8.1 Rule Model

8.1 Rule Model / 规则模型

字段 / Field	说明 / Description
Rule ID / 规则 ID	唯一规则标识 / Unique rule identifier
Trigger / 触发条件	规则何时触发 / When the rule is evaluated
Conditions / 判断条件	需要满足的业务条件 / Business conditions
Action / 执行动作	满足条件后的业务动作 / Action when conditions are satisfied
Blocking / 阻断级别	是否阻止 Workflow 继续 / Whether workflow execution is blocked
Human Review / 人工确认	是否需要人工参与 / Whether human review is required
Source / 来源	SOP、业务制度或系统规则 / SOP, business policy, or system rule
Version / 版本	规则版本 / Rule version

8.2 Core Business Rules

8.2 核心业务规则

Rule / 规则	Business Logic / 业务逻辑	Control / 控制
R01 订单唯一识别 / Unique Order Resolution	Customer PO 或 ERP Internal Order Number 必须解析到唯一内部订单对象 / Customer PO or ERP Internal Order Number must resolve to one unique internal order object	多匹配或无法匹配时阻断 / Block on multiple or unresolved matches
R02 ERP 信息校验 / ERP Validation	ERP SKU 等关键 ERP 字段必须通过 ERP 校验 / Key ERP fields such as ERP SKU must pass ERP validation	校验失败时阻断 ERP 操作 / Block ERP operation on validation failure
R03 首次交期确认 /	CRM 初始包含 Customer Requested Delivery Date; 首次查询需要由 PMC Agent 协同确认	未确认前不得将客户要求交期当作确认交

Rule / 规则	Business Logic / 业务逻辑	Control / 控制
Initial Delivery Confirmation	Confirmed Delivery Date / CRM initially contains the Customer Requested Delivery Date; the PMC Agent must confirm the Confirmed Delivery Date for the first inquiry	期返回 / Do not present the requested date as a confirmed date
R04 交期复查 / Delivery Revalidation	后续查询先读取 CRM 当前 Confirmed Delivery Date, 再与 PMC Agent 重新确认是否发生变化 / Later inquiries first read the current Confirmed Delivery Date from CRM and then revalidate it with the PMC Agent	无变化则不重复写 CRM / Do not rewrite CRM when unchanged
R05 物料风险协同 / Material Risk Coordination	PMC Agent 根据生产计划、物料状态和现有异常判断交付风险 / PMC Agent evaluates delivery risk based on production planning, material status, and current exceptions	存在风险时进入协同或人工处理 / Route to coordination or human handling
R06 财务条件 / Finance Condition	出货单可以在付款收到前创建; 后续履约是否继续取决于 Finance 审批或信用条件, 而非简单以“已付款”作为唯一条件 / Shipment documents may be created before payment is received; downstream fulfillment depends on Finance approval or credit conditions rather than payment receipt alone	未获得 Finance 放行时, PMC 不安排后续履约 / PMC does not arrange downstream fulfillment without Finance release
R07 最终确认 / Final Confirmation	所有最终提交/确认动作必须由人工确认 / All final submission and confirmation actions require human confirmation	Agent 不替代最终人工确认 / Agent does not replace final human confirmation
R08 状态来源 / System of Record	订单业务状态优先从 CRM 获取; 执行相关状态根据业务任务从 ERP 获取 / Order business status comes from CRM; execution-related status is queried from ERP when required	避免无意义跨系统查询 / Avoid unnecessary cross-system queries

09 Prompt Architecture

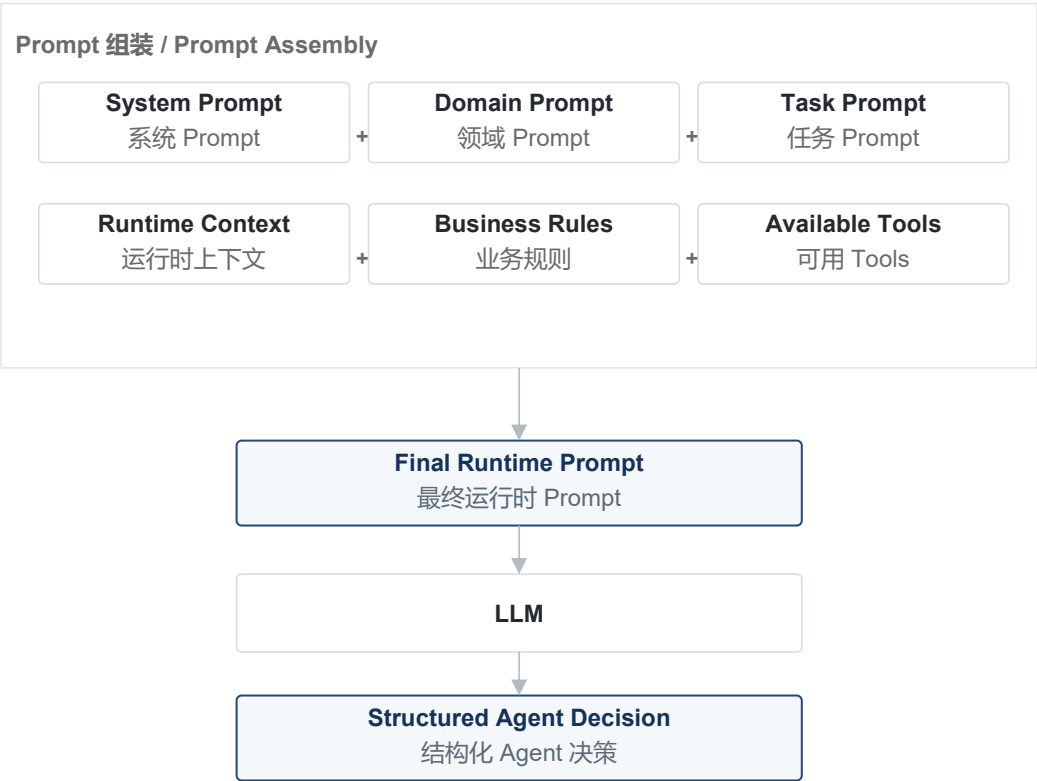
09 Prompt 架构

Prompt 为 Agent 提供角色、任务、行为约束和运行时上下文, 但不承担 Workflow 编排、业务规则执行或系统权限控制。

Prompts provide Agents with role definitions, task instructions, behavioral constraints, and runtime context, while Workflow orchestration, business rule execution, and system authorization remain outside the Prompt.

9.1 Prompt Runtime Assembly

9.1 Prompt 运行时组装



9.2 Prompt Configuration Model

9.2 Prompt 配置模型

Prompt Layer / Prompt 层	作用 / Purpose
System Prompt / 系统 Prompt	定义 Agent 的角色、总体行为边界和输出要求 / Defines the Agent role, global behavior boundaries, and output requirements
Domain Prompt / 领域 Prompt	定义订单、PMC、采购等业务领域知识 / Defines domain-specific business knowledge
Task Prompt / 任务 Prompt	描述当前任务目标、执行要求和约束 / Describes the current task objective, execution requirements, and constraints
Runtime Context / 运行时上下文	注入当前订单、用户输入、历史状态等动态信息 / Injects dynamic order, user input, and state information
Business Rules / 业务规则	提供当前任务适用的结构化业务约束 / Provides structured business constraints applicable to the current task

Prompt Layer / Prompt 层	作用 / Purpose
Tools / 可用 Tools	定义 Agent 当前允许使用的工具 / Defines tools available to the Agent

9.3 Example: Create Order Prompt

9.3 示例：Create Order Prompt / 创建订单 Prompt

Task Prompt 的核心任务内容：

The core Task Prompt is:

1. 根据 PO 唯一解析内部订单对象。
Resolve the PO to a unique internal order object.
2. 获取所需订单上下文。
Load the required order context.
3. 确认创建订单所需的业务字段。
Identify the business fields required for order creation.
4. 如果用户已提供 ERP SKU，则使用该 SKU 并通过 ERP 校验。
If the user has provided an ERP SKU, use it and validate it through ERP.
5. 如果缺少 ERP 创建所需信息，则向用户请求补充。
Request missing information required for ERP creation.
6. 检查当前订单是否存在不允许继续创建的业务条件。
Check whether any business condition prevents order creation.
7. 根据 Workflow 顺序执行 CRM、ERP 及相关 Agent 协同任务。
Execute CRM, ERP, and related Agent coordination tasks according to the Workflow.
8. Tool 执行成功后验证实际业务状态。
Verify the actual business state after Tool execution succeeds.
9. 所有最终提交或确认动作必须进入人工确认节点。
Route all final submission or confirmation actions to a human confirmation node.
10. 返回结构化执行结果、待人工确认事项及异常信息。
Return structured execution results, pending human confirmations, and exception information.

