

牛码新市白皮书 (2026 年 10 月 3 日版)

帽子军团——千军万码微任务

合约经济的度量基础设施：理论、工厂与样本



牛码新市白皮书

帽子军团——千军万码微任务
合约经济的度量基础设施：理论、工厂与样本

Micro Connect New Markets (NUMA) White Paper

HAT over MASS——Micro Agents · Simple & Short tasks
Measurement Infrastructure for the Contract Economy: Theory, Factory, and Samples

框住个体 · 释放群体

Govern The Individual | Unleash The Collective

执行摘要

《帽子军团——千军万码微任务》是牛码新市的第三本白皮书，发布日期为 2026 年 10 月 3 日。三本白皮书独立成篇、内容递进：8·03 提出合同现金流投融资的市场框架，回答“要建立怎样的市场”；9·03 提出双罗盘的度量规则，回答“怎样理解和度量合同与生意”；10·03 进一步回答：这些专业工作，由谁来做，又怎样组织、检验并持续完善。

一、牛码新市是谁

牛码新市是合同证券化市场的登记与度量基础设施提供方。本书把这套基础设施中的六项核心工具比作“仪器”，而制造和维护这些仪器的组织，则比作一座“AI 工厂”——帽子军团。人制定规则、承担责任，AI 按任务分工协作，让工具的建设与维护有据可查、经得起检验。

六件仪器围绕一张合同从进场到存续的全生命周期分工协作，各件仪器的建设状态及开放范围，以正式版本和实际披露为准。

1. 合同终端（Contract Terminal）是合同进场的第一站。系统按要求收集材料、逐项录入并核验，对缺失、不一致或不合要求的内容提出补充与澄清要求，把非标准化商业合同整理成结构化、可核验、符合进场要求的记录。它持续收集和核对资料，保留来源、时点及核验状态；如遇缺失或冲突，系统自动标明，不靠 AI 推断补齐事实。合同本身须合法合规、已生效、正在真实履行，现金流可由第三方核验。满足这些准入条件后，是否适合投资，仍需结合具体证据独立判断。

2. 合同登记簿（Contract Registrar）把经采集和核验的合同整理成可登记、可查询、可比较的标准化记录，载明签约双方、合同约定、当前状态并关联底层材料。条款和状态发生变更时，保留变更依据及历史版本。登记不改变合同权利义务，也不构成合同履约的保证。

3. 合同罗盘（Contract Compass）为登记在册的合同提供度量。通用罗盘看合同结构，行业罗盘看合同背后的生意，两者均设置收益维与风险维，收益与风险分别呈现、不相加、不相互抵消。因子定义、数据来源、规则版本及评分方法在相应正式版本中列明，供专业机构复核。

4. 罗盘读数（Compass Reading）按日更新已登记合同的因子数据，并依据数据性质和新增证据更新读数，持续记录经营和履约变化，标明数据时点与观察期间，保留更新前后的记录。没有新增事实依据时，保留原有读数及其时点，不因每日检查而改写数值。

5. 组合登记簿（Contract Portfolio Registry）在投资者通过持牌机构完成交易和交收后，逐日记录各组合持有哪些合同、持仓数量及其变化。订单登记簿保留原始成交和持仓事实，组

合登记簿在组合层关联这些记录。它负责记录持仓关系与变动，不承担客户资产托管职能；模拟配置与实际持仓分开标识且隔离记录，不会混入同一持仓口径。

6. 罗盘终端（Compass Terminal）是供持牌机构/境外合格投资者使用的专业查询与分析界面。使用者按自身方法论调用读数、比较资料、回查依据，形成自己的判断。专业终端按相应资格和使用授权开放，不面向公众；公众可阅读一般性行业研究资料，两类访问权限分别管理。

牛码新市的全套仪器体系为投资者与融资者双方服务。一边是店铺、工厂、农场等经营实体作为合同发行人，提供经牛码新市整理的合同资料 and 经营信息，使这些资料能够被专业机构理解、度量和核验。另一边是由持牌机构/境外合格投资者组成的牛码新市投资者，他们借助牛码新市提供的工具，把专业判断延伸到数量更多、规模更小的生意。投资决定由具有相应决策权限的主体作出，持牌机构按其承担的职能，在许可范围内履行相应义务。

读数不是评级，罗盘指北不指路。

牛码新市提供可核验、持续更新的记录和数据点，不对合约信用质量作出评级，不向投资者推荐资产。分值不能解释为回报率、违约概率或信用等级；投资者可参考罗盘读数对资产质量作出判断，但投资风险仍由投资者自行承担。

二、三本白皮书与六件套仪器体系

牛码新市用六件仪器完成合同信息的登记、度量和持续跟踪。两端是交互与读取的终端，中间由登记簿保留事实、罗盘提供度量、读数跟踪变化。这六件仪器各有分工，也可以在业务过程中互相调用，并不是必须依次经过的操作顺序；罗盘终端可贯穿专业查阅，组合登记则在交易与交收完成后记录实际持仓关系。8·03 白皮书提出市场框架，9·03 白皮书展开登记与度量相关功能，本书将进一步把这些功能整合为“六件套”仪器体系，并以“帽子军团”说明支撑其建设与维护的组织是如何运转的。

仪器是给人用的；工厂车间是让仪器的制造与维护有规则、有检验、有责任。

三、直面问题：专业能力怎样覆盖小生意

公司证券化以公司为单位，组织资本参与经营所需的权利和责任。当投融资的对象细化到一份合同，尽调、核验、定价和监测仍要逐项完成，专业成本却未必能由单笔小生意承担。合同约定了现金流安排，但约定能否履行、款项能否实际收回，仍须持续核验。过去依靠专业人才逐笔处理，任务越细、数量越多，越难算得过账；AI 有望降低资料整理、规则检查等重复工作的成本，使专业能力有机会覆盖更小的业务。

四、仪器怎么造：帽子军团

帽子军团：框住个体，释放群体

制造和维护这套仪器的工厂叫帽子军团（HAT over MASS）。MASS 由 Micro Agents（微任务智能体）与 Simple & Short Tasks（简单且短时的任务）两组词组成，同时取英文“海量”之意，组织海量 Agent 承接可以独立检验的微任务；HAT 是 Human-Commanded, AI-Operated, Trust 的缩写，即“人管控、AI 运作”的信任机制。规则由人制定、修改和废止，每个 AI 岗位由具名真人认领，AI 在授权内执行。帽子只管边界，不管动作。遇到规则空白、超出授权或需要人作自由裁量的事项，暂停相关执行，提交具备相应授权的真人处理。

这里的“微小”说的是任务粒度，不是模型能力。任务拆小，让边界更清楚、结果更容易检验；海量协作，让完整的模型能力在不同任务中释放出潜力。帽对兵守住规则与边界，兵对兵完成任务派发、执行和检验。

理想是“大财小用”，卡在“大才不能小用”，答案是“大材小用”。

“十二个足够”从任务、底座、编制和检验四个方面约束军团：任务足够小、窄、碎；底座足够多元、对冲、有性价比；编制足够多、细、冲；检验足够密、长、深。任何条件不成立，都须发出警示，查明缺口并重新校准。规则盲区、数据污染、同源偏差和误判传导需要持续识别和处理；无法微任务化的判断，交由具备相应授权的真人承担，AI 可在授权内提供资料与分析支持。

给强大的 AI，戴上人立的铁帽。

两组矛盾及其回应



图 0-1：AI 时代的两组矛盾及帽子军团的回应

五、AI 工厂的现状

牛码新市的 5000 名 AI 员工已经上岗，分布在调度管理（Administration）、业务执行（Business）、管控合规（Control）和部署支撑（Deployment）四类职能中。“上岗”指岗位纳入组织台账，不等于 5000 名 AI 员工此刻都在执行任务。

AI 员工是岗位的对外称呼，AI 岗位是稳定的组织身份：具体任务到来时，系统按岗位要求启动 Agent，调用所需能力完成当次执行。身份编码（Agent ID）连接岗位与具名真人负责人，Agent 能力表现为调度提供依据；每次实例运行形成相应运行编码，并关联所属任务及本次输入输出、调用和检查记录。同一任务可以包含多个实例和多次重做。小队可以解散，身份与记录继续保留。

从接单到归档，七站串起一次交付，也留下每一步的依据与责任。

以连锁便利店罗盘生成任务为例，任务先明确范围和交付标准，再拆成因子识别、口径说明、评分规则等工作单元，经接单、规则匹配、组队、执行、验证、收口、归档七站流转。各站记录输入、规则、产出和异常，发现问题先定位、澄清，确需修改时再退回相应环节修订、

复验。书中SIM-01是说明性模拟，不代表实际运行成果；真实任务的完成进度与检验状态，以留存的运行和验收记录为依据。

六、已形成的成果：500 个行业罗盘

行业各有尺度：五百个行业的罗盘成果

本版行业罗盘规则成果已覆盖 500 个行业。正文详细展开连锁便利店、川渝火锅门店和现制茶饮门店作为三个精读案例，附录一另收录 17 个品类案例，合计 20 例。这 20 例是理解行业罗盘的窗口，不作为统计抽样样本；覆盖 500 个行业的规则成果，不代表已完成相应行业的企业实测评分。当前案例展示指标与因子结构；企业应用前仍须核对适用版本、量化口径、数据来源、观察期间、权重和评分标准，再开展取数、实际评分与回测。已开放的行业成果及版本资料经链接查阅。

一项评分要站得住，先要说清测量什么、怎样计算、观察多久，再说明什么状态落入哪一档。同一组数据应能按同一规则复算出一致结果。收益维与风险维分别表达，不相加、不相互抵消；经营支撑强在哪里，风险又藏在哪里，都要让读者看见。

行业罗盘采用“3 类行业视角 + 3 个模型对冲”：正常经营者、老炮和骗子从不同立场看生意，其中“骗子”是寻找包装与刷分路径的对抗视角；三个模型分别形成判断，再比较假设、依据与分歧。

六鉴与执衡对候选行业因子进行设计审查、资格判断与因子组合筛选。六鉴分别检验经营意义、计算与评分的可复算性、数据路径的可信性、区分依据、相对全池的信息增量与抗操纵性，执衡据此判断候选资格，再从可选池中筛选组合。候选取得“可选”资格，仍不等于最终入选，更不等于完成企业验证。

行业研究拆成微任务并行推进，以多视角提出候选判断、多模型对冲检视，再经六鉴审查、执衡筛选——500 个行业罗盘由此加速形成，每项成果保留检验与修订的依据。

七、下一步：验证、校准与模拟运行

接下来的工作围绕四个方面展开。

一是企业取数与实际评分，让读数来自可核验的真实经营数据。

二是回测与校准，用经营和履约结果检查因子、阈值与系数，发现问题先查明原因，再作相应修订。

三是把整套系统放进明确标注为模拟的环境中持续运转，连接登记、度量、任务交接、检验、预警与处置，并纳入不利情景，从而检验各环节能否衔接、异常能否被发现和处理。

四是每月 3 日向市场公布进展，按实际进度说明做了什么、做到哪一步、还缺什么证据。

本书的 500 个行业规则成果构成了起点，后续还要通过企业应用与持续检验，验证这些规则是否适用、读数是否可靠。

模拟合同、模拟资金、模拟配置及模拟结果均须明确标识，与真实交易及持仓记录分开管理。

9-03 提出的邀请，也是后续模拟体验的方向：

您的订单，您的组合，您的 SHARP。

Your Order. Your Portfolio. Your SHARP.

八、帽子军团如何接受监督与检查

牛码新市主动拥抱监管、服从监管，并为监管检查提供可核查的规则、系统与运行记录。本书从管人、查系统、看标准三个方面说明如何检查帽子军团。管人，核查每个岗位的具名认领、授权，以及相关规则的制定、修改、废止记录。查系统，核查任务输入、规则版本、执行、检验和异常处理记录，并按实际建设情况说明检验接口、过程回放能力及其限制。看标准，核查“十二个足够”是否有证据支持，条件不足时是否警示并重新校准。人管规则与例外，系统管执行与留痕，独立检验检查产出与过程是否符合要求。

人在监管要求下承担责任，以规则和授权管理系统，系统依规组织执行并保留记录。这套分工将系统治理用于 AI 工厂，人的工作集中在规则、授权、监督和例外处理；依法或依制度需要的专业确认和人工裁量仍由具备相应职责的人承担。

九、牛码新市与持牌机构的分工

投资者通过具备相应许可的持牌机构完成交易。投资决定由具有相应决策权限的主体作出；投资者适当性及相关服务、合规义务，由承担相应职能的持牌机构在许可范围内履行。交易与交收完成后，实际持仓关系记入组合登记簿，并关联原始成交记录；合同登记簿记录合同本身及其状态，罗盘读数依据新增事实与证据更新，罗盘终端供有资格和授权的持牌机构/境外合格投资者使用。

牛码新市的职责是登记、度量和提供使用这些资料的基础设施，不代投资者选择资产或配置持仓。合同登记以经采集和核验的合同资料为依据，组合登记以已完成的交易、交收及后续持仓变动为依据，相关规则和记录保留版本与依据。牛码新市不是证券公司，不撮合交易、不挂牌证券、不设客户资金池、不接触客户资产。牛码新市对自身系统治理的健全性与基础设施的可靠性负责：数据错误须修正并保留历史，系统异常须查明影响、完成修复并复验，

因子定义与规则问题须经治理程序修订。不承诺投资结果，并不免除牛码新市在登记、数据、系统和服务方面应承担的责任。

公众可阅读一般性行业研究资料。正式投资参与仅面向符合适用法域要求的持牌机构/境外合格投资者，经具备相应许可的持牌中介进行，不向零售投资者开放，不面向中国境内募资，也不接受境内资金参与投资。

十、如何阅读，如何参与完善

本书正文沿业务难题、罗盘工具与规则、帽子军团理论、AI 工作手册、案例成果、未来展望六部分展开。一般读者从第一、二部分了解问题和仪器，第三部分了解帽子军团的组织方法；第四部分工作手册供 AI 岗位背后的具名真人查阅履职要求，供技术尽调方和 AI 审计者检查组织与运行安排。经营者、行业专家和投资研究者可从第五部分及附录一入手；后续验证安排见第六部分。

信任没有存折，只有日结。

牛码新市公开方法、成果和边界，是为了让市场能更好地提出具体问题。

因子是否抓住影响生意的关键因素，数据是否能够取得并核验，规则是否存在刷分空间，遇到反例又应改哪里，都可以逐项讨论。每一轮有依据的修订，都应让这套方法更接近真实经营；行业覆盖仍不能代替具体项目核验与持续观察。

地图逐步展开，罗盘继续校准，军团已上岗。

牛码新市邀请企业经营者、行业专家、金融机构与 AI 技术伙伴参与共建，并欢迎监管部门提出指导意见，让这套方法在实践与检验中不断完善。

行业罗盘成果入口

从正文 3 例精读案例与附录 17 例品类案例出发，进一步浏览 500 个行业的罗盘成果

读者扫描下方二维码，按提示登记信息并完成验证后，可按行业查看指标与因子结构、已开放的规则资料及版本说明。各行业当前开放内容以正式披露为准。

说明：仅供在线浏览，不提供下载。



行业罗盘网址：<https://microconnect.com/industry-compass>

版本关系与使用说明

版本关系：本白皮书为牛码新市于 2026 年 10 月 3 日发布的第三篇白皮书（下称“本版”），与 8 月 3 日发布的第一篇、9 月 3 日发布的第二篇相互衔接、各有侧重，共同构成牛码新市的公开框架说明。涉及同一事项的明确更新，以本篇相关说明为准。

成果口径：本白皮书介绍“帽子军团”的理论框架、牛码新市的初步组织实践及本轮行业罗盘成果。相关方法与规则将结合实际运行、企业数据及各方反馈持续检验和完善。本版所称“5000 名 AI 员工已经上岗”和“500 个行业覆盖”，分别为岗位纳入组织台账和行业规则成果的口径，不等同于具体任务运行记录或具体企业验证；连锁便利店任务链为说明性模拟。

使用前提：本书及配套行业罗盘、牛码调整系数仅供一般性行业研究、经营分析与方法讨论，不构成投资建议、推荐、任何交易要约、产品发售或投资招揽，亦不构成收益承诺、法律意见或合规批准。相关分值、系数及排序不代表投资回报率、违约概率、信用等级或投资优先顺序。本版展示的行业罗盘案例主要为指标与因子结构，不代表已完成具体企业的取数、评分或验证；说明性模拟不代表实际运行成绩。市场参与者应基于自身的独立判断、尽职调查及专业顾问意见作出决策。AI 仅在授权范围内整理证据、执行检查并形成建议；最终评分及投资、适当性、合规或法律判断，须由具备相应授权与资质的人作出或确认。涉及依法须持牌开展的服务，应由具备相应许可的主体在许可范围内提供；内部授权或人工确认不能替代该等许可要求。

参与范围：本白皮书及配套行业罗盘向公众提供一般性研究信息。公开阅读不构成向读者提供投资服务或确认其投资参与资格。牛码新市的投资参与仅面向符合适用法域要求的专业投资者，并须通过具备相应许可的中介机构进行，不向零售投资者开放。牛码新市的资产端面向全球，资金端面向中国境外符合资格的专业投资者。牛码新市不面向中国境内开展任何募资活动，不接受境内任何资金参与投资。

目录

第一部分 业务难题	14
第一章 从公司尺度走向合约尺度	14
1.1 公司证券化的贡献与处理尺度	14
1.2 合约是理解经营的更小单元	14
1.3 “大财小用”的理想与“大才不能小用”的成本困境	15
1.4 成本下降之后，证据与责任仍须成立	15
第二部分 罗盘工具与规则	17
第二章 登记与度量仪器的用途和分工	17
2.1 合同终端	17
2.2 合同登记簿	17
2.3 合同罗盘	18
2.4 罗盘读数	18
2.5 组合登记簿	18
2.6 罗盘终端	19
2.7 从仪器功能回到本版成果	19
第三章 行业罗盘怎样把生意讲清楚	20
3.1 三层结构	20
3.2 数量与选取规则	20
3.3 评分规则	21
第四章 500 个行业罗盘交付了什么	22
4.1 成果覆盖范围与五百个行业索引	22
4.2 因子定义、单位、观察期间与评分标准	24
4.3 成果版本、证据记录与后续校准	24
4.4 怎样找到本行业成果，并参与完善？	25
第五章 行业成果怎样进入牛码新市	27
5.1 通用罗盘看合同结构，行业罗盘看具体生意	27
5.2 通用码与行业码：各自的收益评分与风险评分	27
5.3 牛码调整系数与跨行业校准	28
5.4 从行业内判断到跨行业比较的阅读路径	30
5.5 使用边界：行业成果、项目核验与持续观察	30
第三部分 帽子军团理论	33
第六章 AI 能力与人类责任的两组矛盾	33
6.1 放开 AI 与盯住 AI 的两难	33
6.2 金融行业的必答题：合规、投资者保护与风险控制	34
6.3 帽子军团怎样同时解开两组矛盾？	35
第七章 HAT 帽子：人掌握规则，AI 依规执行	37
7.1 HAT 的含义：人的最终责任与 AI 的潜力释放	37
7.2 规则先于能力：允许、禁止与升级条件	37
7.3 每顶帽子背后的具名真人：认领与规则制定、修改和废止	38
7.4 帽子只管边界，不管动作	38
7.5 人的治理、监督与自由裁量责任	39

第八章 MASS：强能力怎样进入微任务	41
8.1 MASS 的含义：微任务智能体、简单且短时的任务与海量协作	41
8.2 任务小，兵不小：完整模型能力与有限任务权限	42
8.3 任务拆分三个足够：足够小、足够窄、足够碎，保留完整业务意义	42
8.4 先可验证，再上规模：独立验收、错误定位与局部返工	43
8.5 底座支撑三个足够：多元、对冲与性价比	44
第九章 微任务怎样组成可靠的协作	46
9.1 “帽对兵”与“兵对兵”：规则治理与任务协作两条通道	46
9.2 从拆分到组装：任务依赖、交接接口与成果衔接	47
9.3 从分歧到处理：澄清、复核、退回与人工升级	48
9.4 分项合格与整体交付：调度、验收与回测的分权制衡	48
9.5 每个 Agent 由具名真人认领：临时协作与持续责任	49
9.6 编制扩充三个足够：足够多、足够细、足够冲及其承载边界	50
第十章 怎样让每次交付经得起检验	52
10.1 全检而非抽检：每项产出均接受检验	52
10.2 持续回测：识别短期运气与环境变化	53
10.3 全程溯源：从结果回到输入、调用、规则与责任	53
10.4 检验的三个足够：足够密、足够长、足够深	54
第四部分 AI 工作手册	57
第十一章 5000 名 AI 员工怎样进入组织〔操作程序〕	57
11.1 已经上岗的 5000 名 AI 员工	57
11.2 三层设计：规则层、职责层与编号层	58
11.3 四大职能：调度管理、业务执行、管控合规、部署支撑	59
11.4 多视角席位：3 类行业视角 + 3 个模型对冲、3 个投资者视角与“6 + 1”检验视角	60
11.5 视角与能力分开：模型、技能和工具按任务调用	61
第十二章 一个岗位、一次运行，怎样分清责任〔操作程序〕	63
12.1 AI 员工的岗位与任务实例	63
12.2 AI 员工的身份编码、能力表现与运行编码	64
12.3 AI 员工的岗位台账、状态更新与实例信息分层管理	65
12.4 AI 员工运行编码关联的记录与证据留存	67
第十三章 行业罗盘任务怎样拆解与组队〔操作程序〕	68
13.1 行业罗盘任务的定义：输入资料、交付标准与约束	68
13.2 行业罗盘任务的拆解：因子识别、口径说明与评分规则工作单元	69
13.3 AI 员工配置：按能力表现、权限与可用状态调配岗位	70
13.4 AI 员工组队：专项小队的组建、分歧处理与协作解除	70
13.5 AI 员工七站作业流程：接单、规则匹配、组队、执行、验证、收口与归档	71
13.6 举例：连锁便利店行业罗盘——从一张任务单到一个候选	72
第十四章 长期岗位怎样支持动态协作〔操作程序〕	75
14.1 部门归属线与任务流程线	75
14.2 临时小队覆盖四类职能，长期责任记录保持连续	76
14.3 调度 Agent 负责日常组队、衔接与整体交付协调	77
14.4 具名真人负责规则 授权与超出规则的升级处理	78
14.5 整体交付与分项责任并存，验收和管控保持独立	79
14.6 举例：连锁便利店行业罗盘——分歧出现时，谁澄清、谁修改、谁拍板	79
第十五章 一项交付怎样经过检验与修订〔操作程序〕	82

15.1 验收是任务流转的准入关口	82
15.2 “6+1”检验视角：六鉴与执衡的分工、检验产出及不合格退回	83
15.3 沿运行记录核查输入、产出、规则与责任	85
15.4 事实回测推动能力校准与规则修订	86
15.5 授权核查、版本管理与记录完整性	87
15.6 举例：连锁便利店行业罗盘——V0.2 经过六审，为什么列为“待澄清”	87
第五部分 案例成果	90
第十六章 三个品类精读与案例阅读方法	90
16.1 三个品类精读	90
16.2 十七个案例的选择与阅读方法	94
第六部分 未来展望	95
第十七章 从本轮成果走向下一次交付	95
17.1 本轮的理论、实践与成果	95
17.2 企业取数、校准、模拟与进展披露	95
17.3 模拟环境中的连续运行	96
17.4 行业反馈与共同检验	97
附录一 17 个补充案例与 500 行业成果导航	116
附录二 十二个足够总表	117
一、十二个足够总表	117
二、总表的使用边界	118
附录三 名词释义与中英文对照	119

第一部分 业务难题

第一章 从公司尺度走向合约尺度

1.1 公司证券化的贡献与处理尺度

公司证券化让不同规模的资本能够围绕同一家公司形成投资关系。公司把经营活动组织起来，财务报表、信息披露和专业服务把分散的信息汇成可以检查和比较的对象。这套制度的贡献，是让资本不必亲自经营每一门生意，也能在明确的权利和责任安排下参与经济活动。

但把公司作为主要观察单元，也形成了相应的专业服务尺度。理解一家公司，需要核验资料、分析经营、识别风险并持续跟踪，这些工作有固定成本。规模较大、资料较完整的公司更有条件承担成本，而数量庞大、分布零散的小微生意，则很难获得同等密度的专业理解。制度可以服务更广的经济，专业工作的成本和处理能力却未必跟得上。

1.2 合约是理解经营的更小单元

许多经营活动最终落到具体合约：谁向谁提供什么，满足什么条件后支付，现金流在什么时间、沿什么路径形成。公司的整体表现重要，单份合约的权利基础、履约条件与经营证据也有独立的观察价值。把尺度下移到合约，有助于看见总量数据中容易被掩盖的差异。

合约约定的现金流，并不等于已经兑现或必然兑现的现金流。真实经营会变化，付款会延迟，信息也可能不完整。因此，粒度越细，越需要把合同条款、经营事实、证据来源和观察期间分别记录，才能判断哪些内容已有事实和证据支持，哪些仍然只是条件、预测或待核验事项。

过去的难处，并不是专业人士不懂小生意，而是逐份理解、核验和持续观察的成本，可能超过一笔小额业务能够承受的范围。把服务尺度从公司细化到一份份合约，不能只把原来的工作复制更多遍；必须同时降低重复处理的成本，让每项判断都能对应到具体证据和责任主体。

1.3 “大财小用”的理想与“大才不能小用”的成本困境

在小微生意中，这道成本题尤其具体。我们想做的“大财小用”，是把大财富拆成更小的资金单元，投向小工厂、小商店、小农场等传统金融较难覆盖的实体经济“毛细血管”中。“小用”改变的是资金进入真实经营的尺度，并不降低对证据、风险判断和责任的要求；我们希望更多规模小、分布散的经营主体，也有机会被认真理解、被真正服务。

钱拆小了，判断并不会跟着变简单。一家小店同样有收入、成本、履约和现金回收，一家小工厂同样要看订单、产能、库存和经营风险。每一个小额、分散、非标准化的任务，仍可能需要资料核验、经营分析、定价判断和持续监测。服务对象越多、行业差异越大，需要处理的信息就越细，需要分别作出的判断也越多。

传统专业服务依靠经验丰富的人才逐笔处理复杂事项，其中涉及专业经验与责任的判断，仍需由具备相应能力和授权的人承担。难题在于，如果每一笔小额业务都投入同等规模的高级人才时间，人才成本就会与任务尺度错位。专业人才当然做得了小任务，只是高成本、单线程逐项的工作方式，覆盖不了海量而分散的需求。单笔可以做得很认真，整体却可能算不过账。这就是“大才不能小用”的困境。

AI 让我们看到另一种可能：将复杂工作拆成更细的任务，把强大的分析和处理能力用于每一个任务。资料整理、事实比对、规则匹配、结构化检查，都可以拆成边界更清楚的单元，由机器按照规则反复执行，并在满足质量要求的前提下降低处理成本。过去只有大额业务负担得起的部分能力，因而有机会进入数量更多、规模更小的业务。