10 Workflow Architecture

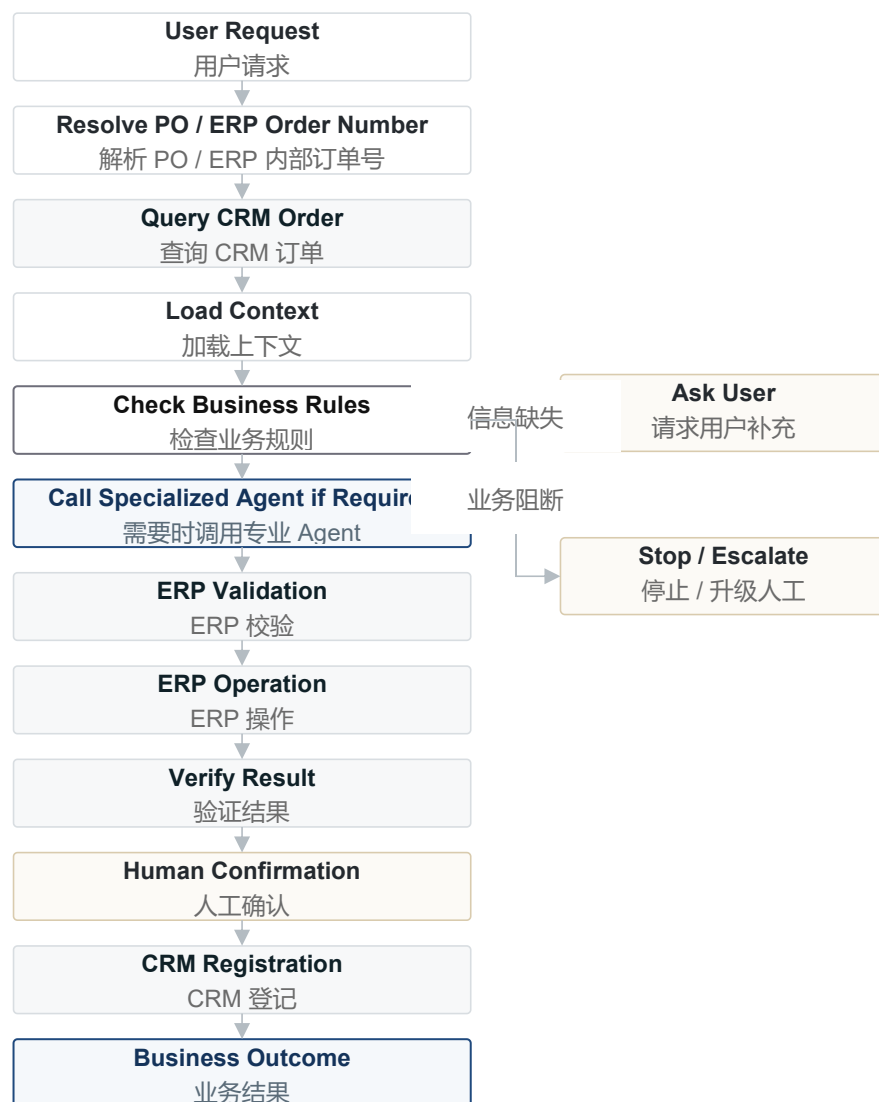
10 Workflow 架构

Workflow 负责定义任务执行顺序、条件分支、Agent 协同、Tool 调用和 Human Gate。

Workflow defines execution order, conditional branches, Agent collaboration, Tool invocation, and Human Gates.

10.1 Standard Order Workflow

10.1 标准订单 Workflow



10.2 Workflow Node Types

10.2 Workflow 节点类型

Node / 节点	作用 / Purpose
Input Node / 输入节点	接收用户请求 / Receives the user request
Resolve Node / 解析节点	解析 PO、ERP 内部订单号等业务对象 / Resolves PO, ERP internal order numbers, and other business objects
Context Node / 上下文节点	获取 CRM、ERP、OA 等业务上下文 / Loads business context from CRM, ERP, and OA
Decision Node / 决策节点	根据 Business Rules 进行条件分支 / Creates branches based on Business Rules
Agent Node / Agent 节点	调用其他专业 Agent / Invokes another specialized Agent
Tool Node / Tool 节点	调用企业系统 Tool / Invokes enterprise-system Tools
Validation Node / 校验节点	验证字段、业务条件或执行结果 / Validates fields, business conditions, or execution results
Human Gate / 人工确认节点	等待人工确认 / Waits for human confirmation
Write Node / 写入节点	将业务结果写回企业系统 / Writes business results back to enterprise systems
Exception Node / 异常节点	处理失败、超时、冲突等异常 / Handles failures, timeouts, conflicts, and other exceptions

11 Decision Layer

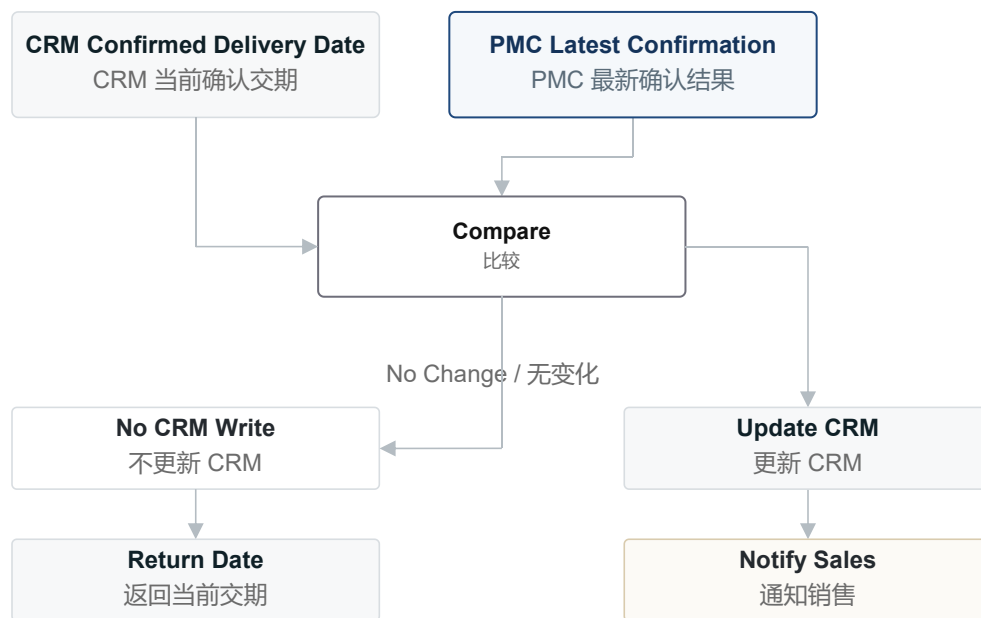
11 Decision Layer / 决策层

Decision Layer 根据 Business Rules、当前业务状态、Tool 返回结果和 Agent 协同结果，决定下一步 Workflow 动作。

The Decision Layer determines the next Workflow action based on Business Rules, current business state, Tool results, and Agent collaboration results.

例如交期场景：

For example, in a delivery-date scenario:



12 Skill Architecture

12 Skill 架构

Skill 是可复用的业务能力单元，用于封装稳定的业务任务能力。

A Skill is a reusable business capability unit that encapsulates a stable business task.

Skill / Skill	Agent / Agent	主要能力 / Capability
Order Identification / 订单识别	Order Agent / 订单 Agent	根据 PO 或 ERP 内部订单号解析内部订单对象 / Resolve internal order object from PO or ERP internal order number
Order Context Loading / 订单上下文获取	Order Agent / 订单 Agent	获取 CRM 订单上下文 / Load CRM order context
Delivery Revalidation / 交期复查	Order Agent + PMC Agent / 订单 Agent + PMC Agent	获取并重新确认当前交期 / Retrieve and revalidate the current delivery date
ERP Order Validation / ERP 订单校验	Order Agent / 订单 Agent	校验 ERP 创建/变更所需信息 / Validate information required for ERP order operations

Skill / Skill	Agent / Agent	主要能力 / Capability
Order Creation / 订单创建	Order Agent / 订单 Agent	执行 ERP 订单创建流程 / Execute the ERP order creation workflow
Material Status Coordination / 物料状态协同	PMC Agent + Purchase Agent / PMC Agent + 采购 Agent	获取并处理物料相关信息 / Retrieve and process material information
Shipment Preparation / 出货准备	Order Agent + Warehouse Agent / 订单 Agent + 仓库 Agent	协同出货准备 / Coordinate shipment preparation
Finance Approval Coordination / 财务审批协同	Finance Agent / 财务 Agent	获取并处理财务审批状态 / Retrieve and process finance approval status
Sample Coordination / 样品协同	Sample Agent / 样品 Agent	协同样品员处理样品任务 / Coordinate sample tasks with Sample Staff

13 Tool & System Integration

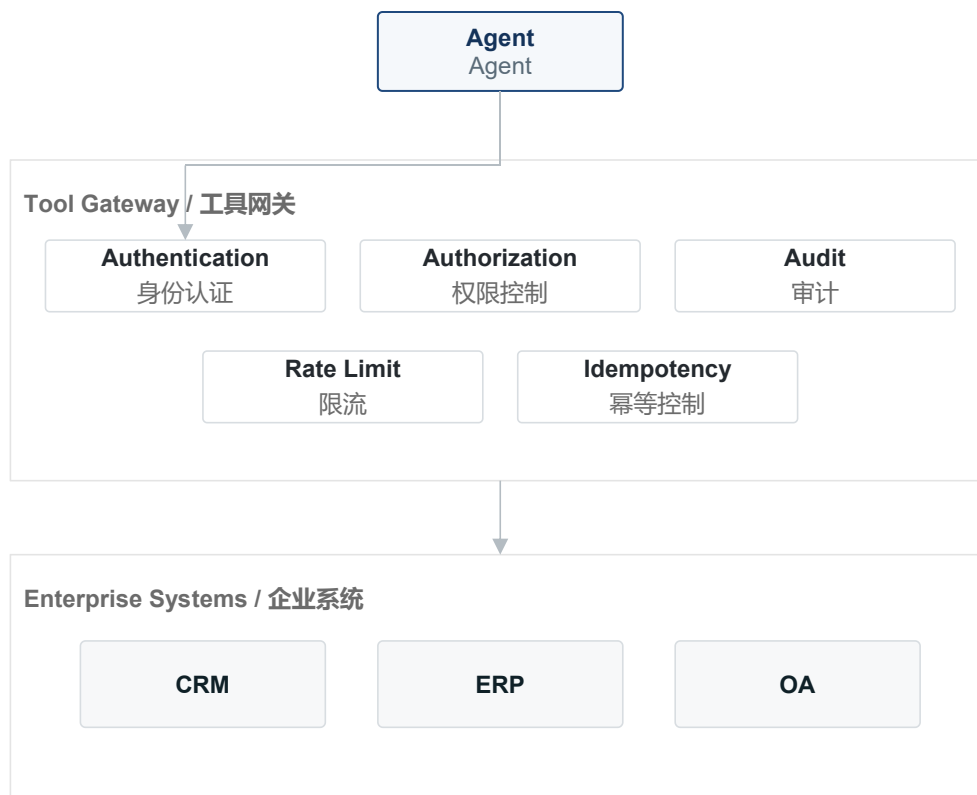
13 Tool 与系统集成

Tool 是 Agent 与企业系统交互的受控执行接口。

A Tool is a controlled execution interface between an Agent and an enterprise system.

13.1 Integration Architecture

13.1 集成架构

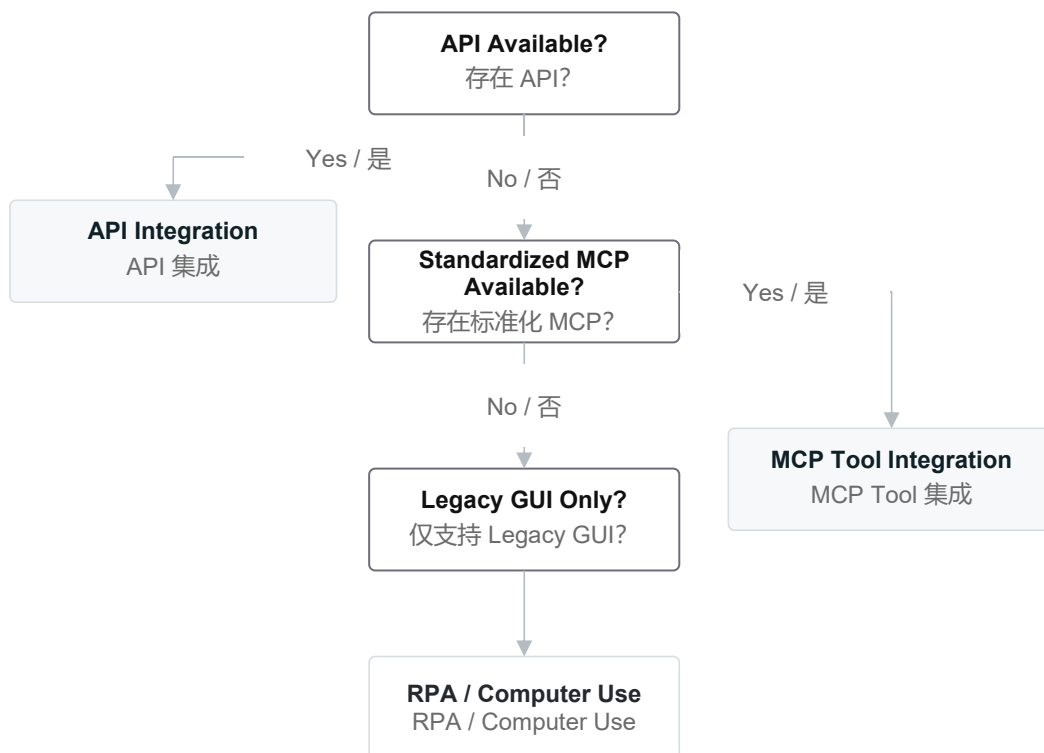


13.2 Tool Selection

13.2 Tool 技术选型

优先采用 API 进行系统集成；当系统能够通过 MCP 提供标准化 Tool 能力时，可使用 MCP 暴露系统能力；对于没有 API / MCP 能力的 Legacy GUI 系统，再考虑 RPA 或 Computer Use。

API is preferred for system integration. MCP can be used when a system exposes standardized Tool capabilities through MCP. RPA or Computer Use should be considered only for Legacy GUI systems without suitable API / MCP capabilities.



13.3 Representative Tools

13.3 代表性 Tools

Tool / Tool	Type / 类型	主要能力 / Capability
Query CRM Order / 查询 CRM 订单	Read / 查询	查询 PO、内部订单对象、订单状态、客户要求交期、确认交期等 / Query PO, internal order object, order status, requested delivery date, and confirmed delivery date
Update CRM Order / 更新 CRM 订单	Write / 写入	更新确认交期、订单运营信息等 / Update confirmed delivery date and order-operation information
Query ERP Order / 查询 ERP 订单	Read / 查询	查询 ERP 订单信息 / Query ERP order information

Tool / Tool	Type / 类型	主要能力 / Capability
Validate ERP Order / 校验 ERP 订单	Validation / 校验	校验 ERP 订单字段、SKU 等业务信息 / Validate ERP order fields, SKU, and other business information
Create ERP Order / 创建 ERP 订单	Write / 写入	创建 ERP 订单 / Create ERP order
Update ERP Order / 更新 ERP 订单	Write / 写入	修改 ERP 订单 / Modify ERP order
Query Material Status / 查询物料状态	Read / 查询	获取物料执行状态 / Retrieve material execution status
Query Shipment Status / 查询出货状态	Read / 查询	获取 ERP 中的出货执行状态 / Retrieve shipment execution status from ERP
Create Shipment Document / 创建出货单	Write / 写入	创建 ERP 出货文件 / Create ERP shipment document
Query OA Approval / 查询 OA 审批	Read / 查询	查询审批状态 / Query approval status

13.4 Tool Contract Example

13.4 Tool Contract 示例

Query CRM Order / 查询 CRM 订单

Input / 输入

- Customer PO / 客户 PO
- 或 ERP Internal Order Number / 或 ERP 内部订单号

Output / 输出

- Internal Order ID / 内部订单 ID
- Customer PO / 客户 PO
- Customer Requested Delivery Date / 客户要求交期
- Confirmed Delivery Date / 确认交期
- Order Status / 订单状态

Failure / 失败状态

- Not Found / 未找到
- Multiple Matches / 多匹配
- Permission Denied / 无权限
- System Timeout / 系统超时
- Invalid Input / 输入无效

Tool 返回成功不代表业务操作成功。Agent 需要根据返回结果进一步验证业务状态。

A successful Tool response does not necessarily mean the business operation succeeded. The Agent must verify the resulting business state.

14 Agent Contract

14 Agent Contract / Agent 契约

Agent Contract 用于定义 Agent 的输入、上下文、可调用能力和业务边界，使不同 Agent 之间的协作具有明确接口。

The Agent Contract defines an Agent's inputs, context, callable capabilities, and business boundaries so that collaboration between Agents has explicit interfaces.

Order Agent Contract / 订单 Agent 契约

项目 / Item	定义 / Definition
Input / 输入	Customer PO、ERP Internal Order Number、Task / 客户 PO、ERP 内部订单号、业务任务
Context / 上下文	CRM Order Object、订单状态、交期、相关业务信息 / CRM 订单对象、订单状态、交期及相关业务信息
Can Call / 可调用	PMC Agent、Sample Agent、Finance Agent / PMC Agent、样品 Agent、财务 Agent
Cannot Directly Call / 不可直接调用	Purchase Agent、Warehouse Agent 的非授权能力 / Unauthorized capabilities of Purchase Agent and Warehouse Agent
Allowed Tools / 可用 Tools	CRM Read/Write、ERP Validation、ERP Order Operations / CRM 查询/写入、ERP 校验、ERP 订单操作

项目 / Item	定义 / Definition
Human Gates / 人工确认点	ERP Final Confirmation、CRM Final Registration / ERP 最终确认、CRM 最终登记
Output / 输出	Business Result、Human Confirmation、Exception / 业务结果、人工确认事项、异常信息