但把任务交给 AI，并不意味着成本、规模和责任问题已经解决。成本是否下降，还要计入检验、返工与运行支撑的投入；责任是否明确，也要有规则、授权和记录作为依据。如果一次交给 AI 的任务过大，或允许它在模糊边界里自由判断，即使生成结果更快、更便宜，错误却可能更难定位，责任也更难落位。“大财小用”不能只靠更快、更便宜的自动化，它需要一种组织方法，把强能力放进可管理、可检验、必要时能升级给人的任务里。让能力强大的 AI 只承担边界清楚的微任务，正是“大材小用”的含义，也是回应“大才不能小用”成本困境的一条路径。

1.4 成本下降之后，证据与责任仍须成立

AI 有望降低部分专业工作的处理成本。能够预先说清输入、规则、输出和检验标准的工作，可以在授权范围内交给机器反复执行；事实核验、信息结构化和持续比对由此有机会扩展到

更多经营对象。机器能够处理更多任务，不等于机器取得发行、交易、投资判断或法律裁量的权限。AI 可按规则整理证据、执行检查并形成建议，需要专业确认或批准的事项，仍由具备相应授权与资质的主体承担。

从业务难题走向工具，先要回答怎样把分散的合同事实记录清楚，怎样按一致口径度量，怎样让新事实能够持续进入读数。本书把承担这些功能的工具比作“仪器”，把制造和维护这些仪器的组织比作“AI 工厂”。第二部分说明这些仪器的用途和规则；第三部分再解释这座 AI 工厂为何采用“帽子军团”的组织方法。把问题、工具和组织分开讲，才能看清每一层究竟解决什么。

第二部分 罗盘工具与规则

第二章 登记与度量仪器的用途和分工

牛码新市是合同证券化市场的基础设施提供方，承担登记与度量功能。我们提供的仪器供持牌机构与金融专业人士使用，把合同条款、经营证据和适用规则连接起来，让使用者能够查到事实、复核读数并形成自己的专业判断。帽子军团是制造和维护这些仪器的组织，不替代投资决策主体与持牌机构履行各自的职责。

8·03 白皮书提出合同现金流投融资的市场框架，9·03 白皮书提出双罗盘的度量规则，并展开登记、观察和组合记录等相关功能。本书沿一张合同从进场到存续的路径，把这些功能归纳为合同终端、合同登记簿、合同罗盘、罗盘读数、组合登记簿和罗盘终端六件仪器。这里说明的是功能与分工，并非固定的操作顺序，六件仪器的建设状态、服务对象及开放范围，须按各自的正式版本与实际披露分别判断。

2.1 合同终端

合同终端是合同进场的第一站，是与发行人或其 AI Agent 持续对话的交互式系统。它按要求收集材料、逐项录入和核验，对缺失、不一致或不符合要求的内容提出补充与澄清要求，将非标准化合同整理为带有来源、时点和核验状态的结构化记录。缺失事实不能靠 AI 推断补齐，材料采集合格也不等于合同已经获准发行、挂牌或交易。

合同终端持续与发行人及承担核验职责的专业人员交互，核对合同是否符合适用要求、已经生效、正在真实履行，现金流是否可由第三方核验。AI 按规则采集和检查；合同解释、法律定性及规则未覆盖的判断，仍交由具备相应权限和资质的人处理。满足准入条件后，是否适合投资，仍须由具有相应决策权限的主体结合具体证据独立判断。

2.2 合同登记簿

合同登记簿记录合同的身份、条款、相关主体、初始画像、运行状态及底层证据。输入是已按相应规则采集和核验的合同信息及后续更新资料，输出是可查询、可比较、可追踪、可回到来源的标准化记录。它保存变更依据和历史版本，使使用者能够区别初始条件、后续变更和当前状态。

合同登记簿承担信息记录功能，登记本身不构成发行，不创设合同权益，不改变原合同的权利义务，也不构成对合同履约的保证。登记状态和证据状态应分别呈现，不能把“已登记”理解成“已经验证全部事实”。

2.3 合同罗盘

合同罗盘为登记在册的合同提供度量框架。通用罗盘看合同结构，行业罗盘看合同背后的具体生意；两者均设置收益维度与风险维度，收益与风险分别呈现、不相加、不互相抵消。输入是具有明确对象、来源和期间的事实，以及适用的因子定义、权重和评分规则；输出是保留证据、版本与使用条件的读数。

罗盘是度量，不是信用评级或投资推荐。它不输出“该买”或“该卖”，相关分值不代表投资回报率、违约概率或信用等级。同一组数据按同一规则版本及相同参数计算，应得到一致的数值和评分档位；证据不足的地方留白。罗盘指北，不指路。

2.4 罗盘读数

罗盘读数呈现已登记合同的因子数据及相应评分，按日检查，并依据数据性质和新增证据更新，将经营、履约及证据变化持续带入观察。输入是新取得的事实、合同事件及相应证据，输出是带有数据时点、观察期间和规则版本的更新记录。更新应保留前后关系，不能用新数据覆盖历史结果所依赖的旧条件。

罗盘读数提供观察依据，不是投资行动建议。每日检查并不意味着每日改写数值；各字段仍按其性质、数据来源和实际变化更新。数据缺失、来源冲突和更新延迟须明确标示，不能把旧读数当作新事实。

2.5 组合登记簿

组合登记簿在投资者通过持牌机构完成交易和交收后，逐日记录各组合持有哪些合同、持仓数量及其变化。在 9·03 白皮书中，成交、持仓、分派和退出事实由订单登记簿记录；组合登记簿在组合层关联这些事实，不取代或改写原始交易记录。

输入是有效的成交、交收和持仓记录，输出是可回到相应交易与底层合同的组合关系。投前扫描与模拟配置应与实际持仓分开；草稿不是已经成交的组合。对组合及其持仓的观察和计算按明确规则进行，平台不替使用者挑选资产；组合登记簿只记关系与事实，不承担客户资产托管或资金划转职能，登记记录不等于保管资金。

2.6 罗盘终端

罗盘终端是供持牌机构/境外合格投资者使用的专业查询与分析界面，连接登记事实、罗盘规则、读数与证据。输入是获授权访问的资料及使用者设定的观察条件，输出是可展开、可比较、可回查的专业视图。使用者按自身方法论调用读数、形成判断；个人参数调整仅作用于其自身观察视图，不改写原始登记事实、默认规则或其他使用者的结果。

终端的服务对象是具备相应资格与授权的持牌机构/境外合格投资者。公众通过行业罗盘网站阅读一般性研究资料，与取得专业终端权限是两种安排。公开阅读不意味着取得投资参与资格，提供查询与分析界面也不意味着牛码新市取得代客决策或交易的权限。具体功能和开放安排以对应服务说明为准。

2.7 从仪器功能回到本版成果

六件仪器把采集、登记、度量、持续更新、组合记录和专业使用连接起来，系统梳理合同底层数据，使其可查询、可理解、可复核。它们的可靠性，取决于事实来源、规则版本、授权边界和运行记录能否相互核对。投资决定由具有相应决策权限的主体作出；投资者适当性、交易执行及相关服务义务，由承担相应职能的持牌机构在许可范围内履行；基础设施提供方仍对自身实际承担的登记、数据和系统职责负责。

本版重点公开行业罗盘的结构、规则成果及制造这些成果的组织方法。覆盖 500 个行业的规则成果，不代表已完成相应行业的企业取数、评分或验证；正文 3 例与附录 17 例，合计 20 例，展示到指标与因子结构。实际使用须核对适用的正式版本及数据条件。以下三章依次说明行业罗盘、成果版本与双罗盘的使用条件。

第三章 行业罗盘怎样把生意讲清楚

行业罗盘是牛码 SHARP 双罗盘中针对某具体行业或经营品类的专属评分规则，与全市场统一的通用罗盘并列：通用罗盘看合同结构，行业罗盘看具体生意，回答该品类哪些可量度的行业事实支撑收益、哪些反映风险，以及应依据什么行业尺度进行评分。规则须经设计审查、对抗性检验与相应的人工复核后，按正式版本公开发布。

3.1 三层结构

行业罗盘采用“维度—指标—因子”三层结构（表 3-1）。维度固定为收益与风险两个方向，分别评分、分别呈现，收益支撑不抵消风险暴露。指标规定每个维度“问什么”，因子是可取数、可核验的经营变量，回答“具体看什么”。

层次	定义	数量规则
维度	收益维度、风险维度两个观察方向	所有品类固定两个维度
指标	维度下的评价项目，规定“问什么”	每个维度对应 3 至 9 个指标，指标内容按品类确定
因子	支持指标判断的可核验经营变量	每个维度对应 3 至 9 个因子；每个指标至少对应 1 个因子

表 3-1：行业罗盘的三层结构

3.2 数量与选取规则

指标与因子的内容和数量均可变。在一个品类的一个维度内，指标为 3 至 9 个，因子为 3 至 9 个；每个指标至少对应 1 个因子，允许只对应 1 个。各品类依据业务复杂程度与标准化程度，选取关键性的指标和因子，形成适合该品类的行业罗盘框架，具体企业的评估还须取得相应数据，并按适用规则评分。

行业因子选取遵循四项约束：每个因子只归入一个指标，同一信息只计分一次；因子须有明确的经营含义和可核验的数据来源，如来源于无法复核的主观判断则不得纳入；指标与因子构成、默认权重随规则版本公开，变更须经治理程序；投资者可在其个人界面中按允许的范围调整因子选取及因子权重，调整仅作用于个人观察视图，不改写默认规则、原始数据或其他使用者的结果。

收益维度常用因子包括需求与收入实现、运营效率与单位产出、收入质量与持续性方面的因子；风险维度常用因子包括成本与执行风险、集中度与依赖性风险、周期与合规冲击风险方面的因子。以上为常用选项，不构成固定清单。

3.3 评分规则

每个因子须明确数据口径与来源、品类刻度与锚点，以及 1 至 3 分的落档条件。两个维度评分方向一致：收益维度因子收益支撑越强，得分越高；风险维度因子风险越高，得分越低。两个维度均以高分表示较好的表现。计算牛码 SHARP 时，风险维度得分须按统一规则转换为相应的风险度量，再与收益侧读数结合；不能直接把风险维度得分作为风险大小使用。本版说明评分方向与转换原则，不展开具体转换公式。当前启动期的 2 分为参考点或估计锚点，经充分样本和独立验证后方可表述为品类中位数。

评分以证据为前提。未提供证据或证据不足以支持落档的因子，应留白并注明原因，不以默认分值补齐。已有评分的证据可信度较弱时，可标灰提示待复核状态，并保留原评分及其依据。

因子评分按公开权重形成指标得分，再汇总为收益维度与风险维度得分，并归一到统一分值域，使指标和因子数量不同的品类可在同一尺度上呈现；维度得分经牛码调整系数完成跨品类校准后形成行业 SHARP，并与通用罗盘共同形成单合同牛码 SHARP。行业罗盘的表现不能抵销通用罗盘的最低准入标准。牛码调整系数的方法见第五章 5.3 节，跨行业比较的条件与使用边界见 5.4、5.5 节。

第四章 500 个行业罗盘交付了什么

方法讲清之后，要打开交付物来看。500 个行业罗盘覆盖哪些经营活动，每套规则包含什么，读者怎样核对定义、单位、期间和分档，又怎样找到它的来源与版本，本章逐项展开。

一份行业成果应让读者从行业名称走到具体规则，看见收益与风险各由什么事实支持。共同结构负责让内容好找，因子负责把生意讲准。覆盖数量说明工作做到了多大范围，明确规则、版本和适用边界，才说明成果怎样被使用。

图 4-1 概括了多方视角汇成行业罗盘成果的路径。



图 4-1：多方视角汇成行业罗盘成果（图中 20 例为正文 3 例与附录 17 例合计）

4.1 成果覆盖范围与五百个行业索引

本轮研究覆盖 500 个行业，形成行业罗盘规则成果。本版选取 20 个品类案例，其中连锁便利店、川渝火锅门店和现制茶饮门店作为精读案例在正文详细展开，其余 17 个品类案例收录于附录一。各行业的经营活动范围、量化口径、数据路径、权重、分档及审查情况，按成果版本披露。覆盖数量不代表已完成对 500 家企业或 500 份合同的评分与验证。

层次	本版呈现	判断边界
组织配置	5000 名 AI 员工已经上岗，四类职能与稳定岗位、当次执行的分层安排	岗位已纳入组织台账，不等于 5000 名 AI 员工同时执行任务；岗位数量须有台账与统计口径支持，实际运行另看任务记录
行业规则	覆盖 500 个行业的规则成果；正文 3 例与附录 17 例合计展示 20 个品类的指标与因子结构	行业覆盖不代表已完成相应行业的企业实测评分或验证
过程说明	连锁便利店 SIM-01 贯穿任务、协作、检验与修订	说明性模拟，非后台真实运行记录
企业应用与校准	给出取数、评分、回测和调整的使用条件	企业实测、系数版本、批准与验证记录按实际证据另行披露

表 4-1：本轮成果与验证状态

先找行业，再确认是否适用。名称相近的生意，收入、成本和风险来源可能不同；同一家企业，也可能经营多项业务。索引因此需要把编号、标准名称、别名、范围与版本连起来，帮助使用者找到真正匹配的那套规则。

一个可供使用的行业索引，需要回答下表这些问题。

索引字段	回答的问题	阅读作用
行业编号	这套成果如何被稳定识别	避免名称变化后失去历史对应关系
标准名称与必要别名	市场通常怎样称呼这门生意	帮助使用者从常用名称进入正确条目
适用范围	规则针对哪种经营模式	防止把名称相近但经济结构不同的业务混用
规则版本与状态	当前看到的是哪一版、处于什么状态	区分历史稿、现行版本及使用限制
更新时间与生效时间	内容何时修改、从何时适用	支持历史结果复核和前后比较
成果内容	包含哪些指标、因子和评分口径	让读者从索引进入具体规则，而非停留在行业清单

表 4-2：500 行业索引的字段与阅读作用

查到行业后，应能继续打开收益维度与风险维度，看到因子、定义、数据口径、单位、期间和评分标准，再返回版本与修改记录。尚未确定或尚未开放的内容，应标明状态。行业拆分、合并或更名，也要保留原编号、继承关系和适用日期，让历史引用仍找得到当时的规则。

行业规则写的是怎样判断，具体读数则要绑定经营对象、事实证据、观察时点和规则版本。500 个行业都有成果，不表示每家企业已有行业码，也不表示各行业经历了同样多的样本和回测。索引应把审查与验证状态如实标出，不能让统一外观掩盖不同状态。

4.2 因子定义、单位、观察期间与评分标准

每个因子完整规则应明确定义、量化口径、数据来源、单位、观察期间与 1 至 3 分标准等要素构成（表 4-3）。同一名称在不同口径或期间下含义不同，评分前须核对所用口径与期间是否符合适用规则。

评分方向与证据要求按第三章 3.3 节执行。正文 3 例与附录一 17 例合计展示 20 个品类样本的指标与因子结构；各因子的量化口径与分档标准在相应品类的正式规则中确定。

成果字段	读者需要核对的问题	常见误读
因子名称与定义	观察对象究竟是什么，包括和排除哪些情形	只凭名称推断计算方式
所属维度及指标	该因子回答收益维度或风险维度中的哪个经营问题	同一事实在多个位置重复计分
量化口径	分子、分母、范围和异常值怎样处理	把相近公式当成同一口径
数据来源	数值来自哪个系统、平台、单据或其他证据	使用无法复核的主观印象代替事实
单位	采用金额、比例、次数、天数还是其他单位	数值相同便认为含义相同
观察期间	是月度、季度、年度、滚动期、旺季还是时点	把不同期间的结果直接比较
评分方向与 1 至 3 分标准	什么表现进入哪个档位，边界如何归属	只凭数值高低判断好坏
规则版本	该口径在何时、对什么对象适用	用现行规则解释历史结果

表 4-3：因子成果的核对字段

4.3 成果版本、证据记录与后续校准

生意会变，因子定义、观察期和阈值也可能需要调整。要让过去的结论仍可解释，每份成果就应留下行业编号、规则版本、状态、更新时间、生效时间和证据范围。

行业编号用来找对象，版本号用来区分同一对象前后的规则。内容何时修改，与何时生效也要分别写清。新版本可以按照其适应范围和生效时间接续使用，旧版本仍须保留，否则历史评分就无法回到当时的依据。

“已形成”“已复核”“已批准”“已生效”和“可用于具体评分”，各有不同含义。表格完整，只能证明内容已经写出。依据推导形成、尚待确认或限制使用的字段，应保留相应状态，并说明依据与限制。

复核一项因子，要能问到规则依靠什么行业事实、口径从何而来、取数需要什么证据、区间为何这样划分，以及修改意见如何处理。版本号帮助定位，实际材料决定能查多深，不能由编号推定研究与审批过程全部完整。

后续经营、履约、逾期、违约、回收和市场使用会带来新信息，也会带来反例。因子解释不了变化、数据取不到、阈值能被规避、范围定得过宽，都应进入复核。修订要说明改动、原因、依据、生效时间与影响，不能只在结果端换一个分数。因子定义、用途或评分规则发生实质变化时，须形成新版本并重新接受完整六审，再由执衡更新资格与组合。

覆盖 500 个行业，代表第一轮规则范围已经建立，同时后续将持续校准，检验这些规则在实际应用中是否适用、是否需要修订。初始成果给市场一个具体讨论对象，新版本把反馈变成更准确的定义和边界。行业覆盖说明规则研究的范围，校准与验证说明规则在具体条件下的表现，两者应分别记录和披露。

4.4 怎样找到本行业成果，并参与完善？

本书正文 3 例与附录 17 例，合计 20 个展示案例，让读者看见一套行业规则怎样组织；统一成果入口则面向 500 个行业，帮助读者找到本行业资料 and 对应版本。本版案例序号仅用于书内检索，不作为行业编号；各案例与 500 行业目录的对应关系，将在完整成果索引中列明。这 20 例按经营场景选取，不代表统计抽样，也不能代替全量成果。

层级	展示内容	读者可以做什么
全书 20 例（正文 3 例 + 附录 17 例）	案例序号、品类简介、维度、指标与因子	按第十六章 16.1 节与附录索引阅读；提出具体字段的修改建议
500 行业统一索引	按行业查找；列行业编号、成果版本、可获取状态	找到本行业，判断对应成果当前开放到哪一步
具体行业页面	已开放获取的罗盘、因子规则、取数清单及版本说明	获取资料，反馈判断，订阅更新或预约交流

表 4-4：500 行业成果导航与获取安排

行业罗盘成果入口为：<https://microconnect.com/industry-compass>。二维码与文字链接指向同一页面。首次进入时需邮箱或微信一种联系方式；身份可选填“经营者 / 投资研究 / 行业专家 / 其他”。

动作	读者获得什么	后续连接
在线查看本行业成果	已开放的指标与因子结构、规则资料及版本说明，仅供在线浏览	登记并完成验证后进入相应行业页面；更新订阅与交流预约另行选择
订阅本行业更新	后续版本变更与解读通知	按所选行业接收更新，了解修订内容与依据
预约行业罗盘交流	讨论如何用于自身经营或投资研究	围绕明确应用需求推进交流与合作

表 4-5：三个参与动作分别选择

第五章 行业成果怎样进入牛码新市

500 个行业成果让不同生意有了各自的表达。接下来还有一个问题：对便利店、制造企业和养殖企业分别得出的判断，怎样放到一起理解？它们收入形成方式、投入结构和周期不同，仅有行业语言，难以比较；只有通用尺度，又容易丢掉影响经营的细节。

牛码 SHARP 双罗盘同时保留这两层。通用罗盘看合同结构，行业罗盘看具体生意，两者各有收益维度与风险维度。通用码和行业码记录读数，牛码调整系数为行业侧的跨品类比较提供环境参照。使用时，评分尺度、经营事实、观察期和证据依据一起保留，比较才有依据。

牛码 SHARP 双罗盘是本书的双罗盘体系；牛码 SHARP 是用于表达和比较合同现金流风险收益结构的中心指标。阅读具体结果时，仍须区分指标框架、原始分数与调整后的读数，并核对适用版本。

5.1 通用罗盘看合同结构，行业罗盘看具体生意

无论哪门生意，一份合同都要说明收益从哪里来，合同安排、相关主体与履约机制怎样支持兑现，又带来什么结构性风险。通用罗盘把这些跨行业反复出现的问题放在统一框架中，让不同合同能用共同语言被阅读。

行业罗盘继续往经营里走：需求怎样转成收入，投入怎样形成产出，结果能否持续，成本、执行、依赖与外部变化又会怎样影响它。便利店要看门店的实际条件，其他行业要看生产单位、周期与数据。共同问题之下，必须保留真正能解释本行业的事实。

合同结构清楚，不等于经营质量已经说明；生意表现良好，也不能代替合同和履约结构的检查。两个罗盘并列，让使用者既看共性，也看具体差异，避免一侧的结论提前覆盖另一侧。

每个罗盘内部都有收益维度与风险维度。通用罗盘的两侧看合同结构，行业罗盘的两侧看实际经营，共有四个应分别理解的观察区域。合同安排可能较稳健，行业风险仍然很高；经营表现可能不错，合同却还有核验缺口。这些信息都应留下。

5.2 通用码与行业码：各自的收益评分与风险评分

罗盘提供观察框架，码记录按相应框架形成的读数。通用码对应合同结构的收益与风险，行业码对应具体生意的收益与风险。行业编号另有用途：它让我们在 500 个行业中找到规则。

通用码让不同合同进入共同尺度，回答合同本身有什么收益支撑、承受什么风险负载。它保留跨行业的结构问题，具体生意在本行业中处于什么位置，还需要行业码继续解释。

行业码把经营差异带回来。便利店的门店条件怎样影响需求、效率、收入质量与风险，要由相应因子说明；换一门生意，同样问题可能需要完全不同的证据。每项读数都应带着定义、单位、期间和评分规则，才知道它究竟表达什么。

两类码都分别呈现收益和风险。较强的收益支撑不会自动消除风险，风险较低也不保证收益较高。分数看起来都是数字，承担的含义却不同，不能直接相加、抵消或混成一个方向。

也不要将通用码和行业码压成一个找不到来源的任意总分。按既定规则形成单合同牛码 SHARP 时，仍须保留所属罗盘、收益维度或风险维度、规则版本和观察期间，使读者能够回到指标、因子与证据，知道结论为何成立。

5.3 牛码调整系数与跨行业校准

行业罗盘先判断一门生意在本行业内表现如何。两门生意即使在各自行业得分相近，所处的行业环境也可能不同。把分数直接并列，容易把“在本行业表现怎样”误读为“放在所有行业中处于什么位置”。牛码调整系数（NUMA Adjustment Factor）补的是行业环境这一层参照：它不改行业罗盘的因子和评分规则；待具体企业完成行业内评分后，再结合行业景气状况观察行业环境差异。企业得分能否用于跨行业比较，还须检验评分尺度与分值含义的可比性。

牛码调整系数不限于单一数据来源。本版以中国国家统计局（National Bureau of Statistics of China）相关数据为主要参照，后续将逐步拓展可采用的数据来源，以适应不同行业、国家和地区的经营环境与数据条件。牛码新市也欢迎市场参与者结合自身研究和使用场景，开发适用的调整系数，并说明数据来源、计算方法、适用范围与版本。不同方案可供使用者按需选择；涉及跨行业或跨地区比较时，仍须核对口径与可比性。

本版主要使用中国国家统计局发布的企业景气指数¹。企业景气调查按季度开展，指数汇集经营者对当前经营状况和未来趋势的判断。按照中国国家统计局的指标解释与编制方法，指数在 0 至 200 之间，以 100 为景气临界值。它反映企业群体对经营环境的感受和预期，不直接代表行业收入、利润或具体项目的回报。

¹国家统计局“企业景气指数”官方说明：<https://www.stats.gov.cn/zs/tjws/zytjzbqs/qyjqzs/>

本文行业罗盘覆盖的 500 个细分行业，与企业景气指数使用的统计行业分类并不总能一一对应。牛码新市依据细分行业的业务构成，将其映射到相关的基础行业指数；一门生意受多个基础行业环境影响时，按预设权重组合这些指数。映射关系和权重说明的是本方案对行业关联程度的判断，将附上设定依据，随计算版本一起保存，供日后复核。

农林牧渔业和金融业在本方案使用的企业景气指数序列中，没有可直接对应的指数。牛码新市因此借助中国国家统计局发布的季度国内生产总值（GDP）初步核算数据作辅助统计。该数据分别列示行业增加值的当季绝对额和当季同比增长速度：绝对额按现价计算，反映当季行业增加值的规模；同比增速按不变价计算，反映剔除价格变化后的行业增加值增长。本方案使用当季值，不拿半年或全年的累计值代替单季数据。

对农林牧渔业和金融业进行牛码调整系数估计时，先从同时具有企业景气指数与季度行业增加值数据的行业取样，把同一行业、同一季度的资料配对。配对前将核对行业分类、统计覆盖范围与数据版本，不能仅凭行业名称相近就直接对应。模型以已公布的企业景气指数为被解释变量，以增加值的当季同比增速和当季绝对额为解释变量，建立跨行业、跨季度的线性回归关系；再把农林牧渔业和金融业同期的增加值数据代入，分别估计其景气指数。增加值规模和增长并不等同于景气程度，这一估计关系能否适用于缺项行业，仍须验证。

模型检验应同时考察后续季度和未参与建模的已知行业，检查估计误差及其对行业排序的影响。根据检验结果调整模型后，还须用未参与调整的数据再次检验。对已知行业的检验，不能代替对农林牧渔业和金融业适用性的判断。这里得到的是模型估计值，应与中国国家统计局直接发布的指数区分。证据不足时，应保留估计限制，不把模型输出当作已经验证的行业读数。

每轮系数计算以最近一个已取得完整资料的季度为基准，模型训练则使用事先确定的历史季度样本。已公布的企业景气指数、用于估计缺项的 GDP 行业增加值，以及后续的映射、组合和排序，都须对应同一基准季度，不把较新的景气指数与上一季度的 GDP 数据拼在一起。本版以 2026 年第二季度为计算基准；未来有了能够完整对应的新季度数据，再统一更新计算和排序。所用数据的发布时点与修订版本也将留存，便于复核。

取得同一季度的行业景气指数后，牛码新市按各行业指数的相对位置形成调整系数。系数以 1 为相对调整的中性基准，靠前或靠后的行业按统一参数作相应调整。这一中性基准不同于企业景气指数 100 所代表的景气临界值，相对靠前也不意味着行业处于景气区间。这个系数表达的是当期行业环境的相对位置，不是相邻行业经营差距的精确刻度；指数很接近时，细

微名次变化不建议过度解读。参与排序的行业范围、系数转换公式、调整幅度及其作用于行业侧读数的方式，须随方法版本说明。

本版展示的是行业罗盘的规则成果，以及牛码调整系数的设计方法。具体企业的实测评分与跨行业校准，仍需在后续应用中完成。实际使用时，每个系数都应能查到所用的数据、观察时期、行业映射与权重，以及模型和参数版本。牛码调整系数为行业侧判断补充环境参照，合同结构仍由通用罗盘审视，组合层的风险留存系数（DR）也有其独立用途。行业环境提供参照，具体生意怎样，最终还要结合企业经营、合同安排、现金流与履约证据来判断。