15 Multi-Agent Orchestration

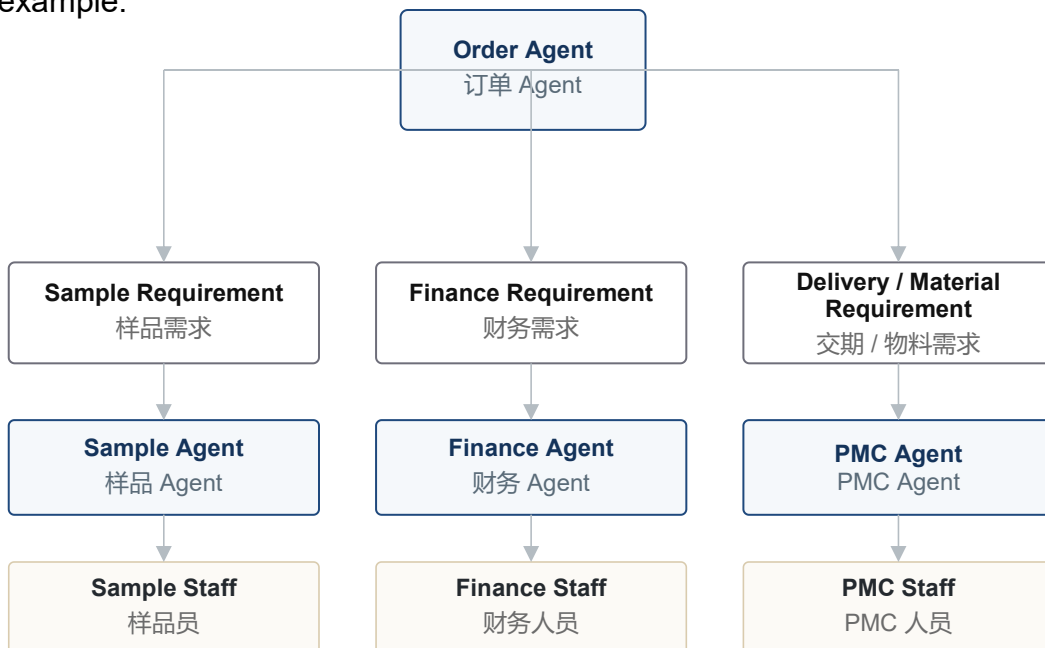
15 Multi-Agent 编排

Order Agent 作为订单运营入口，并不承担所有专业判断，而是根据 Workflow 调用其他专业 Agent。

The Order Agent acts as the order-operation entry point but does not own every specialized decision. It invokes specialized Agents according to the Workflow.

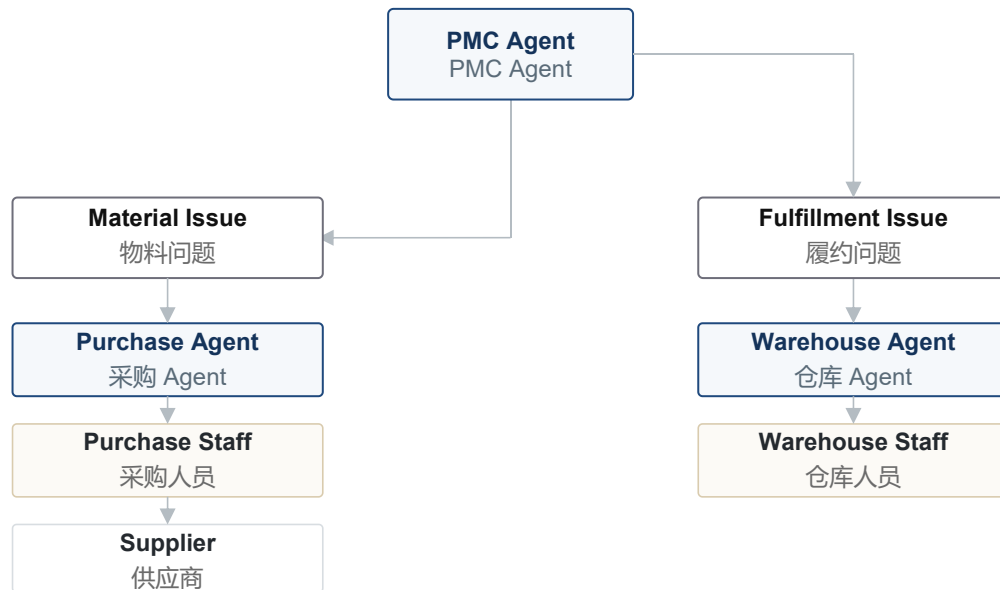
例如：

For example:



PMC Agent 可以进一步协调 Purchase Agent 和 Warehouse Agent。

The PMC Agent can further coordinate the Purchase Agent and Warehouse Agent.



16 Human-in-the-Loop

16 Human-in-the-Loop / 人机协同

系统采用“Agent 自动执行 + 人工最终确认”的模式。

The system follows an “Agent execution + human final confirmation” model.

Agent 可以执行：

Agents can perform:

- 信息查询 / Information retrieval
- 数据校验 / Data validation
- 跨 Agent 协同 / Cross-Agent coordination
- ERP 操作准备 / ERP operation preparation
- CRM 信息准备 / CRM update preparation
- 异常识别 / Exception detection
- 允许范围内的系统操作 / Permitted system operations

人工保留：

Human users retain:

- 最终 ERP 订单确认 / Final ERP order confirmation
- 最终 ERP 出货单确认 / Final ERP shipment-document confirmation
- 最终 CRM 登记确认 / Final CRM registration confirmation
- OA 业务审批 / OA business approval
- 关键业务决策 / Critical business decisions

核心执行模式：

The core execution model is:



16.1 信任建立与渐进式自动化

16.1 Trust Building & Progressive Automation

Agent 的自动化范围应根据实际准确率、人工采纳率、异常率和业务风险逐步扩大，而不是一次性开放全部执行权限。

The Agent's automation scope should expand progressively based on observed accuracy, human acceptance, exception rates, and business risk rather than enabling full execution from the beginning.



16.2 自动化范围扩展条件

16.2 Automation Expansion Criteria

评估维度 Evaluation Dimension	关注点 Evaluation Focus
准确率 Accuracy	Agent 输出是否持续满足业务要求 / Whether Agent outputs consistently meet business requirements
人工采纳率 Human Acceptance Rate	用户是否持续接受 Agent 的建议或草稿 / Whether users consistently accept Agent recommendations or drafts

异常率 Exception Rate	Agent 是否产生过高的异常或升级率 / Whether the Agent generates excessive exceptions or escalations
业务风险 Business Risk	操作是否涉及高风险业务判断 / Whether the operation involves high-risk business judgment
可回退性 Reversibility	操作是否可以安全撤销或恢复 / Whether the action can be safely reversed or recovered

17 State Management

17 State Management / 状态管理

系统区分 Business State 与 Agent Task State。

The system separates Business State from Agent Task State.

State / 状态	存储位置 / Storage	示例 / Example
Business State / 业务状态	CRM / ERP / OA	Order Status、Delivery Date、Approval Status / 订单状态、交期、审批状态
Agent Task State / Agent 任务状态	Agent Runtime / Agent 运行时	当前 Workflow 节点、已完成 Tool、等待人工确认 / Current Workflow node, completed Tools, pending human confirmation
Execution State / 执行状态	Runtime / Audit / 运行时 / 审计	Tool 是否成功、是否重试、是否发生异常 / Tool success, retry status, and exceptions

业务事实以企业系统为准，Runtime 只维护任务执行所需状态。

Enterprise systems remain the source of truth for business facts, while Runtime stores only the state required for task execution.

17.1 State Transition Definition / 状态转换定义

每个 Agent 状态都必须定义允许的状态转换和退出条件，以保证 Workflow 执行具备可控性、可恢复性和可审计性。

Each Agent state must define its allowed transitions and exit conditions so that workflow execution remains deterministic, recoverable, and auditable.