5.4 从行业内判断到跨行业比较的阅读路径

阅读双罗盘，先看合同结构的收益支撑与风险负载，再看因子怎样解释具体生意。两侧各自用了什么资料和期间，都应核对。先把来源看明白，再看后续的排名比较。

随后检查行业侧校准：是否采用相应牛码调整系数，基准覆盖什么对象与时间，结果适用哪版规则。调整后仍应能返回原始因子、未调整系数、证据和基准，分清差异来自经营事实、行业背景，还是期间与版本变化。

比较前，把对象、行业、期间、版本、数据范围与校准条件逐项对照。月度和年度数据的观察范围不同，新旧规则也可能改变定义和边界。数字形式相同，只说明都写成了分数，条件是否可比还需另查。

资料缺失、行业归属不清、主营业务跨行业，或不同来源冲突，都应明确标识，并说明对评分和比较的影响。低分、尚未评分和证据不足是不同状态。不能为凑齐读数而混用业务、期间或口径；需要合并计算时，应说明方法与依据。

阅读顺序由此清楚：先合同、后生意；收益和风险分别看；核对行业侧校准，再在期间、版本与证据条件相匹配时比较。每一步都让读数保留自己的来处。

5.5 使用边界：行业成果、项目核验与持续观察

500 个行业成果为理解生意和评分口径提供了基础，案例帮助读者学会阅读，具体行业仍须沿索引查找。规则覆盖已经形成，各项因子、评分和校准方法在不同环境下表现怎样，还需要持续验证。

进入具体项目，仍要核对企业的业务模式、经营阶段、数据质量与合同安排。多项业务并存时，要说明本次读数覆盖哪一部分。行业名称相近，只能帮助查找，不能代替资料核验和适用性判断。

双罗盘帮助组织信息、指出收益支撑和风险负载，以及下一步应核验的位置。主体、合同、现金流和持续经营变化仍须具体检查。罗盘不构成收益承诺、法律意见、合规批准或自动交易结论；投资决定由具有相应决策权限的主体依据完整资料与适用规则作出，相关专业确认和批准仍按授权及资质要求办理。

经营、外部基准和方法都会变化。新结果写清新条件，历史结果保留旧依据，才能解释前后差异。持续观察的价值，是让判断随事实得到复核和修正，而不把一次评分当成永久结论。

通用罗盘提供合同结构的共同尺度，行业罗盘保留具体生意的差异，通用码与行业码各自记录收益和风险，牛码调整系数为行业侧比较补充环境参照。把对象、期间、版本、证据与适用条件一起交付，500 个行业成果才更容易被读懂、检验和持续使用。

第三部分 帽子军团理论

制造与维护登记和度量仪器的理论来源

第二部分说明了我们建造什么。本部分回答这样的仪器由什么组织制造和维护。答案是帽子军团这座 AI 工厂：它把完整模型能力放进可管理、可检验的微任务，再让大量任务协作完成专业工作。

当 AI 进入真实金融工作，会同时遇到两组矛盾：机器的处理能力可以扩展，责任却须落到具体的人和组织；人的判断与治理责任不能缺席，带宽却无法追上每一次机器动作。以下各章解释 AI 工厂为什么采用 HAT over MASS，以及边界、协作和检验怎样配合。第一部分的成本困境在这里转化为组织设计问题。

第六章 AI 能力与人类责任的两组矛盾

6.1 放开 AI 与盯住 AI 的两难

AI 一旦进入真实业务，问题很快就从“能不能做”变成“我们敢放手到什么程度”。只看能力，它可以同时处理大量信息，让许多任务并行推进，还会随着模型、数据和工具更新继续扩展边界。它不受固定工时和单一岗位带宽的同等约束，我们当然希望把更多工作交给它，让过去因速度和成本做不到的事，第一次变得可做。

能力跑得越快，责任越不能含糊。AI 可以生成判断、提出建议、执行步骤，却不会因为一次结果正确，就自动成为法律和组织意义上的责任主体。谁规定它能做什么，谁划出禁区，谁处理例外与争议，谁对结果负责，这些问题最终都要落到具体的人和组织。所谓“AI 的能力无限、责任有限”，说的正是这种不对称：处理能力和执行规模可以快速扩展，最终责任却不会随着调用次数自动长出来。

人这一端恰好相反。判断力、注意力和工作时间都有上限，不可能跟着海量机器任务一起成倍增长；但到了授权、签署、监督和追责的环节，人又不能因为任务太多就退场。“人的能力有限、责任无限”并非要求个人承担无边界责任，而是在提醒我们：组织决定使用 AI，就必须让责任完整落位，不能把解释不了、检查不到、无人认领的部分留给系统自行消化。

第一条常见路径，是“AI 为主、人为辅”：先让 AI 大规模生产，再让人处理少量例外或完成最终签署。规模很快就能起来，压力却会全部挤到最后一道人工关口。AI 可以并行产生成千上万项结果，人的审核和签署带宽却不会同比增加；审核者来不及回到输入、规则和过程，签字便容易变成形式。AI 多次做对可以积累使用经验，却不能用概率性正确代替过程控制，更不能代替清楚的责任主体。

另一条路径，是“人为主、AI 为辅”：AI 整理资料、做初步分析、给出提示，关键环节仍由人逐项判断。面对重大决策、高度依赖自由裁量或难以预先写成规则的事项，这种安排仍然必要，也让责任关系更直观。可一旦任务变得高频、海量且持续发生，人工审核就会堵住整个流程。AI 虽然进来了，系统仍按人的速度运行，能够扩展的机器能力始终没有变成组织真正的处理能力。

两条路各自守住一端，也各自丢掉一端。一边是“人不够用”，另一边是“AI 不敢用”；把 AI 放开，担心责任失控，把人放在每一步前面，能力又被掐回人的速度。我们要找的答案，必须让能力释放与责任落位从一开始就长在同一套结构里。牺牲其中任何一端，都只是把问题推迟。

6.2 金融行业的必答题：合规、投资者保护与风险控制

很多行业都会遇到这组矛盾，金融行业尤其躲不开。金融活动连接资金、经营和风险，一项结果影响的不只是流程效率，还关系到参与者权益和市场信任。技术可以更新，合规、投资者保护和风险控制却不能等业务跑完再补，也不能因为处理量变大就降低要求。

AI 对金融业务真正有吸引力的地方，恰恰是它处理高频、海量任务的能力。资料要整理，信息要核对，经营要持续观察，已经形成的判断也要随新事实反复检验。若这些工作都由少数专业人员逐项处理完成，服务范围和更新速度必然受人的带宽限制。把可拆分、可重复的工作交给 AI，才可能让过去因成本过高而被挡在门外的任务进入处理范围。

可金融行业引入 AI，绝不能只是把人工按钮换成自动按钮。模型会做一件事，不等于组织已经授权它做这件事。每次执行都要说得清：输入从哪里来，依据哪一版规则，结果接受什么检验，遇到什么情况必须停下并交还给人。少了这些连接，自动化越快，错误、偏差和遗漏就越可能被成批复制。

所以，能力释放和责任控制要在同一结构里一起设计。输入、输出和合格标准明确的任务，可以让 AI 在既定边界内持续执行；每项产出仍要进入对应的检验环节，能够发现、退回和修正错误。需要新判断、碰到规则空白或超出授权范围时，系统必须把问题交给有相应职责的人。人的价值不在于跟在机器后面把每件事重做一遍，而在于制定规则、划定权限、处理例外、监督运行并承担最终责任。

这样一来，“审核”也有了更严格的含义。看到结果后点头不算审核；能够查明它用了什么资料、走过哪些步骤、适用哪一版规则、异常由谁处理，才算把责任接住。自动执行、逐项检验和明确担责必须连成一条链，AI 带来的规模才不会以牺牲合规、投资者保护或风险控制为代价。金融行业真正要答的题，是怎样让效率与责任一起扩展。

6.3 帽子军团怎样同时解开两组矛盾？

帽子军团（HAT over MASS）意在同时解开 AI 能力与人类责任的两组矛盾，并承接第一部分的专业服务成本困境。规则边界把 AI 的能力与人的责任接起来，微任务协作则让专业能力进入海量小任务。两者合在一起，才有机会把“大材小用”变成“大财小用”的组织基础。

“大材小用”是把强大的模型能力用在范围明确的具体任务上。模型可以拥有完整的理解、分析与执行能力，每次任务的目标、范围和权限则须预先界定。我们把复杂工作拆成可独立交付、可检查的单元，让各项任务按依赖关系协作衔接，产出经过检验后进入下一环节；遇到超出规则或授权范围的问题，再交由有权处理的人作出判断。

帽子军团由此而来。“帽子（HAT）”是人设定并维护的规则边界，“军团（MASS）”是大量 Agent 在边界内协作。单个 Agent 领取一项微任务：目标集中，输入输出说得清，周期受限制，结果能检验，过程留下必要记录。这里的“微小”约束的是一次任务的范围和权限，并没有把模型能力一起缩小。

我们框住个体，是为了把每次执行查清楚。任务越大、边界越模糊，出了问题就越难分辨究竟错在资料、规则、执行还是验收。任务拆到可查、可验，异常才能定位到相应环节；碰到规则之外的事项或需要自由裁量，就停下自动推进，升级给人。约束个体的目的，是阻止能力伸进没有责任承载的地方，而非把 AI 困到无效。

群体的释放，来自这些可管理单元的并行与协作。边界、交付和检验条件清楚以后，组织才有底气扩大任务数量，让不同 Agent 分别生产、核对、协作。海量本身不会带来可靠，数量也不会自动变成正确。真正让规模成立的，是每个单元都能检查，分工之间能够制衡，出了问题可以一路回到对应规则和责任。

在这套结构里，人和 AI 各自扛不同的事。AI 在规则范围内执行大量可拆分任务；人制定和修改规则，决定授权边界，处理例外与争议，监督系统表现，并承担不能交给机器的最终责任。人不必亲手完成每一个动作，责任也没有因此消失。它落在更关键的地方：规则有人定，边界有人守，异常有人接，结果经得起复核。

图 6-1 将两组矛盾与 HAT、MASS 的回应放在一起。

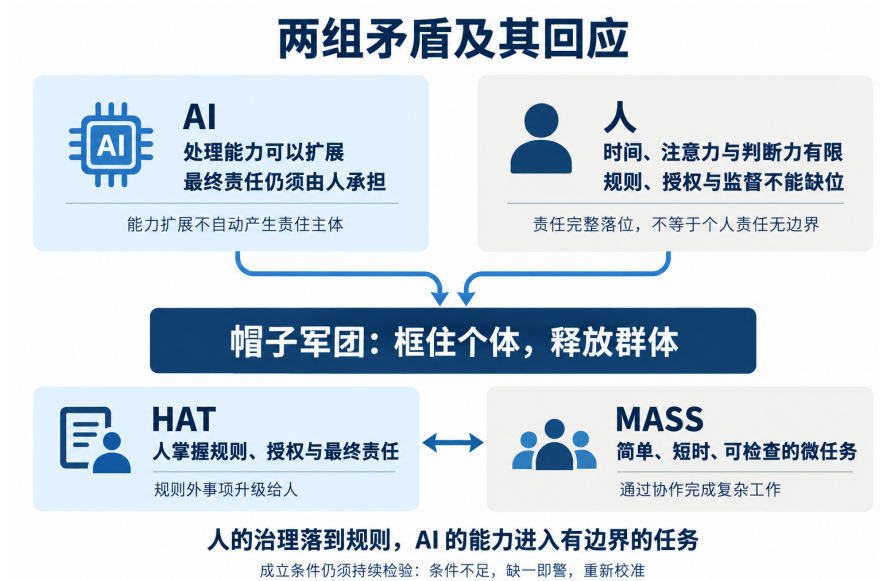


图 6-1：两组矛盾及帽子军团的回应

当然，这套答案有严格前提。任务能否拆到可以独立检验，规则能否覆盖实际能力，记录能否还原过程，不同底座和岗位能否形成真正的检查与制衡，都会决定它是否可靠。规则盲区、数据污染、同源偏差和误判传导，还可能被规模进一步放大。帽子军团提供的是一套边界、协作和责任的组织原则，它从未替任何自动化结果预先背书。

三组概念到这里连成一条线：理想是“大财小用”，卡点是“大才不能小用”，答案是“大材小用”。我们把完整模型能力放进可管理、可检验的微任务，框住个体，释放群体，让仪器的每项产出都能回到规则、证据和责任。接下来说明帽子怎样让人掌握规则，军团怎样以微任务形成海量协作，以及协作和检验怎样决定这座工厂是否可靠。

第七章 HAT 帽子：人掌握规则，AI 依规执行

如果 AI 每走一步都等人审批，海量、高频任务就永远跑不起来；如果只管自动执行，责任又会掉进找不到真实主体的空白。HAT 帽子正是为了解开这个结。它不再纠缠人和 AI 谁做得更多，而是先把两件事分开：**人掌握规则以及规则的变更权责，AI 在规则已经覆盖的范围内持续执行。**

“帽子”把规则、授权边界和升级安排装在一起。它告诉单个 AI 哪些事可以做，哪些事不能碰，走到什么情形必须把问题交还给人。个体边界清楚，边界内的动作才不必次次等人点头；规则始终由人掌握，自动执行也就不会变成责任转移。

7.1 HAT 的含义：人的最终责任与 AI 的潜力释放

HAT 由三个组成部分构成：Human-Commanded（人管控）、AI-Operated（AI 运作）、Trust（信任），本质是“人管控、AI 运作”的信任机制。它同时回应两项要求：人的规则、授权与最终责任能够落实，AI 在明确边界内具有充分的执行空间。信任则由责任安排、过程记录和交付检验共同建立。只强调管控，可能限制 AI 的有效使用；只强调能力，又可能使执行超出组织能够监督和负责的范围。

人类管控，意味着由人制定、修改和废止规则；AI 运作，意味着规则、权限和任务范围明确后，AI 可以自主、持续地完成规则内的工作。人的工作因此转向规则、授权与异常处理，AI 的能力则在可检查的任务中展开。

因此，HAT 里的信任从来不靠一句“AI 会做对”，也不会把一次正确交付当成长期可靠。它落在三个可以检查的条件上：责任找得到真实主体，过程看得见也还原得了，结果能按同一口径复核。信任由此不再只是对能力的感觉，而成为对责任、过程和结果的结构性的判断。

7.2 规则先于能力：允许、禁止与升级条件

能力越强，越要先把行动范围说清楚。引入 Agent 之前，我们先回答三个问题：允许它做什么，明确禁止什么，遇到什么情形必须中止并升级给人。允许事项给出行动空间，禁止事项划出红线，升级条件负责接住规则尚未覆盖的情况。

一句方向正确的原则，还不能直接拿来运行。规则至少要能执行、能检验，还要在系统里形成实际隔离。能执行，意味着对象、任务和适用条件足够清楚，Agent 不必靠临时自由裁量

补上关键空白；能检验，意味着结果可以逐项对照规则，不能只凭“看起来合理”放行；实际隔离则要求未授权的数据、工具、动作和流程在运行中真正受限，不能只在纸面上写一句“不允许”。

规则也不会一次写完、永久有效。模型、工具、数据或任务方式一旦变化，原有帽子还能不能覆盖新能力，就要重新判断。新能力已经出现，并不意味着它适合当前任务或已获准进入正式运行；其适用性、可靠性，以及允许、禁止和升级条件，都需要重新确认。

对资金划转、合同签署、对外承诺等高后果动作，以及涉及许可、资质或安全约束的操作，一旦触发预设暂停条件，应立即暂停相关自动执行，同时记录触发原因并保存证据，再交由具名真人处置。完成整改与复验后，由有权主体批准恢复。恢复不能仅以“错误已改”为依据，还须留下复验结论和批准记录。

7.3 每顶帽子背后的具名真人：认领与规则制定、修改和废止

规则要真正承担治理作用，除了写得清楚，还必须找得到对它负责的人。每个 Agent 由具名真人认领，组织才能明确它适用什么规则、处于什么授权范围、规则变化时谁来处理。“具名”要解决的，是在明确组织和部门职责的基础上，进一步找到负责规则、授权与异常处理的具名人员，避免责任只停留在笼统的系统或部门名称上。

认领绝不是把名字挂上去就算完成。规则为什么制定，后来改了什么，新版本适用于哪些任务，什么时候应停止使用，都要留下可识别的责任与记录。若规则已经变更，却说不清是谁在何时改的、新旧版本各管哪一段，后续执行便无从审查。检查一顶帽子，最终要回到三件事：规则是否清楚，变更是否留痕，执行能否回放。

负责人扛的是规则、授权和异常处理，不是替 AI 把每项任务再做一遍。若具名认领最后变成认领人逐项手工复核，责任虽然有了名字，处理能力仍被人的带宽锁住。有效的认领，应让规则决定有人负责、授权边界有人维护、规则之外的问题有人承接，同时让规则之内的执行继续交给 AI。

7.4 帽子只管边界，不管动作

帽子管的是边界，不是边界内的每一个动作。它规定允许范围、禁止范围和越界后果，让 Agent 在授权内依规执行。任务一旦超出授权、触及禁区或碰到规则没有覆盖的判断，系统

就按预设条件限制行动或升级给人，不能让 Agent 自行扩大权限、改写规则，更不能代替人完成自由裁量。

“不管动作”当然不等于“看不见动作”。规则内执行仍要留下过程信息，接受检查；输入错误、执行偏差或环境变化，也仍可能让结果不合格。帽子改变的是人介入的方式：对规则已覆盖、授权明确的日常任务，人无需逐项审批，而是通过规则、授权、监督和升级处理实施治理；依法或依组织制度须事前批准的动作，仍按相应要求执行。

这样一来，人直接管理的对象便从海量单项任务，收敛为可以维护、审查和修订的规则。任务可以持续增加，背后的规则却必须始终找得到、看得见，也能对应到负责主体。规模化能走多远，因而不只看执行能力增加了多少，更看组织能否维护规则、识别边界变化、接住越界事项。

7.5 人的治理、监督与自由裁量责任

AI 接手规则内的重复执行后，人的工作没有消失，而是上移了：从逐项执行、逐项复核，转向规则治理、授权监督和责任承接。哪些任务可以规则化，哪些风险必须隔离，什么情况下要停止自动执行，哪些判断仍需人凭经验、权限和专业责任作出，都要由人决定。系统做得到，并不等于这些问题可以省略。

治理也不能停在规则发布的那一天。规则进入真实任务后，人要持续感知它运行得怎样。这种“运行体感”不能只是笼统印象，它要落到具体信息：哪些任务反复触发升级，哪些结果经常被退回，哪些异常无法归入现有规则，复核和回测发现了哪些稳定偏差。运行信息不断回到规则层，帽子才能跟着环境、能力和任务变化继续修订。

规则之外的自由裁量，仍然由人承担。事实不足且无法按既定流程补充、目标冲突无法依规则处理，或判断需要超出已有授权时，系统应暂停受影响的自动执行，把已有事实、执行记录和规则缺口提交有权限的人处理。若人的处理方式具有可重复性，还应成为下一轮规则修订的候选依据，不能长期靠临时批准维持。

HAT 是一种持续运行的关系。人给 AI 设定自己能够负责的边界，AI 在边界内发挥能力，运行结果再回到人的监督和规则修订。人无需亲自完成所有动作，却始终握住规则、责任和自由裁量。边界这样立稳以后，微任务才可以安全增加，海量 Agent 之间也才有一套共同遵循的秩序。

涉及资金划转、合同签署、对外承诺等可能产生重大影响的操作时，应事先明确谁有权作出专业判断、谁可以叫停、谁可以批准恢复。触发暂停条件后，按本章 7.2 节所述流程停止操作、保存证据，并交由有权处理的人处置；完成整改和复验后，经批准才能恢复。认领 AI 岗位或参与任务，本身不代表拥有上述权限。

第八章 MASS：强能力怎样进入微任务

HAT 先把边界立起来：什么能做，什么不能做，走到哪里必须回到人。接下来，问题从治理进入组织——怎样让 AI 在这些边界内大规模执行，又不让规模本身制造新的不确定性。

MASS 军团的答案，是把完整模型能力组织成微任务。一个 Agent 可以拥有很强的理解、分析和执行能力，但在一次具体运行中，只领取范围有限、周期较短、能够完整检验的任务。大量微任务可以并发、衔接、相互校验，个体受到的约束才会转化成群体的执行力。

所以，“微小”和“海量”恰好互为条件。微小把单个任务约束到可规定、可检查、可重做；海量把这些单元组织起来，承接过去受限于人类工时而无法覆盖的工作。MASS 不靠简单堆高 Agent 数量，它依靠任务粒度、权限边界、验证条件和底座配置，把大规模执行变成一件能够管理的事。

8.1 MASS 的含义：微任务智能体、简单且短时的任务与海量协作

**MASS: Micro Agents (微任务智能体)、
Simple & Short Tasks (简单且短时的任务)。**

军团里每一个智能体 (Agent) 都建立在前沿大模型的能力之上，但完整的模型能力，不等于无限的任务权限。它领到的是一件微任务：简单、短时，边界清楚到结果可以检查，颗粒细到关键步骤可以回溯。让能力强大的 AI 只承担一件范围很小的事情，是为了使任务能够被组织、被检验，并通过成千上万个微任务的并行协作形成规模化能力。微小让人管得住海量，海量让 AI 释放出潜力。

“简单 (Simple)”要求一次任务只承担一个清楚的判断或交付；“短时 (Short)”要求它能在一次边界清楚的运行中完成，不依赖长期持续且无法追踪的临时判断。周期没有统一的分钟数，但输入、输出、结束条件和检查标准都要预先说清。

复杂目标拆成边界清楚的单元，海量并发才有真正的基础。不同 Agent 可以同时处理彼此独立的任务，也可以沿着明确的依赖关系前后衔接。输入、输出和检查条件能够辨认，并发结果才能被比较、组合或退回；任务本身若含糊，增加 Agent 只会同时制造更多无法判断的输出。

MASS所说的“海量”，是一种把成千上万项微任务组织成协作网络的能力：每个单元有边界，单元之间有关系，结果有检查。这不等于任何系统在任何时点都能无限并发。它表达的是一种扩展方向——只要单元设计和检验条件承受得住，任务数量就不再被单一人类团队逐项执行的速度锁死。

8.2 任务小，兵不小：完整模型能力与有限任务权限

“任务小，兵不小”，先把能力与授权分开。这里的“兵”是 MASS 对 Agent 的比喻，并不意味着它只会一种低水平技能。相反，一个 Agent 可以建立在完整而先进的模型能力上，但一次运行只能使用完成当前任务所需的权限、资料 and 工具，并对明确的交付负责。

给高能力划边界，是组织高能力的前提。若因为模型会做更多，就把资料搜索、判断、计算、修改甚至批准全塞进同一次运行，结果可能看起来完整，错误却很难定位，责任也无法分辨。模型能力有多宽，不能替代任务边界有多清楚。

小任务给错误划出一个可辨认的范围。一次只处理一类问题，错误更容易隔离在对应环节；边界足够窄，检查者才能按预先说明的标准判断是否通过；周期足够短，输入、操作和输出才有条件完整保留、重新检查。这些条件不能保证 Agent 永远正确，却能阻止错误披着无法拆解的“完整答案”直接流入后续工作。

这就是“大材小用”落到任务层面的含义。“大材”是可以被大规模调用的强模型能力，“小用”是让它每次专注于一件可以判断、检查和重做的事。与其让一个 Agent 同时扮演所有角色，我们更愿意把强能力投入清楚的任务，再通过协作完成复杂目标。能力没有被压低，它只是被转换成可管理的执行单元。

8.3 任务拆分三个足够：足够小、足够窄、足够碎，保留完整业务意义

微任务不能只按字数、时长或步骤数来量。它能否进入 MASS，要同时看三个条件：足够小、足够窄、足够碎。三者分别回答三件事——一次承担多少工作，执行时留下多少临时裁量，产出能否独立检验和重放。

足够小，是一次只做一件事。这件事可以有多个步骤，但只能有一个清楚可辨的交付目标。输入从哪里来，允许使用什么，要输出什么，合格标准是什么，都要在执行前说清。若一个任务同时要求搜集事实、制定规则、判断风险和批准结果，它就还没有小到可以清楚验收。

足够窄，是任务不依赖执行时任意增加的自由裁量。Agent 可以在预先确定的范围内选择方法，却不能自行改目标、扩权限，或代替人完成未获授权的判断。事实不足、规则冲突、结果越界，或任务需要引入新的自由裁量时，执行就停在已知边界内，升级给人。升级并不说明任务设计失败，它证明权限没有被悄悄放大。

足够碎，是每项产出能独立检验，关键步骤能重新回放。“碎”并非按段落切文章，也不能把一个判断拆到失去意义。若某个单元必须靠执行者口头补上一大段未记录背景才看得懂，它形式再短，也不是独立微任务。合理拆分要让每个单元仍有完整含义，又能被单独退回、修正或重做。

小、窄、碎，缺一项都不成立。任务很小，却允许无限裁量，它不够窄；边界很窄，产出却无法独立检查，它不够碎；拆成很多份，每一份仍背着多个交付目标，它也不够小。三项条件一起满足，微任务才成为能够大规模复制、分配和检查的组织单元。

条件	具体含义	检查要点
足够小	一次只承担一个清楚的交付目标	输入、输出与合格标准能否事先说明
足够窄	任务只在已授权边界内执行	是否依赖临时自由裁量或自行扩权
足够碎	结果可独立检验，关键步骤可回放	能否单独退回、定位与重做

表 8-1：任务拆分的三个足够

8.4 先可验证，再上规模：独立验收、错误定位与局部返工

复杂目标一旦拆成微任务，数量往往会迅速增加。原本由数人串行完成的一项工作，可能展开成许多信息获取、独立判断、交叉检查和结果组合任务，并分给不同 Agent 执行，形成更大的编制。可编制变大只说明任务被拆开了，并不能证明拆得好。

单个任务没有明确标准，数量越多，未知输出就越多；单个任务无法独立验证，组合后的大结果便说不清哪一步出了问题；一项产出若必须靠执行者临时补充大量未记录背景，另一个 Agent 也无法稳定接手检查或继续执行。边界和证据的缺口，不能靠数量填平。

扩量之前，组织要先把几个基本问题答清楚：每个单元的输入输出是否明确，合格标准能否在执行前说清，过程能否回放，失败后能否只退回相应单元，需要裁量时是否会停下并升级给人。这些问题有了可判断的答案，单个任务才具备复制和并发的条件。

创建 Agent 本身的成本相对有限，真正随规模一起增长的，还有模型调用、数据、工具、调度、检验和异常处理等组织成本。这里不需要先假设一个通用数字。能不能扩展，首先取决于每个执行单元是否可查、可判断、可退回，而不是系统能创建多少个 Agent 名称。无法验证的单元复制得越多，风险越难收敛。

MASS 遵循的顺序很简单：先做到可验证，再扩大任务数量。先证明一个任务能被清楚交付和检查，才把同类任务复制到更大范围。执行规模、检验能力、规则覆盖和责任承载必须一起增长；只有任务越来越多，检查和异常处理仍停在原地，称不上可管理的海量协作。

8.5 底座支撑三个足够：多元、对冲与性价比

任务边界设计清楚，仍不足以支撑海量协作。若所有任务都依赖相同的模型来源、训练偏差和解题路径，单个任务看似独立，大规模输出仍可能在同一处集体出错。因此，MASS 除了要求任务足够小，也要求承载任务的模型底座足够多元、足够对冲、足够性价比。