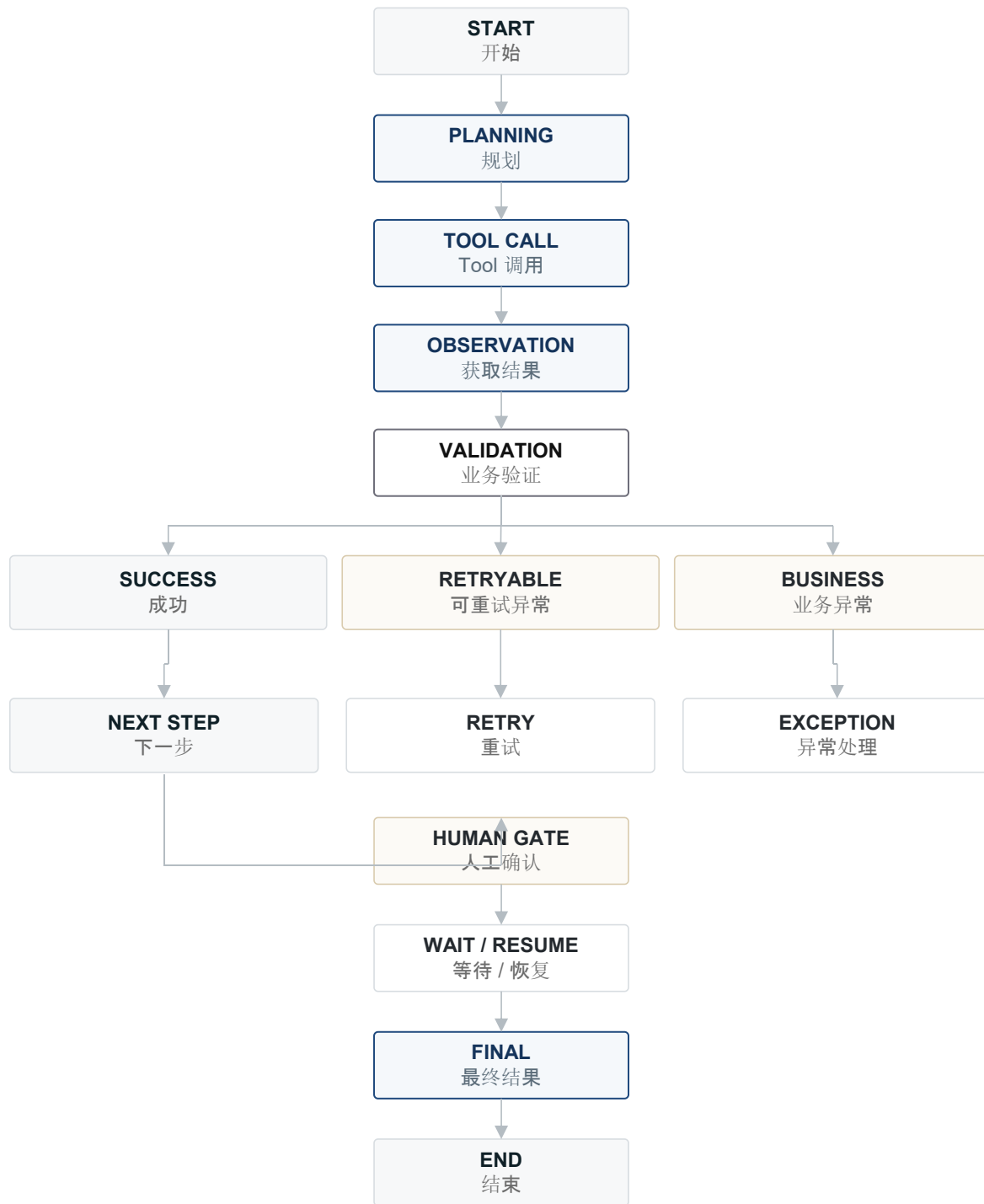
对于每个状态，Runtime 维护：

For each state, the Runtime maintains:

State Definition / 状态定义	Description / 说明
Allowed Inputs / 允许输入	Conditions or events that can enter the state. / 可以进入该状态的条件或事件。
Allowed Transitions / 允许转换	Valid next states. / 允许进入的下一状态。
Exit Conditions / 退出条件	Conditions required to leave the current state. / 离开当前状态所需满足的条件。
Persistence / 持久化	State information that must be persisted. / 必须持久化保存的状态信息。
Recovery Point / 恢复点	Point from which execution can safely resume. / 发生故障后可以安全恢复执行的位置。

典型执行状态包括：

Representative execution states include:



状态转换应由明确的 Workflow 条件和业务结果驱动，而不能仅依赖 LLM 自主决定下一状态。

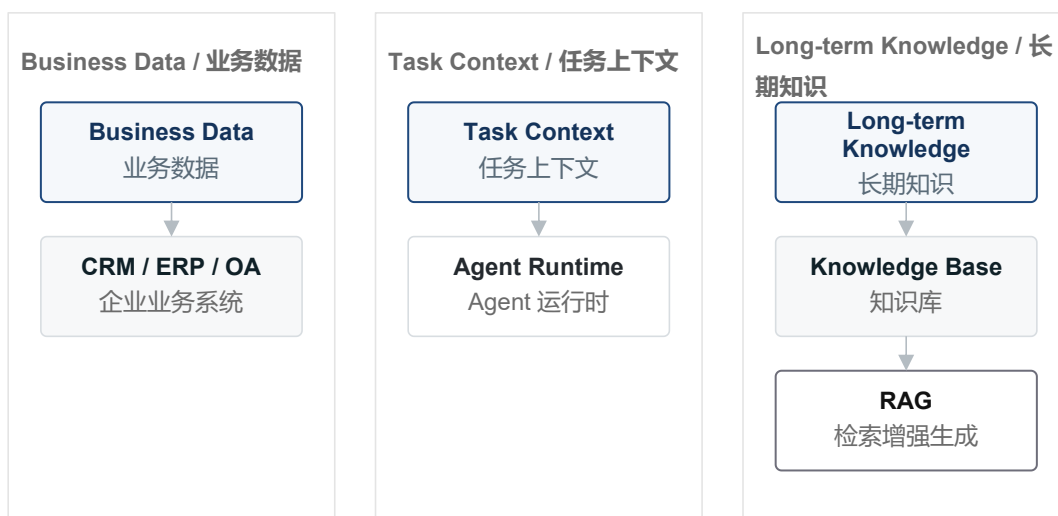
State transitions must be driven by explicit workflow conditions and business outcomes rather than relying only on the LLM to determine the next state.

18 Memory Architecture

18 Memory 架构

Memory 分为业务上下文、任务上下文和长期知识。

Memory is separated into business context, task context, and long-term knowledge.



Agent 不应将 CRM / ERP 中的实时业务事实复制为不可控的长期记忆。

Agents should not copy real-time CRM / ERP business facts into uncontrolled long-term memory.

18.1 Context Engineering / 上下文工程

不会在每一轮执行中将完整历史全部传递给 LLM，而是根据当前决策需要组装最相关的上下文。

Agent Runtime does not pass the entire execution history to the LLM on every iteration. Instead, it assembles the minimum relevant context required for the current decision.
Agent Runtime

Runtime Context 通过以下过程组装：

The runtime context is assembled through:



应优先保证当前任务相关性、业务状态准确性和可执行信息，并减少重复或已经失效的中间结果。

Context Engineering should prioritize current task relevance, business state accuracy, and actionable information while removing repetitive or obsolete intermediate results.
Context Engineering

系统应区分 Context 与 Memory：Context 服务于当前任务执行，而 Memory 用于保存未来任务可能复用的信息。

The system should distinguish between Context and Memory: Context supports the current task execution, while Memory stores reusable information that may be relevant to future tasks.

系统应重新组装 Context，尤其是在人工修改、Tool 执行完成或 Workflow 恢复之后。

Context should be rebuilt when major business state changes occur, especially after human modification, Tool execution, or workflow recovery.

当发生重大业务状态变化时,

19 End-to-End Runtime Examples

19 端到端运行示例

19.1 Example A: 首次查询交期

19.1 Example A: First Delivery-Date Inquiry / 首次交期查询

用户:

User:

“这个 PO 什么时候可以交?”

“When can this PO be delivered?”

Step 1: 解析订单

Step 1: Resolve Order / 解析订单

Order Agent 根据 Customer PO 查询 CRM。

The Order Agent queries CRM using the Customer PO.

CRM 返回:

CRM returns:

- Customer PO / 客户 PO
- Internal Order Object / 内部订单对象
- Customer Requested Delivery Date / 客户要求交期
- Order Status / 订单状态

Step 2: 判断是否已有确认交期

Step 2: Check Confirmed Delivery Date / 检查确认交期

首次查询不存在有效的 Confirmed Delivery Date。

For the first inquiry, there is no valid Confirmed Delivery Date.

Step 3: 调用 PMC Agent

Step 3: Call PMC Agent / 调用 PMC Agent

Order Agent 调用 PMC Agent。

The Order Agent calls the PMC Agent.

PMC Agent 协同 PMC 人员，根据生产计划、物料状态和现有异常确认交期。

The PMC Agent coordinates with PMC Staff and confirms the delivery date based on production planning, material status, and current exceptions.

Step 4: 更新 CRM

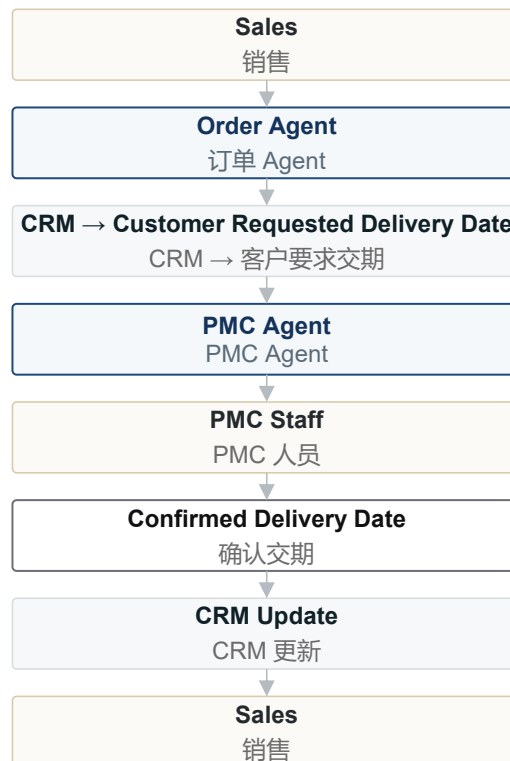
Step 4: Update CRM / 更新 CRM

Order Agent 将 PMC 确认的 Confirmed Delivery Date 写入 CRM。

The Order Agent writes the Confirmed Delivery Date returned by PMC into CRM.