足够多元，是不把全部任务和检查压在同一个底座上。在适用范围内使用不同模型底座和不同供应商，可以减少同源偏差被反复复制的风险。多元追求的是有意义的差异，不是型号越多越好。不同品牌的模型仍可能共享相似数据、检索结果或工具链，不同供应商也不天然等于完全独立。它能降低同源偏差，却不能宣告偏差已经消失。

足够对冲，是让适合并行检查的同一任务由不同底座独立执行，把一致和分歧都摆到台面上。它的价值不在于多数模型投票就一定正确，而在于让原本藏起来的不同判断变成可检查信号。出现实质分歧时，Agent 不能自行抹平差异或挑选有利答案；达到预设条件，就应自动触发升级检查，再依规则进入进一步复核或人工处理。

“触发升级”在这里首先是一项组织原则。落到具体任务，什么差异算实质分歧，由谁处理，怎样记录，后续流程是否暂停，都必须写进对应规则。没有分歧标准，同时产出多份答案也只是并行生产，还没有形成真正的对冲。

足够有性价比，是按任务难度、风险和检查需求配置不同能力与成本的模型。所有任务都调用成本最高的模型，海量协作难以长期维持；只看价格、使用能力不足的模型，失败、返工和检查成本又可能更高。真正的性价比，是在满足任务和检验条件的前提下，让能力、成本与规模相匹配。

多元、对冲和性价比必须一起看。只有多元，没有对冲，不同底座只是各做各的，彼此没有形成可观察的制衡；只有对冲，没有多元，多次执行仍可能重复同一种偏差；前两者都有，

却不考虑性价比，验证要求也很难扩展到海量任务。三项条件共同守住两件事：底座不能成为单一失败来源，资源配置也不能拖垮规模化执行。

MASS 最终不看军团里有多少“兵”，而看每个单元是否被约束在可检验范围内，承载它们的底座是否足够多元、能够对冲、也能在成本上持续。任务的三个足够，让单元本身成立；底座的三个足够，让这些单元在规模中保持可比较、可持续。两组条件一起站住，微小才能真正支撑海量，能力释放也不必拿管理失控来交换。

条件	具体含义	检查要点
足够多元	保留不同模型与能力来源	是否真正存在来源与方法差异
足够对冲	独立形成判断，再比较依据与分歧	检查是否重复了相同假设，分歧是否保留
足够有性价比	能力与任务难度、成本相匹配	把调用、数据、检验与返工成本一起考虑

表 8-2：底座支撑的三个足够

第九章 微任务怎样组成可靠的协作

把任务拆小，只解决了单个 Agent 怎么管，还没有回答成千上万个任务怎么一起跑。规模一上来，三个问题会立刻冒出来：谁派活，谁判断结果能不能过关，规则管不到的地方由谁接住。关系没理顺，Agent 越多，交接、冲突和错误就可能传得越快；可若每一步都退回人工审批，海量协作又会被压回人的速度。

所以，帽子军团不能只会拆任务，还要有一套撑得住规模的组织方法。规则治理与任务执行分开走，调度、验收与回测各自守住职责，每个 Agent 再由具名真人认领。这样设计的目的很明确：让大量 Agent 在边界内真正协作，同时让每一项产出都有人检查、能够复核，也经得起后续事实的检验，而不是再叠加一层组织架子。

9.1 “帽对兵”与“兵对兵”：规则治理与任务协作两条通道

帽子军团的组织原则可以浓缩成一句话：人管帽子，AI 管任务。帽子先把边界说清楚——什么可以做，什么不能做，碰到什么情况必须停下或升级。规则由人制定、修改和废止，规则之外的自由裁量也回到人。Agent 可以带着很强的能力上场，但只能在帽子划定的范围内行动。

这就形成两条通道。“帽对兵”管的是规则与边界：任务能不能进来，可以调用哪些资料和能力，触到哪条线必须限制或升级。“兵对兵”管的是边界内的协作：谁接任务，结果交给谁，怎样验收，怎样复核。一条通道把能力框在可检查的范围内，另一条通道让日常动作不必步步等人的临场指令。

为什么一定要分成两条通道？因为人一旦成为所有任务的传话筒和批准人，就会成为整张协作网络最窄的瓶颈，Agent 的并发能力也就无从释放。但任务协作若完全脱离帽子，规则缺口、错误输入和越界动作又会顺着链条一起跑下去。我们要的，是边界内放开协作，触到边界时能够限制、退回或升级。

人离开的，是逐项执行的位置；人不能离开的，是治理责任。规则是否清楚，过程有没有留下证据，异常有没有被处理，需要自由裁量时谁来决定，这些问题始终要有人回答。Agent 之间的协作也必须建立在明确的任务目标、输入输出和合格条件上。人管帽子、AI 管任务，靠的正是规则把责任与执行重新接起来。

图 9-1 展示规则边界、微任务协作与独立检验怎样接成一条交付链。

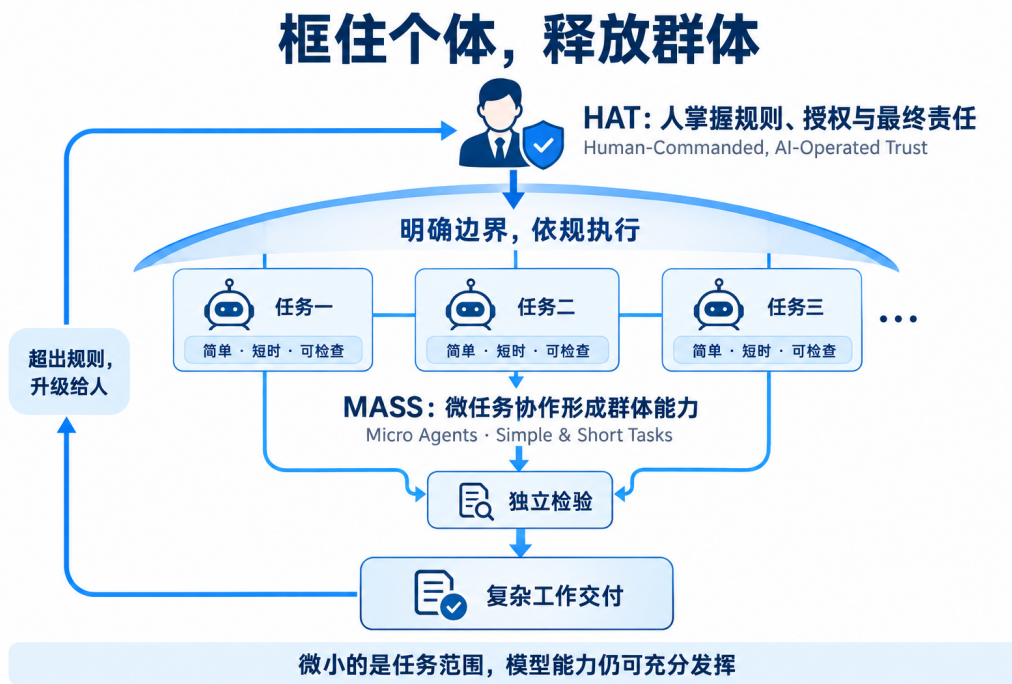


图 9-1：规则边界下的微任务协作与交付

9.2 从拆分到组装：任务依赖、交接接口与成果衔接

任务拆开之后，还要接得回去。一个单元的结果，往往是另一个单元的输入；前一步尚未确认，后一步就不能把它当成已经成立的事实。调度先识别这些依赖：哪些任务可以并行，哪些必须等待前置结果，哪些条件改变后需要重新安排。并发由依赖关系决定，不能只看手边有多少 Agent。

交接接口要把下一位接手者需要的内容说清楚。交付不仅是一段结论，还应带上输入依据、适用规则、输出口径、版本、检查状态和未解决事项。下一环节据此确认能否接收，不必依赖前一位执行者临时补充背景。接口明确，分工才不会把完整问题切成彼此听不懂的碎片。

组装也有自己的验收要求。每个分项都合格，仍需检查它们采用的定义、单位、期间和版本能否对齐，彼此是否矛盾，合起来是否覆盖了最初的交付目标。整体结果不能靠把若干合格答案拼在一起自动成立；组合后的关系，同样需要独立检查。

某一接口未接好，应先定位受影响的环节，再判断哪些后续结果需要重做。调度负责重排依赖，执行岗位修正相应产出，验收岗位检查修改及其传递影响。能局部返工，也能看清返工影响到哪里，微任务的灵活性才不会变成整条链反复重启。

9.3 从分歧到处理：澄清、复核、退回与人工升级

不同 Agent 给出不同答案，先要查清它们是否在回答同一个问题。输入材料、概念定义、观察期间或适用规则不同，都可能产生表面冲突。先把这些依据摆在一起澄清，才能分辨分歧来自口径不同、证据不同，还是判断本身不同。

澄清之后仍有实质分歧，就进入复核。复核者依据既定规则检查事实、计算过程和推理，保留各方意见及其依据。模型数量、表达信心或交付速度，都不能代替这一步。证据不足时，应留下尚未解决的问题，不能为了拼成一个答案而抹掉异议。

只有确认需要修改，才把任务退回对应工作单元。退回时要说明问题、依据、修改范围和重新验收的条件，让执行者知道需要修正什么；不因某个环节有错就推翻所有已合格产出。涉及前置条件或共同口径变化的，还要检查影响是否传到后续任务。

超出现有规则、需要扩大授权或引入新的自由裁量时，Agent 应暂停受影响部分的自动执行，并升级处理。它把事实、已有规则、分歧和可能影响交给有权限的具名真人，由人决定继续、修订还是终止。处理结果再回到任务与规则记录，日常协作才能在清楚的边界内恢复。

9.4 分项合格与整体交付：调度、验收与回测的分权制衡

协作一旦规模化，派活、放行结果和持续检验就不能混成同一种权力。调度、验收与回测都在同一条任务链上，却分别回答不同问题：任务怎样流动，单项产出能不能进入下一环节，既有判断和规则能不能经受更长时间、更多情形与后续事实。

调度负责让工作跑起来。它分配任务、时间和资源，确认输入与前置条件，把工作交给合适的执行单元；结果被退回或条件发生变化时，也由它重新安排。但调度掌握顺序，不等于掌握专业结论的最终判断权，更不能为了赶交付而跨过验收。

验收守的是结果进入下一环节的门。它依据预先设定的标准，看产出能不能被理解、核对和继续使用，而不是听执行者说“我做完了”。合格，才能继续流转；未达到准入条件的结果应暂缓流转，先说明问题并核对依据；明确需要修改时，再退回相应工作单元。执行者不能自己给自己盖章，调度也不能用“任务已完成”代替验收。

分项合格与整体交付要分别判断。独立验收既检查各项产出，也检查汇总后的接口、依赖、口径和目标覆盖是否完整。调度可以组织组装，不能因此自行放行；整体仍有矛盾或缺口，就应按影响范围退回相关环节。

回测把时间拉长。一次产出在当时过关，不能证明规则换到别的环境仍然有效，也不能证明这次正确来自稳定能力而非短期运气。持续检验把观察范围从一次验收延伸到不同时间与场景：历史回测检查明确版本的方法在既有资料中的表现，后续事实验证观察其在新数据中的表现，压力测试检验不利情形，回归检查和异常监测则帮助识别变更影响与运行偏差。这些检验相互补充，各自保留对象、版本、样本和结论。

把三项职责分开，不等于任何组织都必须机械设置三个固定岗位。我们要守住的是判断的独立性：调度不能代替验收，验收不能代替长期回测，生产者也不能用自己的说明结束审查。岗位怎样安排，可以随任务范围和风险而变；生产与审查的分离、异议的保留、问题返回正确位置的能力，不能跟着变。

分权的目的是，是让一条任务链能够自己纠偏，而不是让三方彼此掣肘。调度推动工作前进，验收守住结果准入，回测检查规则和能力是否还有效。验收发现问题，先澄清并定位，确需修改的进入返工；持续检验发现偏差，先区分输入、执行、流程与规则原因，再采取相应调整。涉及规则变更的，由有权主体批准后重新进入调度与验收。

9.5 每个 Agent 由具名真人认领：临时协作与持续责任

Agent 可以承担执行，却不能成为组织责任的终点。每个 Agent 都由具名真人认领，它所戴的帽子、获得的授权，以及规则之外的问题，因而都能回到明确的人。这样，异常发生时，组织面对的不会只剩一个模型名称、一段指令和一份无人接手的结果。

认领人的核心工作是管帽子，不是替 Agent 干活。他需要知道规则是否清楚、授权是否合适、过程是否留痕，回测和异常信号是否说明帽子需要调整。日常任务只要规则覆盖、权限明确、结果可检验，就让 Agent 依规完成；触发升级条件、遇到规则空白或需要自由裁量时，再回到有权处理的人。

具名真人认领最怕走向两个极端。一头是逐项审批：每个动作都等人确认，组织重新被人的带宽卡住。另一头是挂名认领：名字登记了，运行状态却没人真正了解，规则松动和异常信号也无人处理。有效认领要求人离开每一步执行，却始终贴近规则、留痕与回测；在其责任范围内未能识别规则松动和异常状态，也应承担解释责任。

实际运行中反复出现的问题，应成为检查和修订规则的依据。Agent 总在同一环节产生分歧，输入条件已经变化，或现有规则无法明确下一步如何处理，这些情况不能仅靠反复返工解决。

认领人应查明：问题出在执行、授权，还是规则未能覆盖当前情形，再推动相应调整。执行者不能自行修改规则，但应将发现的问题、相关证据和处理记录提交给有权修订规则的人。

具名认领也有自己的边界。认领人负责帽子和授权，不因此自动成为每项任务的调度者，更不能取代独立验收或回测。组织要分别说清谁派活、谁产出、谁判断合格、谁处理规则之外的事项。把这些角色都压成一个人，看似责任集中，实际只会让责任重新变得模糊。

协作小队可以随任务组建，也可以在交付后解散，具名认领关系和工作记录却不能随之消失。岗位身份承接持续职责，每次运行留下对应证据；后续发现问题，仍要找得到当时适用的规则、获得的授权与承担责任的人。

9.6 编制扩充三个足够：足够多、足够细、足够冲及其承载边界

帽子军团的编制，不是一张人数清单。它要解决的是任务有没有人覆盖、错误能不能找到环节、同一种判断有没有力量来挑战。为此，编制要同时做到足够多、足够细、足够冲。三个“足够”分别对应覆盖、归因和制衡，少一个，海量微任务都很难真正组织起来。

足够多，先保证必要任务都有人承担，关键判断也留得出交叉复核的位置。编制不够，同一个 Agent 就可能跨过太多任务边界，重要环节也只能依赖单一判断。但“多”从来不等于无限扩张。每一个新增 Agent 都要有相应规则、检验和责任承接；数量与任务、检查关系对不上，再大的编制也只是数字。

足够细，是把分工切到能够定位错误的程度。即使任务已经很小，若一个模糊角色仍然包办输入选择、形成判断、验收和修订，出了问题依旧说不清源头。分工变细以后，每个环节的输入、输出和责任更清楚，不合格结果也能退回真正需要修正的位置。代价是交接会增加，所以规则和调度必须把这些细单元重新接成一条线。

足够冲，是让不同立场真正形成制衡。“冲”不是制造组织对立，也不是为了争论而争论；它要防的是所有执行者和检查者沿着同一套假设，整齐地得出同一个偏差。不同立场需要独立提出判断、保留异议，再用证据和规则处理分歧。只有这样，系统性偏差才可能在扩散前被看见。

“足够多、足够细、足够冲”分别解决任务覆盖、责任分工与独立制衡的问题。岗位数量充足，但分工不清，错误仍难以定位；分工细致，却缺少不同立场的独立检查，偏差仍可能在多个环节重复出现；不同意见充分，但缺少承接复核的岗位和调度安排，问题也难以得到处理。

三项条件共同成立，大量 Agent 才能在同时执行任务的基础上，形成分工明确、衔接有序、能够相互检验的协作。

扩编的边界，由规则、检验与责任承载能力共同决定。任务种类增加，允许、禁止和升级条件就要跟着维护；产出增加，检查所需的证据和复核能力也要跟上。新增 Agent 只是潜在执行力，规则覆盖与检验能力接得住，才会成为有效协作。

异常还必须有人接得住。责任主体不清，问题会在任务之间来回传递；超出规则的事项全部压到少数人身上，组织又会堵在人的处理带宽上。扩编时要同时检查规则维护、证据复核和人工升级的承载条件，不能让数量跑在它们前面。

编制也不会设计一次就永久有效。任务在变，能力在变，事实反馈也会改变覆盖、归因和制衡的要求。组织要持续看见哪些任务没人接、哪些环节责任过宽、哪些判断缺少独立挑战，再据此调整。下一章讨论的，正是如何用全量检验、持续回测和完整溯源，为这些安排每天补上新的证据。没有持续检验，编制再完整，也只是一张设计图。

条件	具体含义	检查要点
足够多	覆盖必要任务与交叉复核	执行、检验与异常承接能力是否一起扩展
足够细	把交付与责任分到可辨认单元	问题是否能回到具体岗位与任务
足够冲	不同立场与职责形成制衡	生产、检验和协调是否保持独立意见

表 9-1：编制扩充的三个足够

第十章 怎样让每次交付经得起检验

Agent 把人的能力推向过去承担不起的任务规模，也把另一个问题推到我们面前：产出越多，藏在其中而未被发现的错误也可能越多。所以，军团能否扩编，不能只数有多少 Agent 在工作，还要看每项任务有没有规则、每个结果有没有证据、每次异常能不能找到处理责任。上一章把调度、执行、验收与回测分开，最终都是为了回答：我们凭什么相信这一项结果？

这里的信任，落在一项具体结果上，也带着明确条件。它要经得起一连串追问：结果是否符合当前规则，输入是否成立，过程能否还原，后续事实是否继续支持原判断，出了问题谁来处理。检验、回测与溯源各自回答其中一部分，也共同决定这份信任能走多远。

10.1 全检而非抽检：每项产出均接受检验

传统抽检来自一个现实前提：检查资源有限，只能从部分样本推断整体。它可以帮助我们看总体风险，却无法证明每一项具体产出都过了关。Agent 开始海量并发以后，产出更快、更多；如果检查仍停留在少量样本，任何一项没被抽中的错误，都可能直接进入下一环节，被引用、复制，最后被放大。对使用这一项结果的人来说，“整体大致可靠”不够，他需要知道眼前这一项接受过检查。

帽子军团所说的全检，是让每项产出都进入与任务性质相匹配的验收或检验流程，不留下默认免检区。它不要求所有任务用同一种方法，也不要求真人把每项结果从头重做：规则清楚、结构稳定的任务，可以由 Agent 核对完整性、一致性和边界；影响重大、规则未覆盖或需要自由裁量的事项，则要进入更深检查并按条件升级。全量产出检验完成后，授权范围内的人工抽查仍可继续存在，用来核查检验机制、高风险结果和异常处理是否真实有效；人工抽查是追加监督，不替代全检。

AI 同样可以参与检验。重复比对、格式核对、规则匹配、异常发现和交叉复核，都适合由它承担，这让每项产出进入检查流程在规模上成为可能。但成本降下来，边界不能一起降下来。执行者不能自行宣布结果合格，检验者也不能离开既定规则临时发明标准。生产与审查仍要分开，检查权限、升级条件和人的最终责任，仍由帽子所代表的规则体系来定。

每次检验还要留下彼此对得上的记录。结论属于哪项任务，依据哪个规则版本，检查的是哪份产出，异常由谁处理，都要能够连起来。只有一个“通过”，却说不出查了什么、按什么规

则查、问题交给了谁，这不是完整证据。全检的目的也不是堆积形式记录，而是让每项产出都有清楚的检查入口和判断依据。

10.2 持续回测：识别短期运气与环境变化

一次验收通过，只能说明结果在当时的输入、规则和检查条件下过了关。它证明不了能力会一直稳定，也证明不了规则换个环境仍然有效。答案做对了，可能靠的是可靠方法，也可能刚好碰上有利样本；规则在一个时期有效，换到另一种市场、业务周期或边界情形，也可能失去解释力。持续回测，就是为了把稳定能力与短期运气分开。

回测要把一个明确版本的判断，放进更长的历史、不同场景和边界条件里。所谓足够长，没有一把适用于所有任务的尺。短周期任务也许要覆盖不同输入和异常状态；受季节、经营周期或市场变化影响的任务，就要跨过相应阶段。能力只有在多种条件下反复站得住，我们才有理由相信它主要来自稳定方法，而非一时走运。

回测还要发现答案什么时候过期。数据来源会变，业务运行方式会变，市场环境和外部规则也会变。一个结论过去正确，如果支撑它的条件已经变化，就不能凭旧验收继续沿用。新事实与原判断放在一起比较，才能看出偏移来自环境、输入、执行能力，还是规则没有覆盖新情况。原因不同，处理也不同：更新资料、校准能力、修订规则，或交给真人重新判断。

所以，回测不应是发布前做完一次就结束的程序，它是信任继续成立的条件。频率、样本和阈值不必强行统一，但每次回测都要说清楚：检验哪个版本，使用什么期间和事实条件，发现偏移后怎样处理。缺了这些信息，“持续回测”只剩一句好听的原则，分不清能力、运气与环境变化。

10.3 全程溯源：从结果回到输入、调用、规则与责任

检验看见问题以后，下一步要找到问题从哪里开始，不能笼统归结为“模型错了”。只盯着最后提交结果的 Agent 远远不够，一项产出往往同时受输入资料、模型调用、规则版本、任务交接和检验方式影响。全程溯源，就是沿着结果往回走，把形成结果的关键条件一段段找回来，让错误能够被定位、解释和修正。

一条可复核的证据链，要把结果同本次任务的输入、模型或工具调用、适用规则版本、执行路径和检验结论连起来。输入告诉我们当时掌握什么事实，调用记录说明用了什么能力，规则版本交代依据什么边界，执行路径呈现任务走过哪些环节，检验记录则说明结果为何被接

受、退回或升级。再连接身份编码与具名真人认领关系，组织才能同时回答：结果怎样形成，异常由谁处理。

把过程还原出来，才能分清问题的性质。资料不完整或失真，是输入问题；规则没覆盖新场景，是规则缺口；交接中丢了条件，是流程问题；调用的能力或版本不适合，则要校准能力选择。若把这些都算成“模型错了”，真正需要修正的位置反而会被放过；若只为了问责找一个方便的人承担，问题也不会消失。溯源要做的，是让责任落到能够实际纠正问题的地方。

修订以后，溯源还没有结束。输入更新了，调用调整了，规则或流程改了，都要再用同类问题验证一次，看原来的错误是否真的被挡住。修订前后的版本、检验结论和适用范围也要保留，让后来者看得见问题如何发现、如何处理、新方法在什么条件下重新成立。只证明“改过”，还不能证明“改对了”。

全程溯源是一项证据要求，不是一句对系统能力的承诺。每次 Agent 实例运行产生运行编码，并由它关联所属任务；运行编码可以帮助定位这次运行，却不等同于任务号，更不能替代证据。最终能追到哪一步，取决于实际留下了哪些输入、调用、规则、流程和责任记录。没有留下的部分，不能用“系统已经验收”或“理论上应该留存”补出来。

10.4 检验的三个足够：足够密、足够长、足够深

检验的三个足够，从三个方向托住信任：足够密，看覆盖；足够长，看时间；足够深，看能否还原。它们必须连在一起工作，任何一项都不能只停留在口号上。

条件	具体含义	检查要点
足够密	每项产出进入相应验收或检验	是否仅依赖少量抽样，或存在默认免检区
足够长	持续覆盖不同时期、场景和边界	是否把短期运气当成能力，继续使用过期答案
足够深	沿输入、调用、规则与责任回查	能否定位错误来源，并验证修订效果

表 10-1：检验的三个足够

足够密，要求每项产出都进入检查。检查深度可以随任务风险、影响范围和规则成熟度而变化，全量检验之后也可以继续进行授权人工抽查；但任务再多，也不能让一部分结果默认绕过检查。人工抽查用于追加核查，不把全检重新退回抽检逻辑。

足够长，要求回测跨过与任务真正有关的环境与边界。只在一个时点、一个条件下正确，既排除不了偶然，也看不见环境变化后的偏移。多长才够，要由任务对象和它的变化规律决定，一刀切的期限回答不了这个问题。

足够深，要求我们不能停在输出文本、一个运行编码或一个“通过”状态上。运行编码把一次 Agent 实例运行关联到所属任务，却不等同于任务号；要解释偏差从哪里来、该由哪个环节纠正、修订后有没有真正挡住同类问题，还必须继续回到输入、调用、规则、检验和责任记录。

三个足够，少一个都会留下结构性空白。检查很密，时间不够长，连续的短期好运可能被当成稳定能力；回测很长，溯源不够深，发现偏差也不知道从哪里下手；溯源很深，覆盖不够密，仍会有产出完全没进检查。任务、底座、编制、检验四组条件合在一起，才是完整的十二个足够。第四部分把这些设计条件落实为岗位与作业程序。

第四部分 AI 工作手册

供具名真人、技术尽调方与 AI 审计者查阅

本部分是帽子军团的 AI 工作手册，也是制造登记与度量仪器的车间手册。它说明 5000 名 AI 员工的岗位设置、任务拆解、组队规则与七站流转程序，供具名真人、技术尽调方和 AI 审计者沿职责、流程与记录检查。

一般读者可先读各章的“本章概要”，再按问题回查具体步骤。岗位台账、一次任务与一次实例运行分别管理；本书以连锁便利店说明性模拟展示交接和检验，不把模拟链当作后台真实运行成绩。手册的作用，是让执行和责任有据可查。

第十一章 5000 名 AI 员工怎样进入组织〔操作程序〕

本章概要 | 本章回答：5000 名 AI 员工怎样进入组织？

关键规则：

- AI 员工是相对稳定的岗位，由具名真人认领，承接长期职责并积累工作记录；
- 具体任务到来时，再按岗位要求启动 Agent 实例，调用模型、技能与工具完成当次执行；
- 每次运行单独留痕，并关联所属岗位和任务；
- 任务结束后，岗位身份、认领关系和历史记录继续留存。

**一般读者读到这里即可，正文细节供组织者、实践者、技术尽调方与 AI 审计者等专业人士查阅。*

第三部分要求由人掌握规则、授权与最终责任。本章通过三层结构、四类职能和具名认领，将这些要求落实为可辨认的岗位关系（参见第七章 7.3 节、第九章 9.5 节）。

11.1 已经上岗的 5000 名 AI 员工

5000 名 AI 员工已经上岗，分布在调度管理、业务执行、管控合规和部署支撑四类职能中。

“AI 员工”是岗位的对外称呼：每个岗位承接长期职责、具名认领关系和工作记录，Agent 实例则承担当次执行。岗位已经进入组织安排，具体任务运行与验证到了哪一步，还需要相应记录说明。

人数扩大以后，第一个要厘清的并非模型有多少，而是岗位与运行如何区分。AI 岗位是组织内部相对稳定的身份，承载持续的职责、规则、归属和记录；具体任务到来时，系统再按岗位要求和任务条件启动 Agent 实例，调用所需能力完成一次运行。岗位可以先后参加多项任务，也可以在不同时点采用不同能力。实例结束，岗位身份仍然保留。

所以，“已经上岗”和“正在执行任务”不能混为一谈。前者说明岗位已进入组织台账，具备可识别的身份和职责关系；后者说明某个 Agent 实例正在进行一次具体运行。5000 名 AI 员工已经上岗，不代表 5000 个实例始终同时工作，也不能据此推算一项任务调用了多少 Agent。岗位规模、任务组织与单次运行需要分开记录，数据才不会彼此冒充。

一个岗位进入台账，意味着职责、具名真人认领关系和实际任务记录都要随之落位。人类认领者负责规则、授权、监督，以及需要自由裁量时的升级处理；日常派活、组队和检查则交由 Agent 依规协作。我们没有把每一步重新送回人工审批，因为那会让规模化协作重新退回人的速度；我们保留的是人的规则责任和升级责任。

稳定岗位与每次任务之间，还要靠记录连起来。身份编码（Agent Identifier, Agent ID）识别岗位及关联其具名真人认领关系；Agent 能力表现记录实际调用能力的来源、版本与表现；每次 Agent 实例运行形成一个运行编码，并关联所属任务、权限、输入、输出和检查记录。下一章会展开这三类信息。本章先回答更前面的问题：5000 名 AI 员工如何被组织成一支有规则、有分工、能随任务协作的帽子军团。

11.2 三层设计：规则层、职责层与编号层

一份岗位名单撑不起一个大规模 AI 组织。共同规则缺位，通用能力就可能越过授权边界；职责含混，多个 Agent 即使都能完成工作，也说不清该由谁承担；身份和运行记录断开，任务结束后便无法还原过程。牛码新市帽子军团因此设置规则层、职责层和编号层，让行动边界、组织分工与运行证据落在同一套结构里。

规则层先把行动边界说清楚：哪些事可以做，哪些事不能做，出现什么情形必须暂停受影响的自动执行或升级处理。任务允许范围、权限条件、交付标准和异常处理要求，都要转化为 Agent 在接单、调用能力和提交结果时能够执行的约束。遇到规则空白、证据不足或需要价值判断的事项，能力再强也不能自行扩张授权；这时应当停下来，按既定条件交给给人处理。

职责层回答谁来承担，以及彼此如何配合。长期组织中有调度管理、业务执行、管控合规和部署支撑，具体任务里又有不同责任位置。“能做”只说明能力，“该做”才涉及责任。一个 Agent 能够生成结果，不代表它有权批准结果；一个岗位完成了流程协调，也不能用“流程已结束”代替质量判断。职责层把调度、生产、检查和支撑分开，同时为跨岗位协作留下接口。

编号层把稳定身份和每一次运行接起来。身份编码回答哪个岗位参加了工作，Agent 能力表现说明本次调度依据了什么能力，每次 Agent 实例运行则形成独立的运行编码，并关联到所属任务及过程记录。岗位身份不会随任务结束，能力表现会在实际使用中积累，而同一任务可以包含多个实例、多个运行编码。编号的价值在于让复核者从结果一路回到岗位、能力配置和那一次具体运行。

三层如何连接，并把检查与回测反馈给人，可对照图 11-1 阅读。

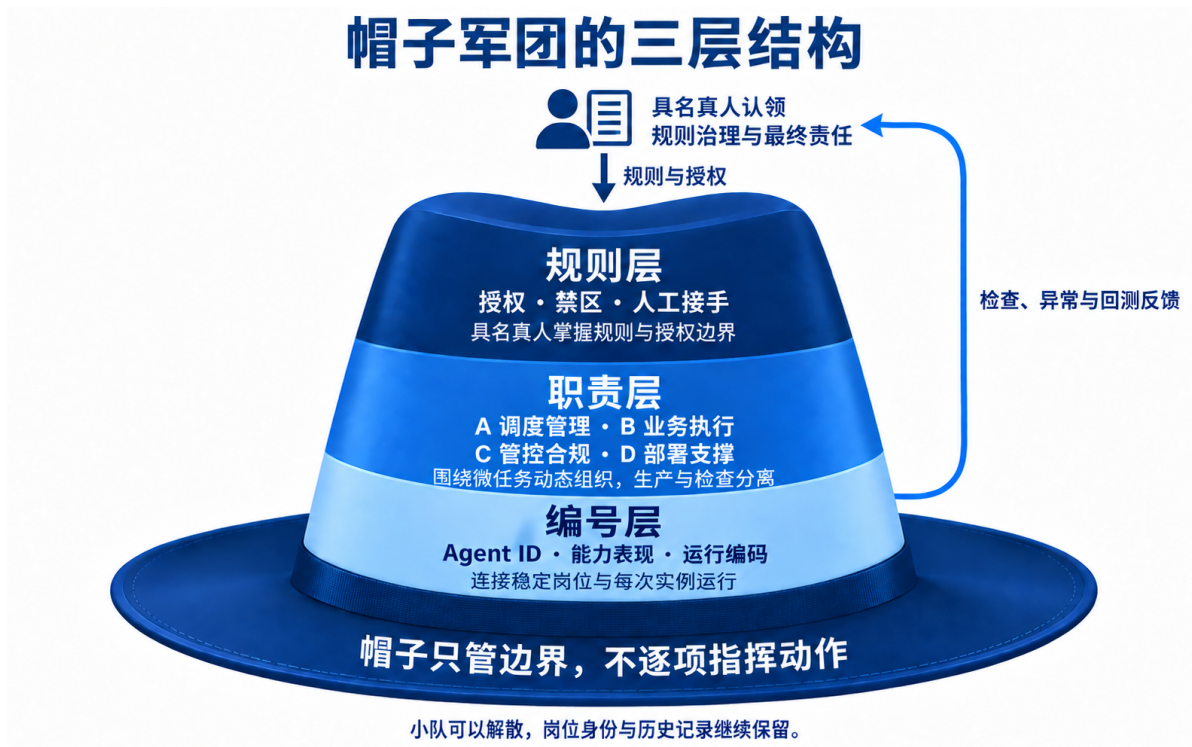


图 11-1：帽子军团的三层结构

三层各管一件事，也彼此制约。规则层划边界，却不能替岗位承担责任；职责层确定分工，却不能证明工作怎样完成；编号层连接记录，也不会自动让规则变得合理、让结果变得正确。三层同时工作，组织才既能保留稳定身份，又能按任务灵活组队，并在出现问题时找到真正需要处理的环节。

“帽子”把这三层关系装进一个可运行的组织单元：它连接岗位角色、规则、权限和责任，却不等同于行政职级，也不是某个模型的别名。同一套基础能力可以服务不同岗位，前提是本次戴什么帽子、拿到什么权限、承担什么交付都足够清楚，调用过程也有记录。这样，我们可以共享能力底座，同时保留岗位之间真实而必要的差异。

11.3 四大职能：调度管理、业务执行、管控合规、部署支撑

三层设计搭起骨架，四大职能回答 5000 名 AI 员工怎样分工。它们并非四组互不来往的 Agent，也不是所有任务都必须机械走过的四道审批。组织长期需要四种责任：有人把目标组织成任务，有人形成业务结果，有人独立检查边界与质量，也有人把模型、数据、工具和运行环境准备好。

A：调度管理（Administration）把目标变成可执行、可检查的工作安排。它识别任务范围、前后依赖和交付要求，再组织拆解、选岗、衔接与验证。调度的本领在于让合适的岗位在合

适的时点进入任务，而非替专业岗位作判断。行业罗盘怎样拆成工作单元、AI 员工怎样按任务组队，后续章节会把这条线具体展开。

B：业务执行（Business） 直接形成与任务目标相对应的分析、处理和交付。行业不同，所需知识、数据和工作方法也会不同，因此业务岗位需要专业分工。它们是结果的主要生产者，却不能自己完成对结果的最终认可。重要判断仍须接受相应检查，支持判断的输入、依据和过程也要一并留下。

C：管控合规（Control） 持续追问三件事：工作是否在授权边界内，证据是否足够，结果是否达到要求。风险、规则、法务、财务、合规及必要的独立检查都在这一职能内展开。业务岗位不能包办对自身结果的全部验证，调度岗位也不能拿“流程走完”替代质量结论。发现问题时，管控合规需要提出澄清、复核、修订或升级要求。

本书所称“管控合规”，是 AI 在授权内取证、核对、提示异常、开展设计审查和质量检查的职能。涉及合规批准、投资判断、适当性判断、对外生效的企业最终评分及法律意见，由具备相应授权与资质的人作出或确认并留痕。AI 仍可按规则计算和形成评分建议；内部计算、独立检查与最终批准分别承担不同责任。

D：部署支撑（Deployment） 准备并维护模型、数据、技能、工具、系统和运行环境。一个 AI 岗位能否稳定工作，取决于数据路径是否可用、工具是否接入、权限是否配置、版本是否受控，而不只取决于模型能力。部署支撑负责让这些条件按任务需要可被调用并留下变化记录；它能证明技术条件是否可用，却不能替业务和管控岗位判断结论是否正确。

四大职能连成一条责任链：调度管理组织任务，业务执行形成结果，管控合规检查边界与质量，部署支撑提供运行条件。具体任务可以依难度、风险和依赖关系并行推进，也可以分阶段进入，无须固化成一条不可调整的审批流水线。长期职能让责任稳定，动态组队让 5000 名 AI 员工围绕目标迅速协作。

11.4 多视角席位：3 类行业视角 + 3 个模型对冲、3 个投资者视角与“6 + 1”检验视角

组织分工解决“谁承担哪类责任”，行业罗盘还要解决“同一门生意该从哪里看”。不同立场会看到不同事实，也会提出不同问题。牛码新市帽子军团因此设置多视角席位，把经营观察、模型对冲、投资判断和设计检验分开。视角是一种职责位置，决定岗位关注什么、提出什么判断；它既不是模型数量，也不是某个模型的名字。

三类行业视角分别设置为正常经营者视角、老炮视角和骗子视角。正常经营者从日常经营出发，判断某个因素怎样影响收入、成本、效率和持续经营；老炮带入长期行业经验，识别纸面规则与真实经营之间的落差，以及周期、惯例和隐性约束；骗子视角专门寻找指标被包装、口径被利用或结果被刷高的路径。这里的“骗子”是设计审查中的对抗席位，不对任何具体经营者作身份判断。三种看法放在一起，是为了减少只按理想流程理解生意的偏差。

三个模型对冲属于底层智力调用，与三类行业视角是两个不同方面。面对同一任务，我们让不同模型分别处理，再比较它们的假设、推理和结论；差异出现时，继续核对输入、口径和依据。多模型有助于暴露单一路径的遗漏，却不会自动形成有效对冲。只有任务条件、输入依据、输出要求和分歧处理方式一致，结果才可比较。正式口径因此是“3 类行业视角 + 3 个模型对冲”，两者不能合称九类行业视角，也不要求模型与岗位——绑定。

三个投资者视角分别关注生意、价格和风险。生意评估追问企业靠什么创造价值、经营机制能否持续；价格评估关注价值判断怎样落到价格条件；风险评估检查不利情形、脆弱环节和潜在损失。行业视角先把生意看懂，投资者视角再追问这门生意、这个价格和这些风险对资本决策意味着什么。三者讨论的对象相连，责任却各不相同。

经营观察和投资判断形成候选因子后，“六鉴 + 执衡”只负责本轮设计审查、资格判断和组合筛选。问业检查候选因子的经营意义；量尺检查定义、计算和评分能否唯一复算；寻源检查可信数据路径与必要条件；分水检查区分不同经营状态的依据；照影检查相对全池因子新增了什么信息；破局检验真实状态没有改善时能否通过口径或输入刷分。执衡再把审查结果归入排除、待澄清、待修订或可选，并在可选池中进行组合筛选。可选只代表具备进入组合讨论的资格，不等于最终入选，更不等于已经通过企业取数或实际验证。企业取数、实际评分、校准、部署和发布属于后续实施。

多视角的价值在于把不同责任照亮。行业视角负责看懂生意，三个模型对冲比较底层能力形成的假设与判断，投资者视角追问资本与决策，六鉴与执衡则审查本轮候选因子的设计和组合资格。意见不同，不必急着压成一个答案；先保留依据、澄清分歧，再按规则决定是否需要修订或升级。这样形成的制衡，来自清楚的责任和处理路径，而非简单增加 Agent 数量。

11.5 视角与能力分开：模型、技能和工具按任务调用

多视角能够运转，关键在于把岗位视角和执行能力拆开。视角规定岗位需要作出哪一类判断，并承担相应责任；模型、技能和工具则是完成判断时调用的能力。岗位的职责位置可以保持

稳定，具体能力根据任务对象、证据条件和权限范围调整。同一模型也可以服务不同岗位，但相同模型不会把不同责任变成同一种责任。

这样做，我们不必把每个模型永久包装成一个“员工”。模型更新、技能调整、工具替换，岗位身份和责任关系仍可延续；岗位进入新任务时，再在规则范围内装配适合的能力。共享底座减少重复建设，差异化配置确保每个 Agent 只获得完成当前职责所需的背景、工具和权限。能力可以变，责任必须始终找得到。

按任务调用依然要守边界。模型、技能和工具的选择必须对应任务目标、允许范围和交付标准；多个模型参与对冲时，输入和输出要可比较，分歧要能被识别；调用数据和工具时，要服从相应权限和使用条件；遇到高风险判断或规则空白，则按条件升级给人。组织调用的是能力，帽子管住的是能力怎样被使用。

能力怎样被选、实际表现如何，都要进入记录。Agent 能力表现需要说明调用了什么能力、来自哪个版本、完成了什么任务、形成了什么结果，又接受了怎样的检查。每次 Agent 实例运行形成运行编码并关联所属任务，复核者由此可以继续追问：问题来自岗位职责、能力选择、输入证据，还是执行与检查环节。记录不会自动带来正确答案，却让验证、回测和修订有地方落脚。

至此，稳定组织与动态能力有了清楚关系。规则层守住允许、禁止和升级边界，职责层确定岗位分工，编号层连接身份与运行；四大职能维持长期责任，多视角席位承担观察、判断和设计审查，模型、技能和工具按任务进入一次次运行。5000 名 AI 员工的协作能力，来自稳定岗位承担清楚责任，也来自动态能力能够在规则内流动，并要求每次执行都能找到对应身份和记录。

这套架构为后续运行打下基础。下一章将具体说明 AI 岗位、Agent 实例、身份编码、Agent 能力表现与运行编码如何连接，让稳定岗位、具名真人认领关系和单次任务记录既能分开管理，又能沿同一条证据链相互核对。

第十二章 一个岗位、一次运行，怎样分清责任〔操作程序〕

本章概要 | 本章回答：一个岗位、一次运行，怎样分清责任？

关键规则：

- 身份编码识别“哪个岗位参与”，能力表现说明“为何调度它”，运行编码记录“这次做了什么”；
- 每次运行都须关联所属任务、当时的权限、能力版本、输入输出与检查结果；
- 返工或重新执行须生成新的运行编码；
- 岗位状态和能力配置更新时，应保留历史版本，不能覆盖过去的执行条件；
- 编码只是查找入口，结果能否复核，取决于关联记录与证据是否完整。

**一般读者读到这里即可，正文细节供组织者、实践者、技术尽调方与 AI 审计者等专业人士查阅。*

第三部分把全过程可追溯作为可信交付的条件。本章通过身份编码、运行编码、版本与证据关联，把稳定岗位和当次执行接起来（参见第十章 10.3 节）；能力发生变化时，记录应支持重新验证。

5000 名 AI 员工进入组织运行后，管理就需要回答一个更实际的问题：长期岗位、可调用能力和某一次执行，究竟怎样区分？如果同一个名称同时指向岗位、配置和临时实例，结果需要复核时，身份、权限、过程与责任就会缠在一起。

本章把信息分成三个层次。AI 岗位承载稳定的组织身份，Agent 能力表现提供检索和调度依据，Agent 实例记录一次具体执行。身份编码与运行编码把它们接起来：临时组队可以解散，长期身份不会中断；一次结果也可以沿记录回到当时所属的任务、配置和证据。

12.1 AI 员工的岗位与任务实例

“AI 员工”是面向组织与使用者的称呼；进入内部管理，还要把岗位和实例分开。AI 岗位是以正式、相对稳定的组织身份进入长期台账；Agent 实例则在具体任务发生时按要求启动，完成一次技术运行。一个岗位可以先后参加多项任务，每次调用不同模型、技能、工具、权限和上下文。实例结束了，岗位仍然存在。

这一层区别守住了责任的连续性。AI 岗位说明执行主体在组织里承担什么职责、由谁认领、当前状态如何；Agent 实例说明它在什么时间、哪项任务和哪些约束下做了什么。长期记录归岗位，单次信息归实例。两者混在一起，实例停止可能带走责任线索，某一次偶然结果也可能被误当成岗位的长期表现。

岗位也不能用模型名称来代替。模型、技能和工具是完成任务时调用的能力，岗位承载的是组织身份与职责。模型、技能和工具等技术配置可以更新，任务中的帽子和权限也会变化；这些变化应进入能力表现与运行记录。如此一来，模型替换之后，组织仍能认出同一个岗位，也能比较它在不同配置下做得怎样。

每个 Agent 实例都要有清楚的任务边界。启动时，它接收任务目标、当前帽子、权限、上下文和交付要求；一次运行随之形成独立运行编码，并关联所属任务、输入输出、调用信息和检查结果。任务结束后，小队可以解散，岗位身份、具名真人认领关系和历史记录继续保留。我们把稳定身份与临时执行分开管理，再用运行记录把它们连回去。

图 12-1 汇总了稳定岗位、按任务组队与交付记录之间的关系。



图 12-1：稳定岗位与任务协作中的交付记录

12.2 AI 员工的身份编码、能力表现与运行编码

身份编码（Agent ID）对应 AI 岗位的稳定组织身份，连接岗位含义、主要责任、岗位状态、具名真人认领关系和历史记录，不因某次实例启动或结束而改变。它为这些复杂信息提供共同地址：岗位做过什么、状态怎样变化、由谁认领，都能沿同一身份归集；编码本身不会取代职责内容。

Agent 能力表现用于说明岗位目前凭什么可以被检索和调度。业务场景、行业知识、专业角色、模型、语言、区域、技能和工具都可以进入记录，但一个静态名称远远不够。能力从哪里来、采用哪个版本、在什么任务条件下产生过怎样的结果，会直接影响下一次能否调用。能力表现因此既服务调度，也接受持续校准。

每次 Agent 实例运行都会形成运行编码，并关联所属任务。运行编码把本次目标、帽子、权限、上下文、输入输出、模型与工具调用、检查结果和时间信息接在一起。相同岗位参加不同任务，会产生不同运行编码；同一任务也可能包含多个实例运行，返工或重新执行同样需要新的运行编码，不能用一个任务号覆盖全部过程。

三类信息分别回答三个问题：身份编码说明哪个稳定岗位参与，Agent 能力表现说明为什么调度它，运行编码说明某一次实例究竟做了什么。缺少身份编码，多次运行无法回归长期岗位；缺少能力表现，调度依据无从判断；缺少运行编码，具体过程又会消失。三者接在一起，身份、调用理由与执行证据才能相互核对。

身份和运行记录也不能替代本次任务的帽子与权限。身份编码保持相对稳定，当前帽子说明岗位在本次协作中的角色，权限限定它能读取、调用和提交什么。岗位可以在不同任务中承担不同角色，但每次变化都要落到相应运行编码，并与所属任务和当时授权条件对应；稳定身份本身不能自动带来当前权限。

12.3 AI 员工的岗位台账、状态更新与实例信息分层管理

AI 岗位台账以身份编码为主线，保存岗位含义、主要责任、具名真人认领关系、岗位状态、能力表现和历史记录。它不是一张静态名册。组织需要通过台账判断岗位现在能否被调用、职责和能力怎样变化，以及过去的任务经验能否支持新的调度。

Agent 实例记录保存的是一次运行：本次戴什么帽子、拿到什么权限、使用哪些上下文、模型、技能和工具，形成哪些输入输出，又接受什么检查。这些信息随任务产生，不能直接覆盖长期身份。一次临时调用工具，不代表工具永久进入岗位能力；一次检查未通过，也不等于岗位身份被取消。是否更新状态或能力表现，要回到规则和运行事实。

岗位状态描述组织层面的真实情况。它需要反映岗位是否仍在有效管理之下、是否符合调用条件，以及暂停、恢复等变化在何时、依据什么发生。状态不能靠一次任务结果自动推断，也不能脱离认领关系、配置情况和运行记录随意改写。把这条线守住，“已经上岗”和“此刻正在运行”才不会被当成同一件事。

Agent 能力表现也要保留来源与版本。行业知识、模型、技能或工具更新后，可调用能力可以随之调整，但历史运行必须留住当时实际使用的版本和表现。否则，新记录会覆盖旧条件，复核者无法判断一项结果究竟由哪版能力产生。每次状态或能力表现的更新都应有可核对的运行事实，并保留前后版本关系。

长期台账与实例记录分层，还能把责任和信息边界分开管理。具名真人认领关系连接稳定岗位，用于查找规则、授权和异常处理的责任主体；具体任务资料、上下文和输入输出则服从本次实例权限。组织无需为每项任务重造身份，也不会因为岗位已有身份就默认它可以读取所有资料、执行所有动作。

这套分层最终要回到调度。系统看到的应是岗位当前状态、能力来源与版本，以及真实运行表现共同形成的调用依据，而非一个孤立标签。新任务启动后，每次实例运行形成运行编码并关联所属任务，实际配置和结果再沉淀回台账，让岗位状态和能力表现能够在事实上更新。

12.4 AI 员工运行编码关联的记录与证据留存

运行编码是一项索引：它把一次实例运行中分散的信息接起来，让组织从结果回到所属任务、当前帽子、权限、上下文、输入输出、模型与工具调用、检查结果和时间信息。索引并不等于证据。记录缺失、版本不明或输入输出无法对应时，即使有运行编码，也还原不了完整过程。

一份可复核的记录要回答：实例接收了什么，在什么约束下处理，调用了哪些能力，形成了什么结果，又经过哪些检查。资料来源或任务口径一旦变化，也要进入同一次运行的关联记录。复核者由此看到的不只是最后的文本或评分，还能核对输入、规则、授权条件是否前后一致。

放到行业罗盘任务里，身份编码先找到参与任务的稳定岗位及其认领关系；Agent 能力表现说明调度时依据了哪些行业知识、专业角色、模型、技能或工具条件；运行编码则对应一次实例运行，并关联所属任务、本次资料、配置、产出和检查结果。三类信息合在一起，才能区分“岗位具备某种能力”和“这次是否按相应条件完成了工作”。

模型与工具调用也属于运行记录。岗位在不同任务中可以采用不同模型，同一任务的不同环节也可能使用不同技能和工具。留下调用信息，是为了保存结果产生时的真实条件，而不是用技术名称替代责任判断。模型、工具或知识版本变化以后，历史运行仍能对应到当时配置，后续比较才有基础。

任务结束，实例可以停止，小队可以解散，运行记录及其证据关联仍要保留。后续验收、退回、回测和责任追溯都从这些记录起步。能够查到哪一步，只能由实际留存的证据决定；设置了运行编码，不等于已经具备完整的过程还原能力。

第十三章 行业罗盘任务怎样拆解与组队〔操作程序〕

本章概要 | 本章回答：行业罗盘任务怎样拆解与组队？

关键规则：

- 行业罗盘任务须先明确观察对象、资料范围、交付标准与约束，再拆成可独立执行和检查的工作单元，按能力、权限与可用状态配岗组队；
- 任务沿接单、规则匹配、组队、执行、验证、收口与归档七站推进，各单元按依赖并行，提出候选不等于通过检验，经营意义、统计口径与评分依据仍须接受审查；
- 出现分歧先澄清，确需修改再退回对应环节；
- 定义、用途或评分发生实质变化，须建立新版本并重做完整六审，超出规则或授权的事项交由具名真人处理。

**一般读者读到这里即可，正文细节供组织者、实践者、技术尽调方与 AI 审计者等专业人士查阅。*

第三部分要求微任务边界明确、结果可以独立检验。本章通过任务定义、工作单元和七站作业流程，说明这些要求怎样进入行业罗盘交付（参见第八章 8.3 节、第八章 8.4 节）。

一套行业罗盘即使由模型生成了初稿，也需要经过设计审查。我们要把对一门生意的理解，变成可以说明、可以评分、可以复核的因子体系：行业边界要清楚，输入资料要适用，因子要能解释经营，观察口径要可复算，评分规则还要能区分状态。若把这些判断一次塞进黑箱，最后即使得到一张完整表格，也说不清结论从哪里来，更不知道问题卡在哪一步。

帽子军团先把行业罗盘拆成边界清楚的工作单元，再从资源池选择合适的 AI 员工，启动 Agent 实例，组成只服务当前目标的专项小队。稳定岗位提供身份、能力表现和责任记录，本次任务再装配上下文、权限、知识、技能、工具与交付要求。任务结束，小队解散，岗位与记录继续保留。我们由此获得组队的灵活，也守住过程的连续。

13.1 行业罗盘任务的定义：输入资料、交付标准与约束

任务拆解之前，先要把“生成一个行业罗盘”说具体。行业名称只是入口，同一行业还可能包含业态、收入模式、经营范围和观察对象等方面的区别。边界不清，后续 Agent 就可能各自理解为单店、连锁网络、加盟体系或供应链管理；因子和评分规则即使都写得完整，也无法放进同一套罗盘。

一项能够进入执行的行业罗盘任务，至少要写清五件事：本次覆盖哪些经营活动；要形成怎样的行业观察与评分框架；允许使用哪些行业资料、经营记录、规则依据和既有研究；交付

包含哪些因子名称、定义、量化口径、单位、观察期间、评分规则与说明；以及权限、时限、资料边界、适用规则和升级条件。五项合在一起，才构成同一项任务。

输入资料的关键不在数量，而在来源与用途能够对应。行业背景材料帮助理解生意，却不能自动证明评分阈值合理；经营记录可以支持观察，还要说明对象、时间和口径；前版罗盘可以参照，版本变化后却不能继续照搬旧定义。任务开始时先确认资料范围和规则版本，是为了让所有工作单元面对同一个问题，执行途中不再悄悄移动目标。

交付标准也不能只写一句“提交完整罗盘”。所谓“完整”，是每项产出都有明确位置，名称、定义、口径和评分能够互相核对，关键判断找得到依据。任务开始时还要说明哪些结果需要独立验证，哪些分歧先按现有规则澄清，哪些事项超出规则后必须升级给具名真人。边界写在前面，后面的协作才有共同语言。

13.2 行业罗盘任务的拆解：因子识别、口径说明与评分规则工作单元

任务定义清楚以后，再把它拆成可以独立执行、独立检查的工作单元。最基本的三类工作是因子识别、口径说明和评分规则：因子识别解释一项经营事实为何值得观察；口径说明界定观察对象、计算方法、单位和期间；评分规则再说明观察结果如何按行业标准转成分值。它们彼此依赖，但不能压成一次含混判断。