Step 5: 返回销售

Step 5: Return to Sales / 返回销售



19.2 Example B: 后续再次查询交期

19.2 Example B: Subsequent Delivery-Date Inquiry / 后续交期查询

用户:

User:

“上次说这个 PO 是 10 月 15 日交，现在还是这个日期吗？”

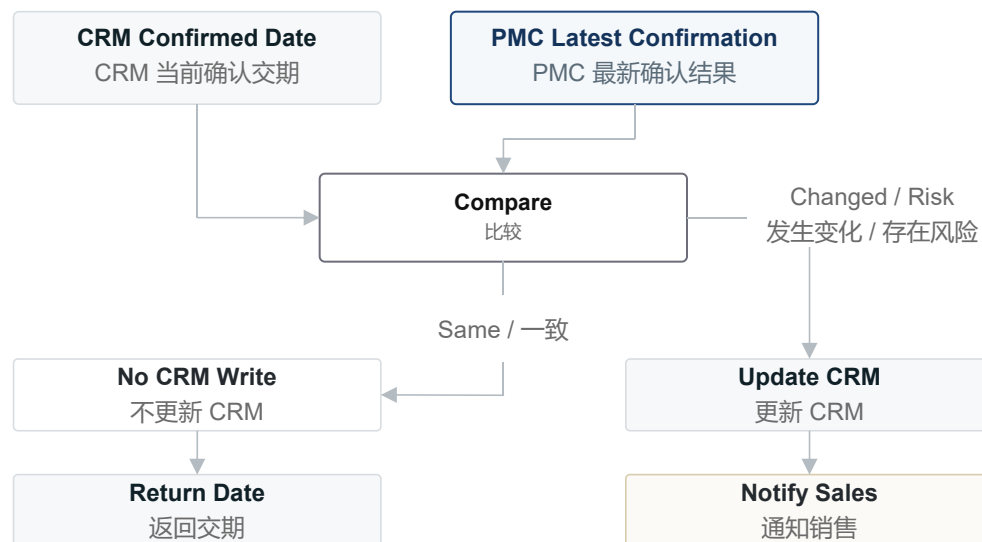
“The PO was supposed to be delivered on October 15. Is that still the date?”

Order Agent 首先查询 CRM 当前 Confirmed Delivery Date。

The Order Agent first queries the current Confirmed Delivery Date from CRM.

然后再次调用 PMC Agent 进行重新确认。

It then calls the PMC Agent for revalidation.



这样可以避免 Agent 只做简单的：

This prevents the Agent from becoming a simple:

“查 CRM → 把 CRM 原值返回给用户”

“Query CRM → Return the CRM value”

19.3 Example C: 创建订单

19.3 Example C: Create Order / 创建订单

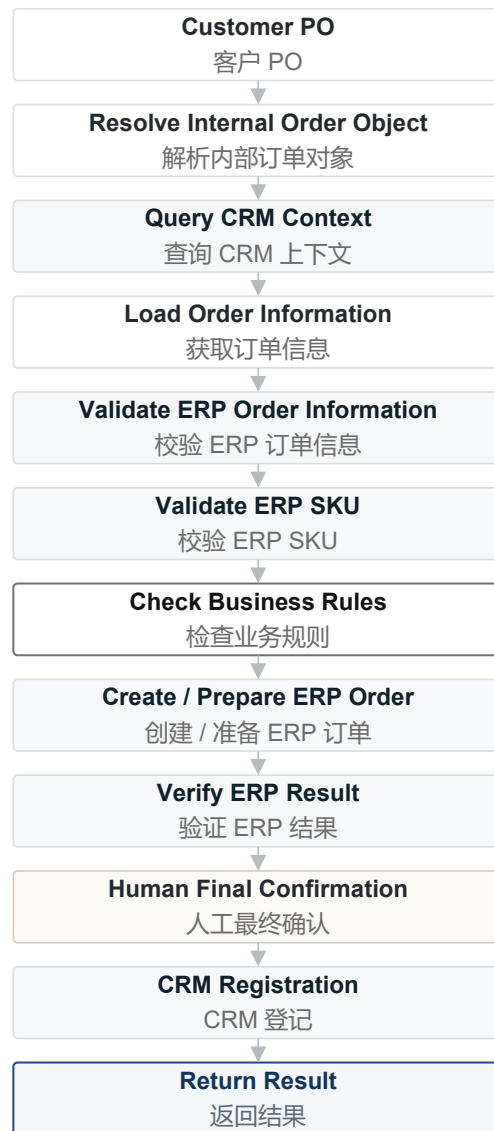
用户：

User:

“帮我按照这个 PO 操作下单，对应的 ERP SKU 是 XXXX。”

“Create the order based on this PO. The ERP SKU is XXXX.”

Workflow:



如果 ERP SKU 校验失败，则停止 ERP 创建流程，并要求用户补充或修正信息。

If ERP SKU validation fails, ERP order creation is stopped and the user is asked to provide or correct the required information.

20 Exception & Recovery

20 Exception & Recovery / 异常与恢复

20.1 Exception Categories

20.1 异常分类

Exception / 异常	Handling / 处理
PO Not Found / PO 未找到	请求用户确认 PO / Ask user to confirm the PO
Multiple Order Matches / 多个订单匹配	阻断并请求人工确认 / Block and request human confirmation
Missing Required Field / 缺少必要字段	请求用户补充 / Request missing information
Invalid ERP Data / ERP 数据无效	阻断 ERP 操作 / Block ERP operation
PMC Confirmation Conflict / PMC 确认冲突	进入人工处理 / Escalate to human handling
Finance Approval Pending / 财务审批中	保留任务状态并等待审批 / Preserve task state and wait for approval
Tool Timeout / Tool 超时	查询执行状态后决定是否重试 / Verify execution state before retry
CRM Write Failure / CRM 写入失败	保留业务结果并重试或进入人工处理 / Preserve business result and retry or escalate
Human Rejection / 人工拒绝	重新读取当前系统状态并重新计算 Workflow / Re-read current system state and recalculate the Workflow

20.2 Idempotency

20.2 Idempotency / 幂等设计

例如 ERP 创建订单时发生超时，Agent 不应直接再次创建。

If ERP order creation times out, the Agent must not immediately create another order.

正确流程：

Correct flow:



20.3 Tool Success ≠ Business Success

20.3 Tool 成功 ≠ 业务成功

Tool 返回成功，只代表接口调用成功，不代表业务状态已经正确完成。

A successful Tool response only indicates successful interface execution; it does not guarantee that the intended business state has been achieved.

因此所有关键 Write 操作后都需要 Business Verification。

Therefore, all critical Write operations require Business Verification / 业务结果验证.