工作单元可以用“小、窄、碎”检验。“小”表示一次只承担一个清楚可辨的交付目标；“窄”表示裁量限于预先确定的规则与授权范围，不能自行改目标、扩权限，遇到超出边界的情形应停止执行并升级处理；“碎”表示每项产出可以独立检验，关键步骤可以重新回放。拆分后的每个单元仍须保有完整含义，能够被单独退回、修正或重做。

以候选“门店密度”为例，提出它作为候选因子，需要完成以下几项工作：界定门店和商圈、选择观察期间、判断密度变化与经营的关系。第一项解释经营意义，第二项形成可复算口径，第三项需要证据支持评分方向。拆开之后，先澄清问题落在哪一层；只有口径明确需要修订时才退回口径单元。若经营意义本身不成立，后面的评分工作也没有继续的必要。

每个进入调度的工作单元都要带着输入、目标、输出、依赖和检查条件。因子识别从行业经营问题和已有证据出发，形成候选因子及理由；口径说明承接已确认的因子，形成定义、计算方式、单位与观察期间；评分规则再以确认后的定义和证据为基础，形成方向和区间。把输入、输出与验收标准一起写清，单项结果才能独立检查，才能成为下一环节可靠的输入。

13.3 AI 员工配置：按能力表现、权限与可用状态调配岗位

工作单元清楚以后，系统才知道该调什么岗位。因子识别需要理解行业经营，口径说明要处理定义、数据和计算，评分规则把事实观察转成稳定刻度。不同工作还可能需不同的资料权限、行业知识、工具条件和独立检查能力。只看岗位名称，或按部门平均派发，都无法完成这样的匹配。

配置岗位时，需要同时查看 Agent 能力表现、可调用知识、权限范围、历史记录和当前可用状态。能力表现要回答岗位过去在相关任务里实际做得怎样；可调用知识说明它能依据哪些行业和规则材料；权限决定可读取和操作的范围；历史记录暴露质量与偏差；当前可用状态则说明岗位是否处于暂停、待校准等状态，避免同一岗位既产出又审查自己的结果。因此，调度依据并非是一句“擅长研究”，而是一组事实。

模型、技能和工具属于本次任务配置，不等于岗位本身。正式组织口径是 3 类行业视角 + 3 个模型对冲：行业视角规定从什么经营立场理解生意，模型负责对同一问题形成可比较的底层判断。系统先根据工作单元确定所需视角，再选择适合的模型、技能和工具。

岗位选定后，系统为本次工作启动 Agent 实例。每次实例运行形成独立运行编码，并关联所属任务、目标、输入、权限、规则版本、输出和检查记录。同一岗位可以先后参加不同任务，同一任务也可以包含多个实例运行；发生重做时，前后运行各自保留，不能用一个任务号或最后结果覆盖每一次过程。

这一步把长期岗位与当次执行分开。稳定身份让能力、经验和责任可以积累；按任务启动实例，则把上下文与权限收进当前边界。动态组队要做的，不是临时找几个模型共同写答案，而是把已经定义的工作、适合的岗位和一次次运行条件准确接上。

13.4 AI 员工组队：专项小队的组建、分歧处理与协作解除

多个工作单元各自选岗后，还要组成一支专项小队。小队只为本次行业罗盘目标存在，不另造永久部门。成员保留原有身份编码、能力表现和真人认领关系，同时获得本次任务需要的上下文、帽子、权限、知识、技能、工具和交付标准。

调度 Agent 维护整项任务的连续性。它掌握总体目标、工作单元依赖和当前状态，确认前一步输出是否足以支持下一步，也在岗位或模型替换时完成交接。调度负责把路接起来，不负责任路上所有判断：业务结果仍由相应岗位形成，验证也不能交还给结果提供者自行完成。

分歧在行业研究中很正常。不同视角可能对经营意义、数据可靠性或评分方向作出不同判断。调度 Agent 先把分歧定位到具体工作单元，组织各方说明依据并澄清概念；只有确认内容确需修改时，才退回相应环节。证据不足就回到资料工作；定义、用途或评分发生实质改变，则建立新版本并重做完整六审。现有规则无法覆盖，或涉及权限与评分原则变化，则升级给具名真人裁量。调度不能为了赶进度自行改规则。

任务完成后，专项小队解除，岗位回到资源池等待下一次调用。协作关系结束，身份和历史不能随之清空。每次实例运行的运行编码、所属任务、输入输出、分歧处理、修订记录和最终交付继续保留，真人认领关系也保持连续。经验证的经验可以进入能力校准和规则修订，但单次合作并不会自动变成永久编制。

13.5 AI 员工七站作业流程：接单、规则匹配、组队、执行、验证、收口与归档

工作单元和专项小队需要一条共同任务线。行业罗盘沿接单、规则匹配、组队、执行、验证、收口与归档七个站点推进。七站的意义不是多设七道审批，而是把每一步接收什么、形成什么、满足什么条件才能继续，以及问题出现后回到哪里写清楚。

站点	主要工作	形成的结果	异常处理
接单	确认行业范围、任务目标、输入资料和预期交付	边界清楚的任务记录	范围或资料不清，退回补充
规则匹配	确认适用规则、版本、权限与升级条件	本次执行采用的规则和约束	规则缺口或权限冲突，提交处理
组队	按工作单元匹配岗位，配置上下文并启动 Agent 实例	专项小队、分工和运行编码	能力、权限或可用状态不满足，重新配置
执行	按依赖完成因子识别、口径说明和评分规则等工作	附带“来源”和“版本”的分项产出	输入冲突或前提变化，返回对应前序单元
验证	按检查条件核对分项结果	通过意见、异议或退回要求	不合格结果退回相应环节重做
收口	汇合已经通过验证的结果，处理未决事项	结构完整的交付内容	仍有未决分歧，不进入归档
归档	保存任务状态、输入输出、运行记录和责任关联	可供后续复核的任务档案	记录不完整，补齐后再完成归档

表 13-1：行业罗盘任务的七站作业

七站真正要解决的是状态可见。工作单元开始执行前，应知道依赖是否满足；结果进入收口前，应知道验证是否通过；发现问题时，先澄清问题所在，确需修订才退回具体环节，不能

笼统要求“全部重做”。同一任务的不同单元可以按依赖并行，但每项产出都要回到自己的输入、规则与运行编码。

验证站核对结果是否达到任务预设的检查条件，不替代六鉴与执衡。六鉴与执衡只审查本轮候选因子的设计、资格和组合；企业取数、实际评分、校准、部署及发布另属后续实施。收口不负责消灭所有差异，它要区分已经成立的内容、已澄清事项和明确需要修订的部分。若定义、用途或评分发生实质改变，应形成新版本并重做完整六审。归档则保留整条任务链，让任务结束后仍能找到每项输出对应的实例、处理过程和责任主体。

图 13-1 把七站流转、候选设计审查和过程记录放在同一条任务链上。

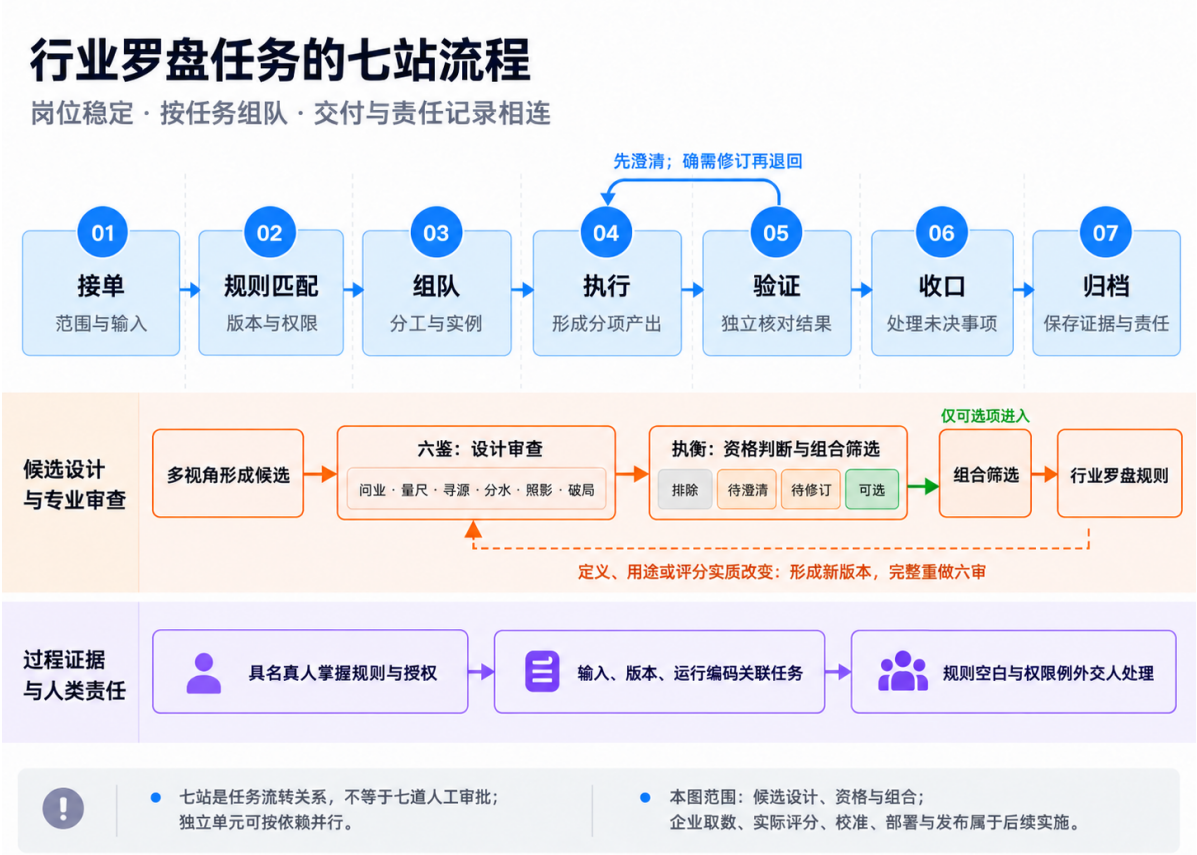


图 13-1：行业罗盘任务的七站流程

13.6 举例：连锁便利店行业罗盘——从一张任务单到一个候选

本章依次说明任务怎样定义、拆解、配岗、组队，又怎样沿七站推进。下面用一个例子把它们串起来，以“SIM-01”为共用任务标识，看一个行业罗盘怎样接单、怎样在协作中处理分歧、怎样接受检验。SIM 仅表示模拟，不是后台真实任务编号；V0.1、V0.2 也是本例的说明性版本，不是已批准的行业成果版本。

接单：先把“连锁便利店”说窄

连锁便利店可以指一家门店、一个区域网络，也可以指品牌总部或加盟体系；观察单位不同，日均销售额、毛利率、库存周转天数算的就不是同一件事。接单站因此先把范围收窄，再把输入、交付和约束写在同一张任务单上。规则匹配站确认本轮评分规则版本，并将统一商圈划定标准等未覆盖事项列为升级条件。

项目	本例设定
任务	SIM-01：生成连锁便利店的候选因子、量化口径与 1 至 3 分的评分规则
观察对象	连锁体系中的门店及其加盟网络，以月度经营事实为主；供应链融资与总部资本结构不在本次范围
输入	行业背景资料、公开经营常识与既有研究；本例不代入企业实测数据
交付	收益维度与风险维度候选因子，每项含定义、用途、统计对象、单位、期间、数据路径、评分规则及审查记录
约束	只做规则设计；企业取数、实际评分、校准与发布属于后续实施
升级条件	需要新增现行规则未覆盖的统计标准，或改变评分原则时，提交具名真人
本例跟踪	候选“门店密度”，自 V0.1 起

表 13-2：连锁便利店模拟任务单

组队与执行：多个视角同时开工

组队站里，按工作单元配岗，组成一支临时小队：A-调度掌握任务状态与依赖；B-行业分析、B-口径、B-评分分别承接三类工作单元；C-六鉴与执衡承担设计审查、资格判断与组合筛选；D-支撑维护资料权限、版本与记录关联。岗位进入任务时，系统为其启动 Agent 实例，每次实例运行生成独立的运行编码。

执行站里，三类行业视角与三个模型同时展开并给出判断，投资者视角再从生意、价格、风险三方面追问，收益与风险两侧的候选因子陆续形成，汇成本次任务的候选池。本例不逐项展开，只从池中选取一项跟踪：“门店密度”。

提出“门店密度”，出发点很直观：便利店做的是近距离生意，同一商圈里开着几家店，会影响每家店能分到多少客流。V0.1 因此只写了名称和一句用途——“反映门店所处商圈的竞争与覆盖情况”。

这句话里藏着三个没有回答的问题：计算的是本品牌门店、竞品门店，还是全部便利店？商圈按什么划，由谁、在什么时候划？门店多了，对这家店是好事还是坏事？

到这里，七站推进到第四站：任务有边界，工作有分解，岗位有编码，候选有版本。V0.1带着三个问题来到B-评分（见第十四章14.6节）。

第十四章 长期岗位怎样支持动态协作〔操作程序〕

本章概要 | 本章回答：长期岗位怎样支持动态协作？

关键规则：

- 部门归属线维护岗位的长期身份与责任，任务流程线按交付需要跨部门组队，临时协作不改变岗位归属和具名认领关系；
- 小队须覆盖调度、执行、管控与支撑四类职能：调度负责衔接和整体交付协调，不能代替专业判断或越过独立验收；执行岗位负责分项产出，不能自行批准结果；管控岗位独立检查权限、证据与规则边界，不能因交付进度撤销异议或降低检查条件；支撑岗位保障运行，不能以技术可用代替业务通过；
- 具名真人按授权承接规则、权限与升级事项；
- 整体协调与分项责任分别留痕，小队解散后仍可追溯。

**一般读者读到这里即可，正文细节供组织者、实践者、技术尽调方与 AI 审计者等专业人士查阅。*

第三部分提出，任务协作与规则监督需要分别组织。本章通过部门归属与任务编组的双线安排，说明长期职责如何保持连续、当次协作如何灵活展开（参见第九章 9.1 节）。

行业罗盘被拆成可执行、可检查的工作单元以后，还要把分散结果重新组织成完整交付。资料、行业分析、评分判断、管控检查和技术支撑可能来自不同部门，参与岗位也会随任务改变。只看临时小队，长期身份和责任记录容易被遮住；只看固定部门，跨部门能力又很难及时围绕同一目标集结。

双线协作把这两种需要接起来。部门归属线维护 AI 岗位的长期身份、能力表现和责任归属；任务流程线围绕当前目标调用岗位并启动 Agent 实例。身份编码、岗位台账与运行编码连接两条线：每次实例运行产生运行编码并关联所属任务，输入、判断、检查与交付因此能够回到相应岗位和具名真人。小队可以灵活变化，责任不会随解散而断掉。

14.1 部门归属线与任务流程线

部门归属线管理相对稳定的岗位关系。岗位属于哪个部门、具备哪些能力、做过什么工作、由哪位具名真人认领，都在这里持续记录。归属部门并不限制岗位只能处理本部门任务，它提供的是长期身份和管理基础，让组织能够维护规则与权限、更新能力表现，并把多次任务串成连续历史。

任务流程线只围绕一次具体交付组织工作。行业罗盘进入执行后，系统依据目标、输入、交付标准和约束，从不同部门调用岗位，按需要启动 Agent 实例。每次实例运行获得独立运行编码，并关联所属任务、本次上下文、权限、知识、技能、工具和检查要求。任务结束，临时协作关系解除，岗位回到原有归属，身份编码和历史记录继续保留。

运行编码是两条线之间的桥。身份编码说明长期岗位是谁，运行编码说明这个岗位某一次实例运行做了什么，并关联所属任务。一次行业罗盘任务可以调用多个岗位，也可以包含同一岗位的多次运行；任务号标识整项任务，不能替代各次实例的运行编码。沿流程线可以查看每个环节发生了什么，沿岗位台账也能回看某个 Agent 参加过哪些任务、当时承担什么角色。

对于 5000 名 AI 员工，这一区分尤其重要。只看部门名称，说不清一次交付究竟由谁完成；只看任务清单，又看不见参与岗位长期由谁管理。双线协作分别回答“这个岗位是谁”和“它在这次任务中做了什么”，再由身份编码与运行编码把两个答案准确连接。

14.2 临时小队覆盖四类职能，长期责任记录保持连续

专项小队是一组围绕任务配置的责任组合，不能只列业务执行岗位。行业罗盘既要生产内容，也要组织依赖、检查边界并维护运行条件。因此，每次组队都要根据任务实际覆盖调度管理、业务执行、管控合规和部署支撑四类职能，让结果有人做、过程有人接、边界有人查、条件有人保。

职能	在任务中的工作	不可替代的边界
调度管理 (Administration)	组织组队，维护前后依赖、交付衔接和退回安排	不代替业务执行形成专业判断，不独立验收
业务执行 (Business)	完成资料整理、行业分析、评分判断和分项交付	不能独自完成对自身结果的最终验证
管控合规 (Control)	检查权限、证据、规则及风险边界，保留检查意见	不因交付进度撤销异议或降低检查条件
部署支撑 (Deployment)	提供模型、数据、工具、运行环境和技术保障	掌握系统权限不等于拥有业务判断权或规则修改权

表 14-1：四类职能在动态协作中的分工

四类职能可以在同一阶段并行。业务执行形成分析与评分规则设计时，管控合规同步核对证据和规则边界，部署支撑处理数据与工具条件，调度管理维护输入、输出和交接。职能覆盖是一项责任要求，不预设必须配置多少岗位；是否允许兼任、哪些条件下必须分离，仍服从

正式规则。但两条底线不能退：结果提供者不取代独立验收，技术可行也不代替业务和管控判断。

临时组队改变的是本次参与关系，岗位的长期身份不变。岗位进入小队后继续保留身份编码、部门归属和认领关系；每次 Agent 实例运行则形成运行编码，并关联所属任务、本次角色、权限、输入、输出和检查结果。同一任务若有返工、复核或版本变化，各次运行分别记录，不能压成一条看不出前后的结果。

完整交付也要有明确的整体认领。小队不能用“共同完成”掩盖协调主体，协调主体也不能把所有分项责任收拢到自己名下。资料由谁整理、行业判断由谁形成、评分结果由谁提交、检查意见由谁作出，都需要分别留下。整体认领回答谁把交付组织起来，分项记录回答每个判断怎样形成，两者合起来才是一条完整责任链。

图 14-1 汇总了四类职能、具名认领与任务流转之间的责任关系。

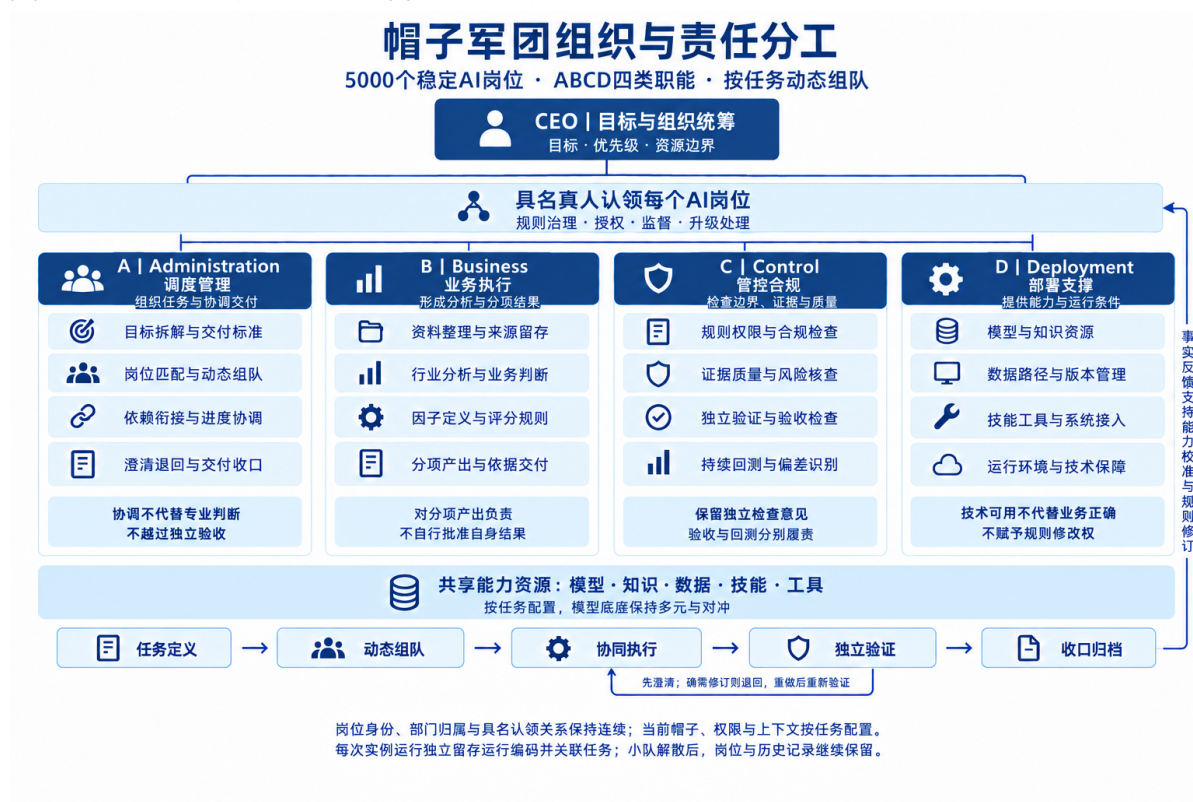


图 14-1：帽子军团的组织与责任分工

14.3 调度 Agent 负责日常组队、衔接与整体交付协调

调度 Agent 的职责，是让任务在既定目标和规则下连续运行。任务开始时，它依据行业范围、交付边界、适用版本和所需职能组织组队，确认每位成员的本次角色。岗位能否进入小队，既看部门和职责，也看能力表现、可用状态、所需权限及与当前任务的匹配程度。

小队组成后，调度 Agent 逐项写清成员接收什么、交付什么、可以使用哪些权限，以及结果要通过哪些检查。有前后依赖的工作，前项输出达到可用状态，后项才具备开始条件。规则版本、输入资料或交付标准一旦变化，调度还要找出受影响的环节，防止不同版本的结果在收口时混在一起。

执行过程中，调度 Agent 持续维护任务关系：跟进分项工作的状态与去向，补上交接断点，组织规则范围内的澄清和复核。发现疑问时先定位原因、再澄清依据，只有确认内容需要修订时才退回相应岗位，并带回失败原因、所用版本和影响范围。修订形成新的运行记录，再重新进入流程，并继续接受相应检查。

调度权不能膨胀为判断权。调度 Agent 可以要求补充依据、组织不同意见复核，也可以暂停尚未满足前提的工作；它无权自行改变帽子、扩大岗位权限或修改评分口径。它负责把分项结果组织成完整交付，却不能改写专业岗位的判断，更不能用交付时间越过独立验收或豁免管控合规检查。

14.4 具名真人负责规则 授权与超出规则的升级处理

每个 Agent 都由具名真人认领。认领的意义，是把规则、授权和规则之外的处理落到明确主体，而非让真人逐项代替 Agent 操作。认领者需要知道 Agent 依据哪些规则工作，可以调用什么资源，遇到什么情形必须停止执行并升级处理，以及规则变化会影响哪些任务。

规则已覆盖且权限充分的日常任务，交给 Agent 依规完成。调度 Agent 组织组队和衔接，业务岗位形成分项结果，管控合规与验收岗位完成检查。若真人因为认领关系重新审批每一次资料整理、分析和评分，海量协作仍会退回人工指挥，Agent 在规则内的运行边界也会被抹掉。

这里的评分，是规则内的计算和评分建议。对外生效的企业最终评分、合规批准、投资及适当性判断，仍由具备相应授权与资质的人作出或确认；内部检查不能代替最终批准。

真人主要处理规则之外的事项。新的业务情形超出现有规则、岗位权限不足，或多个合规要求需要自由裁量时，Agent 应停止自行延伸，把事实、现有规则、分歧和潜在影响一并提交。具名真人据此决定继续、修订还是终止，并把决定写回所属任务与相关运行记录。

认领还要求真人持续了解规则执行、留痕和回测状态。这里的“了解”并非逐项接管，而是判断现有边界是否仍然有效，是否出现反复失败、规避规则或环境变化。真人守住的是“Agent 始终有人认领、规则能够解释、异常可以升级”的边界。

高后果动作的叫停与恢复，应遵循第七章 7.2 节所述处置顺序，并依据组织正式授权表确定处置人员与批准主体。认领关系、专业判断权、叫停权和恢复批准权应分别说明。

14.5 整体交付与分项责任并存，验收和管控保持独立

一份行业罗盘交付包含许多分项判断。调度 Agent 组织整体交付，不会让资料、分析、评分和检查岗位的责任消失。每个岗位仍对自己的输入、判断、结果和适用规则负责；调度记录则说明这些结果怎样被连接、怎样澄清或修订、又怎样重新回到交付中。

具名真人的责任同样分层。真人对所认领 Agent 的规则、授权边界和升级事项负责，却不能用一条认领记录代替本次任务的运行记录。一次交付需要同时看见调度主体、参与岗位、每次实例运行的运行编码、所属任务、适用版本、检查意见和人工升级事项，日常执行、整体协调与最终自由裁量才能各归其位。

管控合规保持相对独立，是对协调权的必要限制。调度 Agent 可以要求检查意见指向具体规则、证据或风险，也可以组织补充材料和重新执行，却不能因为进度压力删除尚未解决的异议。执行岗位可以解释判断，不能自行宣布通过；部署支撑可以证明系统或工具按要求运行，不能替代业务或合规结论。

AI 管控合规的职责边界见第十一章 11.3 节。独立质量检查、合规或法律判断以及对外生效的批准，须分别留存依据和责任主体。

独立验收也不能由结果提供者包办。验收依据预设标准判断结果能否进入下一环节，发现疑问先澄清，明确需要修订才回到对应岗位。六鉴与执衡将在下一章展开，但范围只限本轮候选因子的设计审查、资格判断和组合筛选；企业取数、实际评分、校准、部署与发布属于后续实施。资格分为排除、待澄清、待修订和可选，可选不代表最终入选或得到企业验证。这里先守住权力边界：协调不等于批准，执行不等于验收，认领也不会抹去分项责任。

14.6 举例：连锁便利店行业罗盘——分歧出现时，谁澄清、谁修改、谁拍板

B-口径尚在界定统计对象时，B-评分一侧已先卡住：要写 1 至 3 分，先得知道密度高算好还是算坏，而三类行业视角给出的回答并不一致：

正常经营者视角认为，同一商圈里便利店越多，客流被分走越多，密度应读作压力。

老炮视角指出，连锁品牌常在成熟商圈接连增开自家门店，本品牌门店多，可能恰恰说明商圈被验证过、配送更集约。

骗子视角不谈方向，只问一件事：商圈范围如果能在看到数据之后再划，同一家店把范围画小一点、竞品就少几家，经营没有任何变化，读数却能随意调整，这个因子还靠得住吗？

每类视角由三个模型分别给出判断，再按视角汇总；同一视角内部，三个模型的输出也没有收敛：一个偏向竞争解释，一个偏向规模解释，一个认为定义不足以判断。分歧本身不是问题，要紧的是它落在哪里。

先定位，再澄清，明确需修改才退回

A-调度不投票，也不替任何一方裁决。它先把分歧拆到具体工作单元：方向之争属于因子识别，边界之争属于口径说明；再请 B-行业分析把三种读法分别写成可核对的经营机制——竞争解释依据什么，规模解释依据什么，操纵质疑指向哪个字段。依据摊开后，澄清的结果是：V0.1 把两件事装进了一个名字——竞品带来的分流，与本品牌增开门店带来的覆盖。两者方向可能相反，混在一起的读数无法唯一分档，于是退回 B-口径。

V0.2 把统计对象限定为同一商圈内非本品牌便利店的门店数，本品牌自身的门店数另作观察，不在本例展开；同时要求商圈边界在取数前固定并留存版本，门店清单注明来源与版本，观察期间统一。评分方向不预设，B-评分登记为“待区分证据”，不自拟阈值。D-支撑同步将 V0.1 与 V0.2 分别存档，保存各方意见和每次实例运行的运行编码，并在记录中标明：竞品门店清单的数据路径尚未接入，属于待补的真实数据缺口。

还有一件事小队自己定不了：统一的商圈划定标准，现行规则没有覆盖。B-口径不能为全行业新设标准，A-调度也不能为赶进度代为决定，这一缺口登记为升级事项，提交具名真人。本例不设定批准人和批准结果；在真人裁量之前，V0.2 只写“边界须事先固定并留痕”这一控制要求，不写具体半径。

修订完成后，V0.2 提交 C-六鉴与执衡。六审意见形成前，执衡暂不判定该候选资格（见第十五章 15.6 节）。

环节	模拟角色	本例中做了什么	权限边界（本例不设定批准人与结果）
定位	A-调度	将分歧拆到因子识别与口径两个单元，组织澄清	不能替任何一方裁决，也不能替检验一方下结论
修订	B-行业分析 B-口径 B-评分	B-行业分析写明三种读法的依据；B-口径形成 V0.2，限定统计对象，要求边界事先固定、清单与期间留版；B-评分登记“待区分证据”	不能批准自己的修订，也不能自拟阈值
留痕	D-支撑	保存 V0.1、V0.2、各方意见及各实例运行编码	不能把“系统已保存”写成“业务已通过”
升级	具名真人	承接“统一商圈划定标准”这一规则缺口	本例不编造批准人与批准结果
复审	C-六鉴与执衡	接收 V0.2，准备完整六审	不能只复查最初出问题的一项

表 14-2：连锁便利店模拟任务的岗位分工

小队会随任务解散，记录却不随之消失。表中各岗位的每次运行都挂着运行编码，升级事项则落到具名真人名下。因此，谁提出问题、谁修改定义、谁承接升级，事后都能分开认定。这种认定依靠的是每次运行的实际记录；具名真人的认领关系只说明岗位由谁负责，不能代替这些记录。

第十五章 一项交付怎样经过检验与修订〔操作程序〕

本章概要 | 本章回答：一项交付怎样经过检验与修订？

关键规则：

- 验收须依据事先约定的标准，写明检查对象、结论及允许进入的下一环节；
- 六鉴分别审查候选设计，执衡据此判断资格与组合；完成六审不等于通过，可选不等于入选，入选也不等于完成企业验证；
- 未决事项须注明缺口与承接安排，随版本和审查意见一并归档；
- 定义、用途或评分规则实质改变，须以新版本重做完整六审；
- 后续事实出现偏差，先区分资料、能力与规则原因，再作相应修订，保留历史依据。

**一般读者读到这里即可，正文细节供组织者、实践者、技术尽调方与 AI 审计者等专业人士查阅。*

第三部分将独立检验与全过程追溯作为可信交付的条件。本章进一步说明，这些要求如何落实到验收环节、运行记录和责任追溯之中（参见第十章 10.1 节、第十章 10.3 节）。

一份候选因子表写完，工作只完成了一半。还要检查它是否说清了经营问题、能否计算、有没有数据路径，以及放进组合后究竟增加了什么。只有这些问题得到回答，产出才能进入下一环节。即使通过当次审查，后续出现的新事实也可能要求我们重新判断。

本章先划清本轮工作的范围：六鉴与执衡负责候选因子的设计审查、资格判断和组合筛选。企业取数、实际评分、校准、部署和发布另属后续实施。后文谈回测与实际得分核查，是说明成果怎样在后续使用中接受检验，这不代表把全部实施责任归给六鉴与执衡。

15.1 验收是任务流转的准入关口

验收先回答一个基础问题：交来的东西，是否完成了原先约定的工作？任务目标、行业范围、输入资料、规则版本、权限和交付标准，都是判断依据。这些条件应在开工前讲清，不能等结果出来后再按结果修改标准。

因子识别、口径说明、数据路径判断和评分规则形成，交付物各不相同，检查时却有共同要求：回应了原任务，采用了适用规则，定义清楚，证据可查，能够复核。上游资料或规则已变化，下游仍沿用旧版，即使表格完整，也不能按新条件通过验收。

验收还要看结果是怎样做出来的。本次用了哪些输入、知识、模型和工具，是否在授权范围内，规定的检查是否完成，都要有记录。执行者可以解释自己的结果，独立检查仍须由承担相应职责的岗位完成。具备生成能力，并不等同于取得批准权限。

验收结论要与任务、产出版本、适用规则和检查记录一起保存，并写明允许进入哪一环节。进入候选组合、形成交付与归档，各有自己的条件。未满足条件时，先说明问题与依据，再确定澄清、修订或排除；需要退回的，列明修改范围和重新提交要求。规则变更、权限例外与自由裁量，交由具名真人处理。

这样，后续环节接收到的就不只是一个“通过”标记，而是一项来路清楚、状态明确的产出。进度可以催促工作，不能替代交付标准。

15.2 “6+1”检验视角：六鉴与执衡的分工、检验产出及不合格退回

一个因子能否进入罗盘，不能只问“看起来重不重要”。经营意义、可计算性、数据条件、区分依据、全池关系和抗操纵性，需要分别检查。六鉴各守一项检验责任，执衡据此判断候选资格，再选择组合。七个位置的职责必须分清。

名称	检验问题	对应产出
问业	经营上是否有意义	经营问题与机制图谱
量尺	能否唯一计算和评分	定义、参数与边界表
寻源	是否有可信数据路径	数据路径与条件表
分水	是否具备区分依据	区分假设与证据表
照影	相对全池新增什么信息	全池关系与映射表
破局	能否在状态未改善时刷分	攻击、防御与依赖表
执衡	谁具备资格以及如何组合	资格台账与最终方案

表 15-1：六鉴与执衡的检查问题及产出

六鉴是问业、量尺、寻源、分水、照影、破局六个审查视角的名称；六审是这六项审查的工作合称。执衡随后判断资格与组合，不作为“第七鉴”。

问业先把因子带回生意本身：它对应什么经营问题，通过什么机制影响结果？名称再专业，也要说得出这层联系。经营问题与机制图谱把联系展开，供后续定义、数据与评分讨论使用。量尺要让同一份材料只得出一个计算结果、一档评分。定义、分子分母、单位、观察期间和异常值处理须对齐；评分区间的端点不能重叠或遗漏，小数怎样处理也要写清。净利润等分

母为零或负数时，须有明确规则，不能直接套用正值条件下的分档。定义、参数与边界表须足够完整且无歧义。

寻源追问数据从哪里来、怎样取得、能覆盖什么范围。一个因子再重要，如果来源说不清或取数条件不具备，就缺少计算基础。数据路径与条件表需要把来源、取得条件和限制放在一起；列出一个网址，不等于数据路径已经成立。

分水检查数值变化是否对应有意义的经营差异。数据有波动，只说明它在变化，未必说明它能区分好坏。因此，区分假设与证据表要分开写：我们认为它能解释什么，已有事实支持到哪里，还有什么需要验证。

照影把候选因子放回全池，问它究竟增加了什么。全池关系与映射表逐一说明候选与已有因子之间是重复、包含、依赖还是互补的关系，以便帮助组合留下真正新增的信息。

破局尝试从规则的缝隙进入：经营状态没有改善，能否只改统计范围、记录方式或依赖关系就刷高分？提前找出这些路径，才知道哪里需要防护。攻击、防御与依赖表同时记录刷分办法、阻断条件以及防护所依赖的其他规则。

执衡接收六鉴的完整意见，先将候选列为“排除、待澄清、待修订、可选”四类：不具备资格的予以排除，材料或理解尚不清楚的先澄清，需要调整设计的待修订，满足本轮资格条件的列为可选。资格台账保留依据与分歧，最终方案另行说明从可选池中选了谁、怎样组合以及为什么。可选不等于最终入选，最终入选也不等同于已经完成企业验证。

图 15-1 把候选形成、六鉴审查与执衡筛选连接起来，便于回看各方意见怎样进入最终方案。

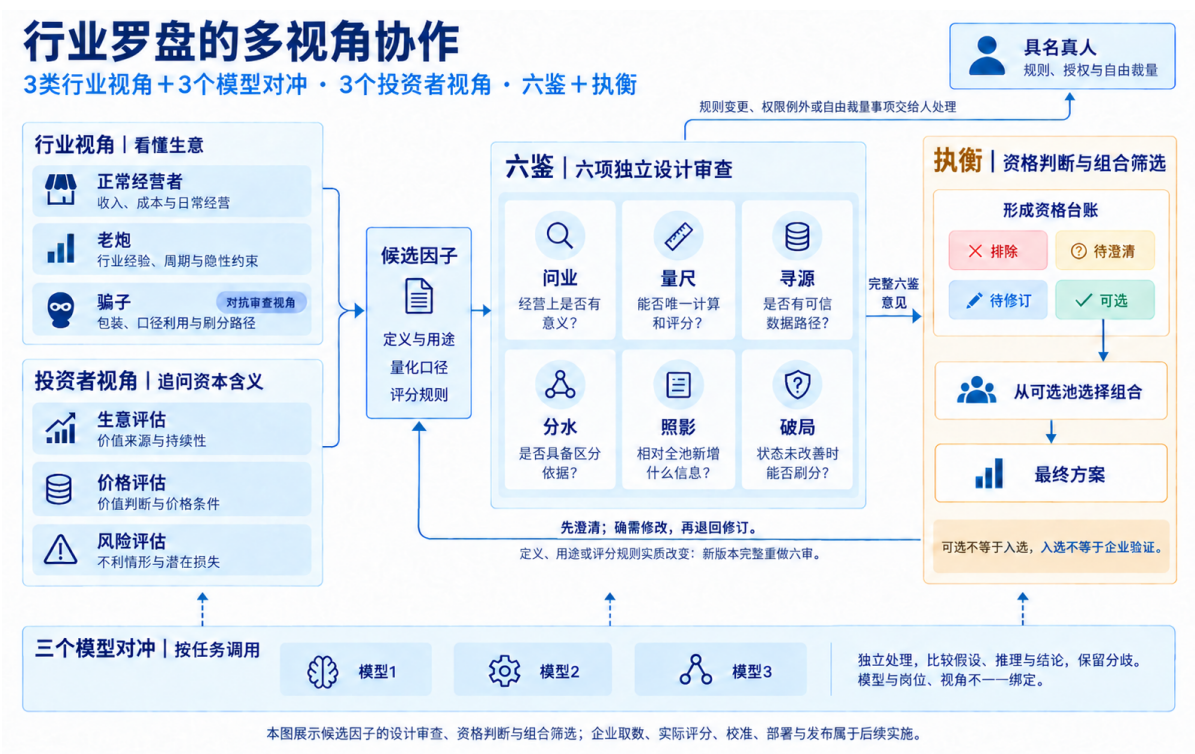


图 15-1：行业罗盘的多视角协作与设计审查

有分歧，先核对事实、定义和各方意见，不能一发现分歧就退回重做。澄清后确需修改，再退回相应工作单元，并留下问题、依据、原版本和重新提交条件。候选的定义、用途或评分规则发生实质变化，须形成新版本，重新接受完整六审；不能只复查最初发现问题的一项。原意见、修订内容和新一轮结论都要保留，再由执衡更新资格并审视组合。

遇到新行业情形、权限例外或规则本身需要改变，六鉴与执衡应把争议和证据讲清，升级给具名真人裁量。专业检验要能提出挑战，最终裁量也必须落在既定授权之内。

15.3 沿运行记录核查输入、产出、规则与责任

验收说明结果可以走到哪里，追溯说明结果怎样来到这里。两者要指向同一套记录。每次 Agent 实例运行形成相应运行编码，并关联所属任务、输入输出、调用、规则版本、检查结果与时间信息；一个任务下可以有多次实例运行，不能把任务编号当成单次运行的追溯编码。

复核一项结果，先找到对应运行，再逐步看输入资料、当时的定义和评分规则、中间产出、检验意见，以及是否发生澄清、退回和重做。之后资料更新、规则修改，都不应覆盖这次运行的原始条件，否则历史结论就无从解释。

运行记录再向上连接稳定的 AI 岗位。身份编码（Agent ID）告诉我们哪个岗位参与，Agent 能力表现说明当时据什么条件调用它。长期身份与单次执行分别留存，才能在小队解散后继续查到任务和责任。

沿岗位记录，还应找到真人认领关系及本次授权。参与者、检查者、任务负责人和最终责任人，各自能做什么，要分别核对。出现在名单上不代表有批准权，负责协调也不代表有权改规则。追溯要回答的是：谁在什么权限内作了什么决定。

状态也要有对应证据。写“已完成”，应能找到交付和收口记录；写“已退回”，应能看到问题及重新提交条件；写“待处理”，应说明卡在哪里、由谁承接。最新状态可以更新，此前的异议、退回和版本变化仍须保留。

运行编码只是查找入口。输入、调用或检验记录没有保存，编码再齐全也还原不了过程。记录支持到哪一步，就说明到哪一步；缺少什么，要明确写出。一个“验收通过”的标记不能补上证据缺口。

15.4 事实回测推动能力校准与规则修订

候选设计通过审查，只能说明它满足当时的资格与组合条件。进入后续企业实施后，取数、实际评分和校准会带来新的事实，原来的假设也可能遇到反例。事实回测就在这一阶段检查：已经交付的判断，是否仍能解释真实经营。

回测先锁定对象与版本，再对应事实来源、样本范围、观察期间、原规则与原结论。条件不同，不能急于把新旧结果的差异当成支持或反驳。回测频率、样本和阈值应有相应专业依据，不能用一套数字覆盖所有行业。

发现差异后，先找原因。原事实条件变化了，应更新输入；已有信息被遗漏了，应检查 Agent 能力与调用条件；执行完全合规却仍产生偏差，则可能是规则留下了空白。原因没分清，改动就容易落错地方。

各类原因要分别处理。更新资料、校准能力、修改规则，不能互相替代。规则变更须说明原因、新旧版本、生效时间和影响范围，并与依赖旧规则的产出核对。若候选定义、用途或评分规则发生实质改变，应以新版本重新六审，再更新资格和组合；企业实施端据此处理受影响的评分。

改完之后，还要用同类案例看原问题是否被阻断，有没有出现新的盲区。调整前后的输入要求、能力配置、规则、检验意见与结果同时保留，才能知道改动真正解决了什么。

经验证的经验再进入下一轮任务，用于校准能力表现、任务模板和检查要求；尚有争议的判断则继续保留依据与边界。历史成绩可以帮助调度，但不能成为免检资格，每次运行仍要接受当次检验。

15.5 授权核查、版本管理与记录完整性

记录里可能有业务资料、权限信息和真人信息，核查须按授权开展。白皮书说明责任怎样关联，并不公开敏感任务和个人信息。获授权的检查者应取得完成核查所需的证据，公开范围另按相应规则确定。

规则、能力与成果要分别管理版本。规则说明当时允许什么、怎样计算和检查；能力记录岗位当时的调用条件；成果说明交付内容依据什么形成。三者相互关联，不能用一个成果文件版本号代替全部变化记录。

重要变更要留下内容、原因、生效时间、适用范围和责任关系。复核者才知道历史结果还能不能用、哪些任务需要重查。旧文件要保留，并且要说清为什么改。

记录是否完整，要逐项看证据：输入可查吗，调用有记录吗，六鉴与执衡的意见留存了吗，澄清与返工接得上吗，裁量在授权内吗？这些答案共同决定能追溯多远，界面状态和最终文件只能说明其中一部分。

没有保存的资料，就说明缺什么、影响哪段判断；恢复不了的历史版本，就保留这一限制。缺口也告诉我们下一轮应补强什么留存要求。

任务经验进入长期能力库，也要经过授权和版本管理。新知识、检查方法与风险提示，应先确认来源、适用范围、验证状态和敏感信息边界，再决定哪些可以复用。

15.6 举例：连锁便利店行业罗盘——V0.2 经过六审，为什么列为“待澄清”

V0.2 改动了统计对象，用途也随之收窄。这类改动已触及因子定义，按本轮规则，须以新版本提交完整六审。

名称	检验问题	本例意见
问业	经营上是否有意义	需说明竞品分流的机制影响的是客流、客单还是品类结构
量尺	能否唯一计算和评分	统计对象方向已明确；边界须事先固定“已写明”，但具体划定标准待升级裁量，暂不能完整复算
寻源	是否有可信数据路径	竞品门店清单从哪里来、多久更新、覆盖是否完整，尚未落实
分水	是否具备区分依据	密度高低是否对应经营好坏，尚无事实支持；方向不预设
照影	相对全池新增什么信息	分流的后果已部分体现在日均销售额的变化上，过密布点的代价可见于闭店率与新店经营盈亏平衡周期；密度作为“原因层”信息新增了什么，需要说明
破局	能否在状态未改善时刷分	边界事先固定、变更留痕后，改画商圈的空间收窄；但清单来源若可自选，仍有漏洞
执衡	谁具备资格以及如何组合	未见应予排除的根本缺陷，但多项证据待补、增量存疑，不具备进入可选池的条件，列为“待澄清”

表 15-2：V0.2 的六审意见与执衡结论（本例设定）

验收验了什么

本次模拟验收的对象是修订过程本身：V0.1 的问题能指向明确字段，V0.2 写清了改动的控制要求，新旧版本能够一一对应，六审意见完整留存。它确认的是修订过程有据可查；候选能否由执衡列为可选、选入组合，乃至完成企业评分，各有其后续条件。

收口与归档：停在哪里，也要写清

收口站把三类内容分开写入交付：已经成立的（统计对象、边界固定要求），已经澄清的（方向之争源于定义混装），仍待补的（划定标准、数据路径、区分证据、增量说明）。“门店密度”以“待澄清”离开本次任务，不进入组合讨论。

归档站把两个版本、分歧记录、升级事项与六审意见关联到 SIM-01 下的各个运行编码。小队解散了，这条记录仍在；下一轮任务补齐证据时，能从它停下的地方接着做，而不必从头猜当时为什么没过。

追溯与回测：后续实施接手下一段检验

候选日后若补齐证据、进入企业实施，回测会接着检验：数据与判断不一致时，先查原因、再作修订，并沿运行编码回溯全过程。

这条模拟链的结论是具体的：问题要能定位，修改要形成版本，验收要说清验收了什么。一个候选没有走到最后，不说明流程失败；恰恰因为它停下的位置和原因都写得清楚，下一步该验证哪里、澄清哪里、修改哪里，才有了答案。

5000 名 AI 员工如何协同上岗，最终要体现在具体任务与交付中。本部分说明岗位怎样按任务调用能力、动态组队，完成行业分析、因子定义与评分规则设计，再通过独立审查和组合筛选形成成果，落实相应责任。连锁便利店说明性模拟在第十三、十四、十五章沿同一张任务单接续展开：先拆解任务，再定位和处理分歧，最后让修订接受检验并留下记录。本例特意选取一个未走到最后的候选，用来展示流程在哪里停、为何停。本轮已形成覆盖 500 个行业的行业罗盘规则成果；其结构和使用条件已在第二部分说明。下一部分以三个品类精读及附录导航展示这些成果怎样被阅读，既让读者看到选择了什么，也保留未完成企业取数、实际评分和验证的边界。

第五部分 案例成果

本部分精读连锁便利店、川渝火锅门店和现制茶饮门店，示范一套行业罗盘从哪里读起：先确认经营范围，再沿收益与风险查看指标、因子及其证据要求。便利店看高频交易与加盟网络，火锅看桌台产出与预收资金，茶饮看出杯、渠道和收现；相近的生意，也需要不同的观察变量。

三个案例与附录一的 17 个补充案例，合计 20 个展开阅读的窗口，不是统计抽样。案例呈现指标与因子结构，不代表已经取得具体企业数据或完成实际评分。读懂结构后，仍须沿 500 行业成果入口核对对应行业、规则版本和当前开放内容，再判断能否用于具体对象。

第十六章 三个品类精读与案例阅读方法

本章以连锁便利店、川渝火锅门店与现制茶饮门店三个品类说明行业罗盘的读法：先看品类简介，再看收益维度与风险维度下的指标与因子，最后分维度简析。连锁便利店承接第十三章 13.6 节、第十四章 14.6 节、第十五章 15.6 节的模拟任务。三个品类仅展示结构，不展开各因子的量化口径与评分标准。

16.1 三个品类精读

连锁便利店 十个指标 十七项因子

便利店交易高频、小额、分散，经营结果受品类结构、夜间经营、库存周转、收入勾稽与加盟体系等多方面影响，适合用于说明行业罗盘的组织方式。

维度	指标	因子
收益维度	门店销售规模与增长	日均销售额 连带件数（单笔平均购买件数）
收益维度	毛利结构与鲜食贡献	综合毛利率 鲜食销售占比 烟草销售占比
收益维度	夜间经营贡献	夜间销售占比
收益维度	库存与营运资金效率	库存周转天数
收益维度	加盟体系健康	加盟店存活率 加盟商续约率
风险维度	租金与固定成本	租售比 固定成本负担

维度	指标	因子
风险维度	损耗与品控	鲜食报损率 全店综合损耗率
风险维度	夜间运营风险	夜班到岗率
风险维度	收入真实性与内控	POS 与支付入账差额率
风险维度	加盟与扩张风险	闭店率 新店经营盈亏平衡周期

表 16-1：连锁便利店的指标与因子

收益维度：从卖得多，到赚得住

便利店面积小、客单低，收益靠高频交易累积，单看营业额无法判断经营质量。规模之外须同时看单笔交易的厚度，因此日均销售额与连带件数并列；利润主要来自鲜食等高毛利品类，烟草也是重要参考品类，三者放在一起才看得清利润结构；24 小时营业是业态特征，夜间时段对销售的支撑决定延时经营是否成立；商品品类多、资金主要压在货上，故看库存周转；连锁扩张主要依靠加盟，加盟店能否存活、加盟商是否续约，决定门店网络能否复制。

风险维度：成本、损耗与扩张

便利店生意模式较为单一，风险维度与收益维度相互对照。便利店的租金与人工支出为刚性成本，而销售随商圈波动，故以租售比与固定成本负担衡量盈亏线的位置；鲜食卖得越多，报损压力越大，损耗与品控是鲜食贡献的另一面；夜间营业依赖排班落实，夜班人手不足会同时放大安全、服务与收银风险；交易小额分散、现金与多渠道支付并存，收银记录与实际到账能否勾稽，是识别收入虚增或漏记最直接的检查；加盟扩张的代价最终体现为闭店，新店盈亏平衡周期则反映扩张节奏是否超出单店的造血能力。上述因子风险越高，得分越低。

表中看不到的：一个候选为什么不在这里

第四部分的说明性模拟任务 SIM-01 中，“门店密度”因证据不足停留在“待澄清”，未进入可选池；该模拟仅说明候选与成果的区别，并非实际生成记录。商圈竞争的影响可通过其他因子，如日均销售额、闭店率与新店经营盈亏平衡周期观察。

川渝火锅门店 十二个指标 十六项因子

川渝火锅门店以川渝麻辣锅底和鲜切食材为特色，收益取决于翻台率、人均消费与锅底毛利，团购外卖和储值是重要促销及现金流改善手段，主料价格波动、租金人力刚性成本及桌台收入真实性是核心关注点。

维度	指标	因子
收益维度	翻台与桌台产能	翻台率
收益维度	客单与毛利结构	人均消费 锅底毛利率
收益维度	获客与渠道	团购券核销率
收益维度	复购与会员	复购率 会员消费占比
收益维度	采购与成本效率	主料采购价差 全店综合损耗率
收益维度	增长与收现	销售额稳定性
风险维度	预收与退款	储值余额/月营业额倍比
风险维度	平台依赖与账期	团购/外卖渠道收入占比
风险维度	主料价格与供应	前三大供应商集中度
风险维度	租金与固定成本	租售比 刚性成本率
风险维度	合规与溯源	主料溯源完整度
风险维度	耗用勾稽与现金核验	现金流质量

表 16-2：川渝火锅门店的指标与因子

收益维度：桌台产出与客群黏性

火锅是重资产堂食，桌台数量一般较为固定，收益上限主要由翻台率决定，结合人均消费即可测算单桌收入，结合锅底毛利率可大致推测出真实收益水平；团购、储值是主要引流手段，但只有核销才转为实际收入；复购与会员消费决定获客成本能否被摊薄；主料采购和鲜切食材的损耗直接侵蚀毛利，采购与成本效率因此归入收益维度；销售额稳定性用于识别依赖节日与促销的短期繁荣。

风险维度：预收、供应、固定成本与现金核验

火锅的风险集中在预收款项、渠道集中、供应波动、刚性成本、合规溯源与收入真实性。团购与储值形成大量预收资金，储值余额与月营业额之比越高，兑付压力越大；团购与外卖渠道占比过高，会削弱定价权并受平台账期约束；主料价格波动大，供应商集中会放大涨价与断供冲击；租金与人工等刚性成本在客流下滑时放大盈利压力；主料溯源完整度用于核验食材来源与合规底线；堂食收入以现金和多渠道支付为主，须以现金流质量核验账面收入。上述因子风险越高，得分越低。

现制茶饮门店 十一个指标 十七项因子

现制茶饮门店以奶茶、果茶等现制饮品为主，依靠高频消费、快速出杯与加料连带提升单店产出，外卖平台和会员小程序是主要渠道，需关注券码核销、储值负债、原料损耗与出杯量真实性。

维度	指标	因子
收益维度	单店杯量与产出规模	日均付费出杯量 同店净收入增长率
收益维度	客单价与产品价值	实收客单价 单杯毛利率
收益维度	会员复购与获客效率	90 日会员复购率 会员消费额占比
收益维度	渠道与合同收入质量	外卖平台综合实收率
收益维度	门店单位经济性	新店现金投资回收期
风险维度	收入虚增与出品真实性	超产能时段杯量占比 杯量-原料勾稽偏差率
风险维度	券码与退款风险	退款率 秒退/极速退款率
风险维度	储值风险与偿付	储值负债覆盖率
风险维度	原料与损耗风险	原料成本率 全店综合损耗率
风险维度	结算与资金真实性	渠道到账流水勾稽差异率
风险维度	合规与经营底线	证照完备率

表 16-3：现制茶饮门店的指标与因子

收益维度：杯量、客单与单店回报

现制茶饮单价较低，主要依靠高出杯量形成单店营收基础。同店净收入增长率能够反映现制茶饮门店的未来增长潜力；加料可以提高客单价，但原料成本同步上升，须将实收客单价与单杯毛利率并看；茶饮消费频次高，90 日复购与会员消费占比反映客群黏性；外卖平台抽佣与补贴使名义收入与实收差距较大，需看综合实收率；新店投资回收期决定单店模型能否支撑扩张。