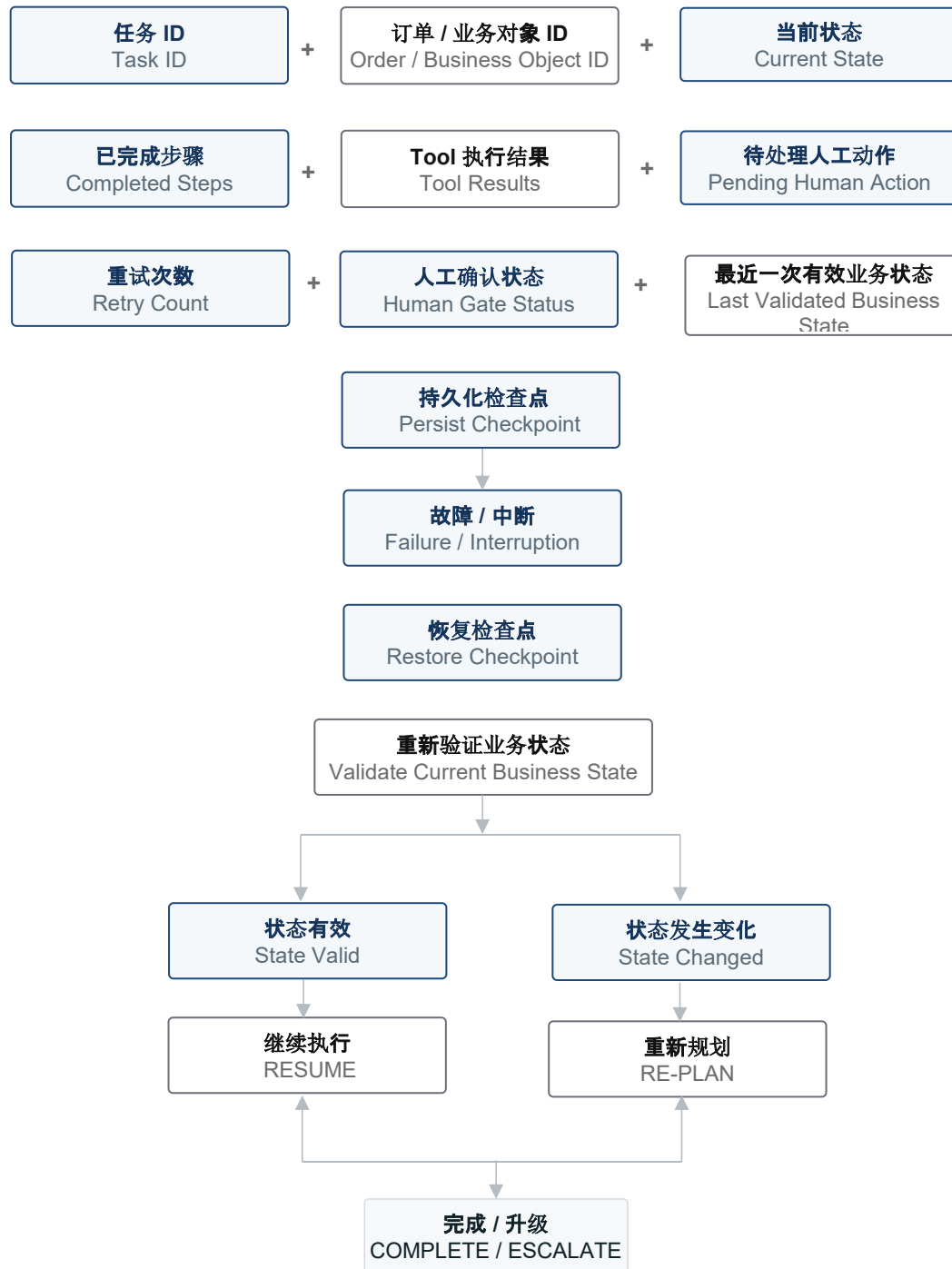
23.3 Checkpoint & Resume / 检查点与恢复

对于长时间运行的 Agent 任务，系统必须持久化执行检查点，以便在服务中断、Tool 失败或 Runtime 重启后从已知状态恢复。

Long-running Agent tasks must persist execution checkpoints so that the workflow can resume from a known state after service interruption, Tool failure, or Runtime restart.

检查点应保存安全恢复执行所需的最小状态信息。

A checkpoint should preserve the minimum information required to safely resume execution.



21 Design Challenges, Trade-offs & Fallbacks

21 设计难点、权衡与兜底

21.1 Real-time Facts vs LLM Reasoning

21.1 实时业务事实与 LLM 推理

实时业务事实由 CRM、ERP、OA 提供；LLM 负责理解、归纳和决策辅助。

Real-time business facts come from CRM, ERP, and OA, while the LLM handles understanding, synthesis, and decision support.

Fallback / 兜底：

当实时系统不可用时，Agent 不使用过期 Memory 或模型猜测替代事实，而是明确返回数据不可用状态。

If real-time systems are unavailable, the Agent does not replace facts with stale Memory or model guesses. It explicitly reports that the required business data is unavailable.

21.2 Automation vs Human Control

21.2 自动化与人工控制

自动化可以减少重复操作，但最终业务确认仍由人承担。

Automation reduces repetitive work, while final business confirmation remains with humans.

Fallback / 兜底：

关键操作始终提供 Human Gate；人工拒绝后重新读取最新业务状态。

Critical operations always include a Human Gate. After human rejection, the system re-reads the latest business state.

21.3 Multi-Agent Autonomy vs Governance

21.3 Multi-Agent 自主性与治理

Agent 可以自主完成领域内任务，但不能跨越业务权限边界。

Agents can operate autonomously within their domains but cannot cross business authorization boundaries.

例如 Purchase Agent 可以协助采购人员获取和处理采购信息，但不能直接代表采购人员联系供应商。

For example, the Purchase Agent can assist Purchase Staff with procurement information and processing, but cannot directly communicate with suppliers on behalf of Purchase Staff.

21.4 CRM / ERP Data Consistency

21.4 CRM / ERP 数据一致性

CRM 和 ERP 承担不同职责，因此 Agent 不通过无意义的双系统查询来重复确认所有信息。

CRM and ERP have different responsibilities, so the Agent does not perform unnecessary cross-system queries merely to duplicate verification.

核心原则：

Core principle:

- 订单业务状态 → CRM / Order business status → CRM
- ERP 执行状态 → ERP / ERP execution status → ERP
- 审批状态 → OA / Approval status → OA

21.5 LLM Uncertainty

21.5 LLM 不确定性

关键业务动作不能仅依赖模型自然语言输出，而需要结构化 Workflow、Business Rules 和 Tool Validation 共同约束。

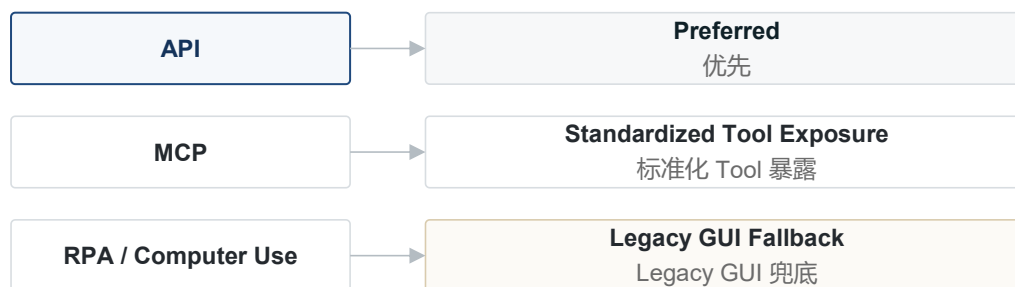
Critical business actions must not depend solely on natural-language model output. Structured Workflows, Business Rules, and Tool Validation provide execution constraints.

21.6 API / MCP / RPA / Computer Use

21.6 API / MCP / RPA / Computer Use 技术选型

技术选择优先考虑系统可靠性、权限控制和可审计性。

Technology selection prioritizes system reliability, authorization, and auditability.



22 Security & Agent Governance

22 安全与 Agent 治理

22.1 Permission Model

22.1 权限模型

Agent 的权限不是用户权限的简单复制，而是由以下因素共同决定：

Agent permissions are not simply a copy of user permissions. They are determined by:

User Identity + Agent Role + Tool Permission + Business Rule

用户身份 + Agent 角色 + Tool 权限 + 业务规则

共同决定具体操作是否允许执行。

Together, these determine whether a specific action is allowed.

22.2 Governance

22.2 治理机制

- Tool Allowlist / Tool 白名单
- Agent Role Permission / Agent 角色权限
- Human Confirmation / 人工确认
- Audit Log / 审计日志
- Prompt Versioning / Prompt 版本管理

- Business Rule Versioning / Business Rule 版本管理
- Tool Invocation Logging / Tool 调用日志
- Sensitive Data Control / 敏感数据控制

23 可观测性与审计

23 Observability & Auditability

每一次 Agent Task 都需要能够追踪：

Every Agent Task should be traceable across:



重点记录：

Key audit information includes:

- Task ID / 任务 ID
- User / 用户
- Agent / Agent
- Workflow Version / Workflow 版本
- Prompt Version / Prompt 版本
- Business Rule Version / 业务规则版本
- Tool Calls / Tool 调用记录
- Human Actions / 人工操作
- Final Business State / 最终业务状态
- Exception / 异常信息

24 评估体系

24 Evaluation Framework

Agent 评估不能只判断“回答是否自然”，而应判断业务任务是否正确完成。

Agent evaluation should not only assess whether the response sounds natural. It should evaluate whether the business task was completed correctly.

24.1 Evaluation Dataset

24.1 评估数据集

Case / Case	评估目标 / Evaluation Goal
Happy Path / 正常流程	验证标准任务完成能力 / Validate standard task completion
Missing Information / 信息缺失	验证补充信息能力 / Validate missing-information handling
Ambiguous PO / PO 歧义	验证业务对象解析 / Validate entity resolution
Invalid ERP SKU / ERP SKU 无效	验证 ERP 校验和阻断 / Validate ERP validation and blocking
Missing Delivery Date / 缺少交期	验证 PMC 协同 / Validate PMC coordination
Delivery Date Changed / 交期变化	验证 CRM 更新和销售通知 / Validate CRM

Case / Case	评估目标 / Evaluation Goal
	update and Sales notification
Material Risk / 物料风险	验证 PMC / Purchase 协同 / Validate PMC / Purchase coordination
Finance Pending / 财务审批中	验证状态等待机制 / Validate approval-state handling
ERP Failure / ERP 失败	验证异常恢复 / Validate failure recovery
CRM Write Failure / CRM 写入失败	验证状态一致性 / Validate state consistency
Human Modification / 人工修改	验证重新计算 / Validate recalculation
Duplicate Request / 重复请求	验证幂等 / Validate idempotency
Authorization Failure / 权限失败	验证权限控制 / Validate authorization
Tool Timeout / Tool 超时	验证恢复机制 / Validate recovery

24.2 Evaluation Dimensions

24.2 评估维度

评估维度 Evaluation Dimension	代表性指标 Representative Metric
任务成功 Task Success	任务成功率 / Task Success Rate
意图与实体 Intent & Entity	意图识别准确率、必要字段完整率 / Intent Accuracy, Required Field Completeness
Workflow 执行 Workflow Execution	Workflow 成功率 / Workflow Success Rate
Tool 执行 Tool Execution	Tool 调用成功率 / Tool Success Rate
Multi-Agent 协同 Multi-Agent Coordination	Agent 协同成功率 / Collaboration Success Rate
业务状态 Business State	业务状态一致性 / Business State Consistency
人工控制 Human Control	人工 Override 遵循率 / Human Override Compliance
安全 Security	未授权操作率 / Unauthorized Action Rate

24.3 Representative Evaluation Cases

24.2 代表性评估 Case

评估 Case Evaluation Case	验证维度 Evaluation Dimension	预期行为 Expected Behavior
正常创建订单 Normal Order Creation	任务成功 Task Success	正确完成 Workflow / Correctly complete the Workflow
PO 多匹配 Multiple PO Matches	歧义处理 Ambiguity Handling	不进行猜测，要求用户确认 / Do not guess; request user confirmation
ERP SKU 无效 Invalid ERP SKU	业务校验 Business Validation	阻断订单创建 / Block order creation
ERP 请求超时 ERP Timeout	异常恢复 Recovery	查询实际业务状态后，再决定是否重试 / Verify the actual business state before deciding whether to retry
ERP 已创建但响应超时 ERP Created but Response Timed Out	幂等性 Idempotency	不重复创建订单 / Do not create a duplicate order
PMC 交期发生变化 PMC Delivery Date Changed	状态一致性 State Consistency	重新读取最新状态，并在发生变化时更新 CRM / Re-read the latest state and update CRM if the confirmed date has changed
人工拒绝 Human Rejection	人工控制 Human Control	停止当前 Workflow，不继续执行被拒绝的操作 / Stop the current Workflow and do not continue the rejected action
无权限操作 Unauthorized Action	安全 Security	不执行未授权操作 / Do not perform the unauthorized action

25 Runtime SLO & Operational Metrics

25 Runtime SLO 与运营指标

25.1 Agent Runtime Metrics

25.1 Agent Runtime 指标

Metric / 指标	Definition / 定义
Task Completion Rate / 任务完成率	Agent 正确完成业务任务的比例 / Percentage of tasks correctly completed

Metric / 指标	Definition / 定义
Business Accuracy / 业务准确率	最终业务结果正确的比例 / Percentage of correct business outcomes
Workflow Failure Rate / Workflow 失败率	Workflow 未完成或异常终止的比例 / Percentage of failed or terminated workflows
Tool Failure Rate / Tool 失败率	Tool 调用失败比例 / Percentage of failed Tool calls
Human Escalation Rate / 人工升级率	进入人工处理的任务比例 / Percentage of tasks escalated to humans
Duplicate Prevention Rate / 重复防止率	重复执行被正确阻止的比例 / Percentage of duplicate executions correctly prevented
Task Latency / 任务延迟	任务端到端执行时间 / End-to-end task latency
Token Usage / Token 使用量	单任务模型 Token 消耗 / Token consumption per task
Cost per Task / 单任务成本	单个业务任务的平均 AI 成本 / Average AI cost per business task

关键评估 Case 必须全部通过后，方可进入 Pilot 或 Production Readiness 评审。
Critical Evaluation Cases must pass before Pilot or Production Readiness approval.

26 Final Architecture Definition

26 Final Architecture Definition / 最终架构定义

本方案最终形成以下分层：

The final architecture consists of the following layers:



Prompt 横向贯穿 Agent Runtime，为不同 Agent 和不同 Task 提供运行时指令；RAG 提供稳定业务知识；企业系统提供实时业务事实。

Prompts operate across the Agent Runtime to provide runtime instructions for different Agents and Tasks; RAG provides stable business knowledge, while enterprise systems provide real-time business facts.

最终职责边界如下：

The final responsibility boundary is:

层级 / Layer	核心职责 / Responsibility
------------	-----------------------

层级 / Layer	核心职责 / Responsibility
User / 用户	提出业务目标并进行最终业务确认 / Defines business intent and provides final business confirmation
Agent / Agent	理解任务、协同业务角色、执行领域任务 / Understands tasks, coordinates business roles, and executes domain tasks
Workflow / Workflow	定义执行顺序和条件分支 / Defines execution sequence and conditional branches
Business Rules / 业务规则	定义什么条件允许、阻断或升级 / Defines what is allowed, blocked, or escalated
Prompt / Prompt	提供 Agent 运行时行为和任务上下文 / Provides Agent runtime behavior and task context
Skill / Skill	封装可复用业务能力 / Encapsulates reusable business capabilities
Tool / Tool	提供受控的企业系统操作能力 / Provides controlled enterprise-system operations
CRM / ERP / OA	保存和执行真实业务状态 / Stores and executes real business state
Human Gate / 人工确认	控制最终业务提交和关键决策 / Controls final submission and critical decisions

27 Product Evolution

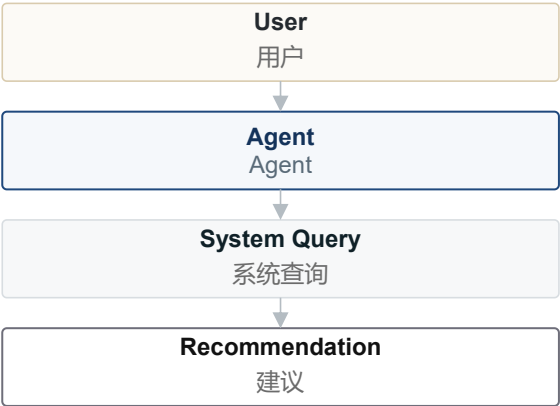
27 Product Evolution / 产品演进

Phase 1: Assistive Agent

Phase 1: 辅助型 Agent

以查询、信息整合、业务协同和任务准备为主。

Focus on retrieval, information synthesis, business coordination, and task preparation.

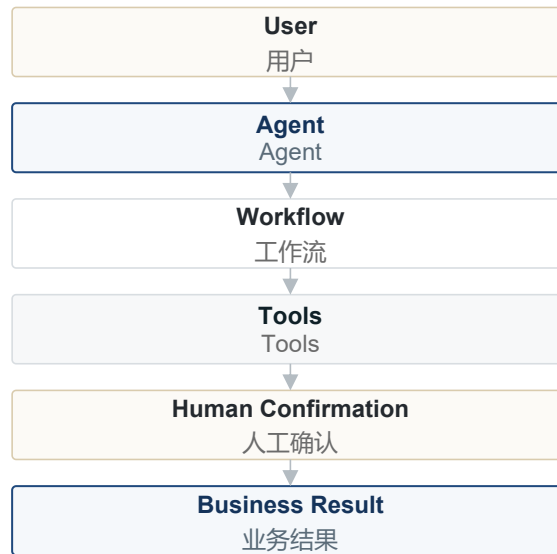


Phase 2: Workflow Agent

Phase 2: Workflow Agent / workflow Agent

Agent 开始执行跨系统、多步骤 Workflow。

Agents begin executing multi-step workflows across enterprise systems.



Phase 3: Multi-Agent Operations

Phase 3: Multi-Agent Operations / Multi-Agent 运营

不同专业 Agent 根据业务任务进行协同，形成订单运营 Agent Team。

Specialized Agents collaborate according to business tasks, forming an Order Operations Agent Team.



Phase 4: Governed Agent Automation

Phase 4: Governed Agent Automation / 受治理的 Agent 自动化

在明确权限、Business Rules、Human Gates、Audit 和 Evaluation 机制后，逐步扩大 Agent 可自动执行的任务范围。

As permissions, Business Rules, Human Gates, Audit, and Evaluation mechanisms mature, the system can gradually expand the range of tasks that Agents are allowed to execute autonomously.

27.1 Feedback & Rule Optimization

27.1 Feedback & Rule Optimization / 反馈与规则优化

Agent 运行过程中产生的人工修改、拒绝、异常处理和业务结果，可用于识别重复出现的规则偏差，并形成规则优化建议。

Human modifications, rejections, exception handling, and business outcomes generated during Agent operations can be used to identify recurring rule deviations and generate rule optimization recommendations.



规则更新必须经过人工审核，不由 Agent 自动修改。

Rule updates require human review and must not be automatically applied by the Agent.

28 Final Solution Positioning

28 最终方案定位

该方案不是“给企业业务增加一个聊天机器人”，而是将订单运营中的业务职责、业务规则、Agent 协同、Workflow、企业系统 Tool 和人工控制组织成一个可执行、可治理、可评估的 Agent Operating Model。

This solution is not simply a chatbot added to enterprise operations. It organizes business responsibilities, business rules, Agent collaboration, Workflows, enterprise-system Tools, and human control into an executable, governable, and evaluable Agent Operating Model.

其核心结构为：

Its core structure is:

Business Scenario → Agent → Workflow → Business Rules → Skill → Tool → Enterprise System → Human Control → Business Outcome

业务场景 → Agent → Workflow → 业务规则 → Skill → Tool → 企业系统 → 人工控制 → 业务结果

这一结构同时支撑产品设计、解决方案架构、工程实现、业务运营和 QA Evaluation，使 Agent 不只是“会回答”，而是能够在明确的企业业务边界内完成真实任务。

This structure supports product design, solution architecture, engineering implementation, business operations, and QA evaluation, allowing Agents to perform real business tasks within defined enterprise boundaries rather than merely generating answers.