风险维度：出杯真实性与经营底线

现制茶饮的风险重心在收入真实性与预收资金。出杯量可能被虚报，超产能时段的杯量与原料耗用勾稽是识别刷单的实物锚点；退款与秒退是刷单、套券的典型痕迹；会员储值形成负债，须看覆盖能力；原料成本与全店综合损耗决定毛利能否守住；多平台结算使到账流水与账面收入容易脱节，须逐渠道勾稽；证照是经营底线。上述因子风险越高，得分越低。

三个案例同属餐饮零售，指标与因子的构成各不相同：便利店侧重夜间经营与加盟体系，火锅侧重桌台产能与预收兑付，茶饮侧重出杯真实性与券码退款。

16.2 十七个案例的选择与阅读方法

完成三个品类精读后，可继续查看附录一收录的 17 个正式品类案例。附录案例覆盖餐饮零售、住宿与到家服务、出行与空间运营、内容制作、算力与技术服务等经营形态。

经营场景	案例编号
餐饮与零售	01 汉堡快餐门店；02 面包烘焙门店；04 量贩零食零售；05 连锁药店门店；08 茶馆运营
出行、住宿与生活服务	03 汽车租赁；06 机构型养老院；07 经济型商旅酒店；09 共享按摩椅运营；10 奢侈品综合养护；11 宠物上门照护服务
即时零售与内容服务	12 小型沉浸式演艺场馆；13 即时零售前置仓；14 电子竞技赛事承办；15 短剧制作
算力与技术服务	16 智能计算中心托管；17 人工智能模型部署与运维

表 16-4：附录一 17 个案例的经营场景

附录一所列 17 个案例为展示样本，各案例的指标与因子内容和数量均按品类确定。例如，汉堡快餐门店以月坪效、日均单量衡量选址与客流承载，汽车租赁以出租率、残值实现率衡量资产产出与退出，机构型养老院以收费入住率、护理员照护比衡量变现与承载，智能计算中心托管以图形处理器（GPU）计费利用率、应收账款周转天数衡量资源变现与回款。

阅读时，请先按照业务类型确认所属品类范围，再分别查看收益维度与风险维度的指标，以及各指标下支持判断的因子。附录一仅提供指标与因子结构；量化口径、数据来源、单位、观察期间、权重与 1 至 3 分标准，须在各品类正式规则版本中明确后方可评分。

第六部分 未来展望

第十七章 从本轮成果走向下一次交付

牛码新市提供合同证券化市场的登记与度量基础设施，帽子军团是制造和维护六件仪器的 AI 工厂。本书从业务难题出发，说明工具的用途、工厂的理论、组织的理论与作业方式、案例的读法，以及持续接受检查的治理关系。8·03 提出合同现金流投融资的市场框架，9·03 提出双罗盘的度量规则，10·03 进一步公开仪器由谁制造、怎样检验和维护。接下来要做的，是把已形成的规则放进具体事实与连续运行中验证。

17.1 本轮的理论、实践与成果

理论上，我们提出帽子军团。HAT 把规则、授权与最终责任留给人，MASS 把完整模型能力放进边界明确的微任务；协同、独立检验与“十二个足够”说明规模化的条件。由此得到的判断是：能否扩大 AI 的使用，不能只看能力与数量，还要看任务、检验和责任能否一起扩展。

实践上，5000 名 AI 员工已经上岗。岗位已纳入组织台账，但不代表同时执行任务。稳定岗位、动态组队、四类职能和七站作业把理论要求落实为组织安排。连锁便利店的说明性模拟案例进一步展示了澄清、修订和验收怎样分别留下依据；验收了修订记录，并不等于候选取得可选资格，更不等于企业验证完成。真实运行仍须以各项任务的运行与验收记录为依据，并按相应披露与授权安排供核查。

成果上，行业罗盘规则已经覆盖 500 个行业，本书通过正文 3 例与附录 17 例，合计展示 20 个品类的指标与因子结构，让行业判断落到可以逐项追问的经营变量上，为牛码新市冷启动建立第一轮讨论与应用基础。企业取数、评分、回测和跨行业校准仍须以具体版本和证据说明为准。

17.2 企业取数、校准、模拟与进展披露

下一阶段的工作围绕四个方面展开，按任务依赖相互衔接，进展披露贯穿其中。

一是企业取数与实际评分。先确认经营对象、适用品类和规则版本，核对量化口径、来源、期间、权重与分档，再取得和核验证据。资料没有到位的字段继续留白，不能为了交付完整分数而补写事实。

二是回测与校准。将明确版本的因子、阈值和读数与真实经营及履约结果对应，区分输入变化、执行偏差和规则问题。每轮修订保留依据、新旧版本和影响范围，牛码调整系数及跨行业比较也按实际验证进度说明。

三是模拟环境中的连续运行。用明确标注的模拟合同和模拟资金连接登记、度量、任务交接、检验、预警和处置，并加入不利情景。模拟要说明假设、限制和未完成环节，结果不作为真实交易业绩。

四是持续披露进展。牛码新市将于每月3日向市场公布进展，按研究和验证的实际进度说明完成了什么、证据支持到哪里、仍有什么缺口，并保留版本与修订记录。规则覆盖、企业验证、模拟运行与正式服务分别说明状态；具体开放范围以相应披露为准。

17.3 模拟环境中的连续运行

有了地图、罗盘和队伍，读者自然会追问：这些东西一起转起来，究竟是什么样？合同不断进入，经营情况每天变化，风险逐渐显现，这支AI军团能否及时发现问题，沿着规则完成预警和处理？这些问题，需要把系统放进连续运行的环境中检验。

我们计划用模拟合同、模拟资金，让已建成的系统功能与作业流程在模拟环境中运转起来，贯通合同进入、度量、模拟配置、持续跟踪、预警与处置。模拟合同、模拟资金、模拟配置及结果均须明确标识，与真实交易及持仓记录分开管理。先检查任务交接是否顺畅、规则能否执行、异常能否被发现并交由相应岗位处理，以及每项读数能否回到证据与责任主体，再按实际业务功能、许可和条件确定后续安排。

这项模拟对内是一场系统检验，对外则计划提供一间供参与者体验的模拟驾驶舱。随着建设与验证推进，参与者将能够输入条件、调整参数、尝试不同组合，观察系统在不同经营情景下怎样运转。行业与技术同行也可以在相应授权范围内沿任务记录检视：哪些规则有效，哪些数据难以取得，哪些判断仍需要澄清。具体实施范围与开放安排，将在相应条件具备后说明。

除了正常经营，模拟还将纳入集中违约、资金紧张、欺诈等不利情景，检查预警是否及时、处置是否依规，以及系统在压力下暴露了哪些不足。模拟身份将进行明确标注，确保其假设及修改依据可供查验。信任要在看得见、查得清、亲手试过的过程中建立。

17.4 行业反馈与共同检验

你的生意，收益来自哪里？风险藏在哪里？

欢迎从本书正文 3 例与附录 17 例出发，通过 500 行业罗盘成果索引找到与你的经营活动相关的罗盘。同时也欢迎您就指标与因子提出具体反馈：哪些变量抓住了经营关键，哪些定义仍需澄清，哪些数据难以取得。500 个行业罗盘是这一轮共同讨论的起点；反馈回到具体行业、具体字段与具体版本，才能推动规则继续完善。

9·03 留下的邀请，也将在下一阶段的模拟探索中逐步获得更具体的表达：

您的订单，您的组合，您的 SHARP。

Your Order. Your Portfolio. Your SHARP.

地图逐步展开，罗盘继续校准，军团已上岗。接下来，登记与度量仪器要在同一套系统中协作，把每条读数轨迹、每次异常处理和每轮修订变成可以查验的记录。规则覆盖、企业验证、模拟运行与正式服务分别说明状态；每向前一步，都要让证据和责任跟得上。

牛码新市邀请企业经营者、行业专家、金融机构与 AI 技术伙伴参与共建，并欢迎监管部门提出指导意见。通往下一次交付的路上，牛码新市期待与您同行。

行业罗盘网站已上线：<https://microconnect.com/industry-compass>。可点击进入，或扫描执行摘要后的二维码。

附录一 17 个补充案例与 500 行业成果导航

第十六章 16.1 节集中精读连锁便利店、川渝火锅门店与现制茶饮门店；本附录收录其余 17 个品类案例，独立编号为 01—17，以下索引均对应本附录。

案例范围与阅读方法

行业罗盘成果已经覆盖 500 个行业。本书选取其中 20 个品类：3 例在第十六章 16.1 节精读，其余 17 例收录于本附录，均展示品类简介，以及收益维度和风险维度下的指标与因子。案例编号用于本书内检索，不等于 500 行业目录中的行业编号。

各例均采用“维度—指标—因子”三层结构。每个维度设置 3 至 9 个指标，并包含 3 至 9 项因子；每个指标至少对应一个因子；具体选择随品类特点而变化。收益支撑越强，因子得分越高；风险越高，因子得分越低。两侧分别观察，不以收益支撑抵消风险暴露。

使用边界

本附录展示 17 个品类的指标与因子结构，未列示量化口径、数据来源、单位、观察期间、权重与 1 至 3 分的具体分档。实际应用前，须在相应品类的正式规则版本中明确上述内容，核对业务范围、适用版本与证据，再完成企业取数、核验、评分和回测。

本附录 17 例不是代表性统计抽样，不能代替 500 个行业完整成果，也不能推定其他行业采用相同指标或因子。跨行业比较仍须结合归一方法、行业侧校准及相关验证依据，不能从本附录的结构表直接得出可比结论。

500 行业成果导航

统一成果入口显示行业编号、成果版本、开放状态和可在线浏览资料，具体行业页展示已开放的罗盘结构、规则资料与版本说明。本书案例序号仅用于书内检索，不作为行业编号。

500 行业状态栏：是否调整、调整版本、验证情况和可比范围须逐行业披露。完整索引与开放范围以网站披露为准；本版未披露可核验的逐行业调整版本，不能据此确认跨行业可比性。

案例索引

案例	品类
01	汉堡快餐门店
02	面包烘焙门店
03	汽车租赁
04	量贩零食零售
05	连锁药店门店
06	机构型养老院
07	经济型商旅酒店
08	茶馆运营
09	共享按摩椅运营
10	奢侈品综合养护
11	宠物上门照护服务
12	小型沉浸式演艺场馆
13	即时零售前置仓
14	电子竞技赛事承办
15	短剧制作
16	智能计算中心托管
17	人工智能模型部署与运维

表 附 1-1：17 个品类案例索引

01 汉堡快餐门店

以汉堡、炸鸡、薯条及饮品为核心的标准化快餐业态，依靠高客流选址、套餐组合与快速出品获取收益，外卖平台与会员复购共同支撑单店营收，经营质量取决于出品效率、食材成本与渠道结构。

收益维度 7 个指标 9 项因子

指标	因子
一、选址与客流承载	月坪效 日均单量
二、菜单结构与客单价	客单价
三、出品效率与产能	平均出品时长
四、渠道结构与获客成本	外卖订单贡献率
五、供应链与食材成本	食材成本率
六、复购与会员资产	月复购率
七、单店现金贡献与回报	单店经营利润率 单店回本周期

表 附 1-2：汉堡快餐门店——收益维度指标与因子

风险维度 7 个指标 9 项因子

指标	因子
一、订单刷量与虚假交易风险	重复账号/设备下单占比
二、出品-原料耗用冲突	全店综合损耗率异常
三、储值团购提前确认	储值未核销占比
四、资金控制	非对公收款占比
五、现金与退款风险	收银差异率
六、库存与临期风险	临期原料占比 原料周转天数
七、渠道与合同依赖	最大平台净收入占比 最大供应商采购占比

表 附 1-3：汉堡快餐门店——风险维度指标与因子

02 面包烘焙门店

以现烤面包、蛋糕及西点为主的连锁门店业态，前店后厂或中央工厂配送，收益来自现制产品溢价、会员储值与私域复购，需重点关注当日报损、原料耗用与储值团购带来的收入确认问题。

收益维度 5 个指标 9 项因子

指标	因子
一、单店产出与坪效	可比同店实收销售额增长率 坪效 客单价
二、产品结构 with 贡献	SKU 动销率
三、复购与会员粘性	90 日付费复购率 会员销售贡献占比
四、渠道与可复制性	私域/自有渠道营收占比 满 12 个月门店盈利门店占比
五、效率与履约	人效

表 附 1-4：面包烘焙门店——收益维度指标与因子

风险维度 7 个指标 9 项因子

指标	因子
一、刷量检测	同收货地址关联订单占比
二、守恒勾稽	月末原料盘点差异率
三、预付违规	储值卡充值当日即全额确认收入比例 储值未核销余额占比
四、平台清分	逾期结算及经营应收占比
五、产销平衡	全店综合损耗率偏离度
六、现金链	现金-日结-缴存勾稽差异率 秒退订单占比
七、现金流真实性	异常支付及期后冲回率

表 附 1-5：面包烘焙门店——风险维度指标与因子

03 汽车租赁

以车队资产为核心的短租与长租服务业态，收益来自车辆出租率、日均租金与租期管理，单车现金回报和残值处置决定资本效率，押金预收资金安全、车辆状态真实性及事故保险勾稽是主要风险来源。

收益维度 4 个指标 8 项因子

指标	因子
一、车队出租产出与利用效率	出租率 单车日均租金
二、单车现金贡献与资本效率	单车月度毛利 车辆投资回收周期
三、合同现金流与回款真实性	租金收现率 订单-合同-收款三流一致率 复租/续租订单占比
四、资产退出与偿付安全	残值实现率

表 附 1-6：汽车租赁——收益维度指标与因子

风险维度 5 个指标 9 项因子

指标	因子
一、车辆资产真实性与排他性	车辆状态与系统台账一致率
二、押金与预收资金安全	押金/预收资金隔离 押金及预收转收入占比
三、收入与现金流真实性	租金收入与银行流水匹配率 个人账户收款占比 关联方交易租金占比
四、事故停运与保险勾稽	保险报案与系统记录匹配率
五、残值与资产退出风险	残值高估/退出净回收偏差 骗租失联率

表 附 1-7：汽车租赁——风险维度指标与因子

04 量贩零食零售

以大包装、低价格、高周转为特点的零食连锁业态，总部集中采购并向加盟门店供货，收益依赖商品毛利、库存周转与网络扩张速度，加盟商压货、返利依赖、应收账款及收入真实性需重点核验。

收益维度 6 个指标 9 项因子

指标	因子
一、单店零售产出	单店日均销售额 月均坪效
二、商品毛利与供应链结构	门店综合毛利率
三、库存周转与商品效率	全链库存周转天数 SKU 动销率
四、仓配履约效率	单店配送成本率
五、现金循环与资金效率	现金转化周期
六、网络扩张与加盟商质量	净开店率 新店首年销售达标率

表 附 1-8：量贩零食零售——收益维度指标与因子

风险维度 6 个指标 8 项因子

指标	因子
一、门店扩张与集中依赖	年度闭店率
二、货权与压货	加盟店库存深度 加盟商退货率
三、库存与临期损失	库存损耗率 账实库存差异率
四、应收与账期风险	加盟商平均授信账期
五、预收与资金依赖	加盟商预收及保证金占总负债比
六、采购集中与返利依赖	TOP5 供应商采购占比

表 附 1-9：量贩零食零售——风险维度指标与因子

05 连锁药店门店

以处方药、非处方药及健康产品销售为主的连锁零售业态，收益来自门店销售、会员复购与慢病管理，医保结算和线上 O2O 渠道占比较高，需关注医保合规、药品效期库存、执业药师配备及现金流质量。

收益维度 5 个指标 9 项因子

指标	因子
一、门店销售产出能力	单店日均净销售额 客单价
二、毛利与品类结构	综合毛利率 高毛利品类销售占比
三、会员复购与慢病价值	会员销售占比 会员 90 日复购率
四、门店网络盈利复制	同店销售增长率 盈利门店占比
五、合同收入与现金实现	净销售现金实现率

表 附 1-10：连锁药店门店——收益维度指标与因子

风险维度 6 个指标 8 项因子

指标	因子
一、库存与效期风险	库存周转天数
二、合规与专业运营	医保违规处罚次数 执业药师在岗率
三、现金流与流动性	经营性现金流净额/净利润 医保应收周转天数
四、收入集中度与依赖	医保销售占比
五、扩张与关店	亏损门店占比
六、数据与核验风险	盘点差异率

表 附 1-11：连锁药店门店——风险维度指标与因子

06 机构型养老院

为失能、半失能及自理老人提供长期住养、生活照护与护理服务的机构业态，收益取决于收费入住率、护理等级定价与增值服务，预收资金安全、护理人员配置、服务记录真实性及资质合规是核心风险。

收益维度 4 个指标 6 项因子

指标	因子
一、床位入住与变现效率	收费入住率 单床月综合服务收入
二、护理等级与定价结构	护理费收入占比 中重度失能在住占比
三、合同收费与现金流韧性	可管理续住率
四、获客转化与增值服务	增值服务收入占比

表 附 1-12：机构型养老院——收益维度指标与因子

风险维度 5 个指标 7 项因子

指标	因子
一、护理承载与排班真实性	护理员照护比
二、服务记录完整性与补录	照护记录异常补录率
三、预收资金与退款安全	预收资金专户监管覆盖率
四、医疗边界与合规许可	经营许可完备性 消防安全验收状态
五、暂离真实性与异常事件	关联方资金占用比 异常事件

表 附 1-13：机构型养老院——风险维度指标与因子

07 经济型商旅酒店

面向商务出行与大众住宿的中端及经济型酒店业态，收益取决于入住率、平均房价与直销会员渠道，OTA 佣金、租金和人房比决定盈利空间，需核验房晚真实性、储值收入确认及协议客户挂账回款情况。

收益维度 3 个指标 8 项因子

指标	因子
一、客房收益质量	入住率 净均价
二、渠道与会员	直销/协议渠道占比 会员房晚占比 储值递延确认比例
三、成本与盈利	能耗成本率 人房比 租售比

表 附 1-14：经济型商旅酒店——收益维度指标与因子

风险维度 4 个指标 7 项因子

指标	因子
一、房晚与入住真实性	超可售库存房晚率
二、收入确认与粉饰	储值当期一次性确认收入占比 退款后收入未回冲率
三、渠道结算与佣金	OTA 佣金漏计率 渠道对账未解释差异率
四、协议挂账与回款	协议挂账未回款占比 协议挂账逾期 90 天以上占比

表 附 1-15：经济型商旅酒店——风险维度指标与因子

08 茶馆运营

以茶饮消费、茶叶零售与社交包间为核心的空间经营业态，收益来自桌时利用率、客单价与会员储值，预订和包场锁定部分收入，需关注储值负债、高值茶叶库存、个人账户收款等内控问题。

收益维度 5 个指标 9 项因子

指标	因子
一、获客与进店转化	团购券核销率 平台订单毛利率 直接渠道收入占比
二、桌台与空间利用	全时段付费桌时利用率
三、客单价值与产品结构	客单价 综合毛利率
四、会员储值与渠道质量	储值客户续充率 储值消费转化率
五、经营效率与收入兑现	同店净收入增长率

表 附 1-16：茶馆运营——收益维度指标与因子

风险维度 5 个指标 7 项因子

指标	因子
一、收入确认与储值风险	未履约款误计收入率 储值负债账实差异率
二、履约真实性与核销	原料耗用吻合度
三、支付现金与核销风险	个人账户收款占比 现金缴存差异率
四、出品原料与库存风险	高值茶叶盘点差异率
五、内控权限与合规	数据缺失日占比

表 附 1-17：茶馆运营——风险维度指标与因子

09 共享按摩椅运营

在商场、影院、交通枢纽等场所投放按摩椅并按次扫码收费的无人值守业态，收益取决于点位流量、单椅使用频次与分成条款，需关注刷单与补录订单、点位合同稳定性、设备运维及结算资金归集。

收益维度 5 个指标 9 项因子

指标	因子
一、点位单产	单椅日付费订单数 实收客单价
二、流量与履约	同点位实收净额增长率 设备在线可售率
三、单位经济	单椅现金回收期 正贡献点位占比
四、合同收益权	到期点位续约率 运营方实收留存率
五、收入现金化	订单至可支配到账周期

表 附 1-18：共享按摩椅运营——收益维度指标与因子

风险维度 6 个指标 8 项因子

指标	因子
一、收入真实性	离线补录订单占比
二、合同成本完整性	场租及分成计提差异率 无有效合同运营点位占比
三、结算与资金	逾期结算款占比 资金归集完整率
四、点位集中与撤离	前五大点位收入集中度
五、退款与争议	退款金额率
六、设备运维与安全	30 日同故障复发率

表 附 1-19：共享按摩椅运营——风险维度指标与因子

10 奢侈品综合养护

为箱包、皮具、腕表、鞋履等高值奢侈品提供清洁、修复、翻新与保养服务的专业业态，收益来自高客单工艺服务与复购，需重点关注高值品保管与责任追溯、工序真实性、外包管控及赔付风险。

收益维度 6 个指标 7 项因子

指标	因子
一、获客与成交转化	有效咨询付费转化率
二、客单价与定价结构	非基础护理收入占比
三、复购与客户价值	12 个月复购率 客户收入留存率
四、技师产能与履约效率	技师人均月产值
五、单件贡献与成本控制	单件贡献率
六、合同现金流	经营现金转化率

表 附 1-20：奢侈品综合养护——收益维度指标与因子

风险维度 7 个指标 8 项因子

指标	因子
一、高值品保管与责任追溯	高值品盘点差异率
二、授权与工序真实性	开工前书面授权覆盖率 收费工序记录匹配率
三、外包透明度与管控	外包方资质核验率
四、履约质量与返修	返修率
五、赔付暴露与资金安全	赔付损失率
六、资金真实性与反欺诈	非经营主体账户收款率
七、人员与经营连续性	核心技师收入集中度

表 附 1-21：奢侈品综合养护——风险维度指标与因子

11 宠物上门照护服务

为宠物主提供上门喂养、遛狗、洗护及短期寄养的到家服务业态，收益取决于区域订单密度、复购留存与增值服务，需关注服务履约真实性、宠物安全事故、人员稳定以及私单飞单等情况。

收益维度 4 个指标 5 项因子

指标	因子
一、需求密度与订单转化	服务区域订单密度
二、定价与收入结构	增值服务收入占比
三、履约变现与完单质量	服务记录客户确认率
四、复购与客户留存	月复购率 获客后 180 日留存率

表 附 1-22：宠物上门照护服务——收益维度指标与因子

风险维度 7 个指标 7 项因子

指标	因子
一、履约真实性与证据完整性	履约异常单占比
二、宠物安全与异常处置	重大安全事故率
三、承载超限与排班冲突	同人超量接单率
四、人员稳定性与私单	私单/飞单率
五、预收资金与现金流安全	现金覆盖运营月数
六、退款赔付与客诉	赔付率
七、档案授权与合规保险	责任险覆盖订单占比

表 附 1-23：宠物上门照护服务——风险维度指标与因子

12 小型沉浸式演艺场馆

以小剧场、沉浸式戏剧及互动演出为核心的中小型演艺空间，收益来自票房、驻演剧目生命周期与包场团建等非票收入，需关注场次与票务收入勾稽、单场承载安全及演出报批合规。

收益维度 5 个指标 9 项因子

指标	因子
一、内容与剧目竞争力	观众复购率 驻演剧目平均生命周期
二、场次与上座效率	上座率 单座单场票务收入
三、单场盈利与成本结构	单场边际贡献率 场租占收入比
四、多元变现能力	非票收入占比 包场/团建场次占比
五、合同现金流与回款质量	票务结算回款周期

表 附 1-24：小型沉浸式演艺场馆——收益维度指标与因子

风险维度 4 个指标 6 项因子

指标	因子
一、收入真实性	场次-票务-银行三方勾稽差异率 赠票/内部票占比
二、沉浸式承载与安全	单场承载超限比 消防验收与安全许可有效性
三、现金流与合同履约	退票挤兑缓冲
四、资质与合规	演出报批/备案齐全率

表 附 1-25：小型沉浸式演艺场馆——风险维度指标与因子

13 即时零售前置仓

依托社区周边前置仓提供 30 分钟至 1 小时即时配送的零售业态，收益取决于单仓订单密度、客单价与复购频次，履约成本刚性较高，需关注损耗缺货、平台依赖、库存账实差异及单仓现金流压力。

收益维度 6 个指标 9 项因子

指标	因子
一、订单规模与频次	单仓日均订单量 月复购率
二、客单价与毛利结构	客单价 综合毛利率
三、履约交付能力	准时交付率 单均仓内履约成本
四、合同现金流与资金周转	现金转换周期
五、结算真实性与收入质量	结算对账准确率
六、单仓贡献与网络复制	成熟仓正贡献占比

表 附 1-26：即时零售前置仓——收益维度指标与因子

风险维度 6 个指标 9 项因子

指标	因子
一、履约成本刚性	履约费用率
二、损耗与缺货	全店综合损耗率 缺货率
三、退款售后	退款率
四、平台依赖与现金流错配	最大单一平台 GMV 占比 应付账款逾期率
五、内控舞弊与数据一致性	库存账实差异率
六、流动性与单仓盈亏	现金缓冲月数 单仓 EBITDA 率

表 附 1-27：即时零售前置仓——风险维度指标与因子

14 电子竞技赛事承办

承接电竞赛事策划、制作、执行与商业化运营的服务业态，收益来自主办方合同、赞助转播与票务收入，项目回款周期较长，需关注观赛与流量数据真实性、收入确认时点、补贴依赖及关联方资金回流。

收益维度 6 个指标 8 项因子

指标	因子
一、合同现金流与回款质量	到期应收回款率
二、单赛事盈利与成本效率	单赛事毛利率
三、票务变现与观赛质量	付费上座率 单付费观众票务收入
四、赞助与转播商业化	赞助商续约率 赞助权益验收通过率
五、客户复购与赛事持续性	赛事客户复购率
六、收入确认与交付质量	预收款转收入兑现率

表 附 1-28：电子竞技赛事承办——收益维度指标与因子

风险维度 6 个指标 8 项因子

指标	因子
一、观赛与在线流量真实性	线下观赛人次核验偏差率
二、收入确认与现金流错配	未完成履约收入占比 单赛事现金回款与确认收入偏差率
三、成本与负债完整性与合规	奖金池与战队出场费应付款
四、补贴与外部资金依赖	政府补贴占总收入比
五、关联交易与资金回流	关联方回款占比 30 日内回流资金占比
六、证据链与数据口径	口径冲突事项

表 附 1-29：电子竞技赛事承办——风险维度指标与因子

15 短剧制作

从事网络剧、网络电影及竖屏短剧开发、拍摄与发行的内容制作业态，收益依赖平台分账、版权销售与投流变现效率，需关注投流成本与回收、制作预算超支、收入跨期确认、存货减值及版权合规。

收益维度 5 个指标 8 项因子

指标	因子
一、分账回收与兑现	平台结算收入兑现率 分账回款周期
二、投流变现效率	投流 ROI 投流回本天数
三、项目盈利能力	单项目毛利率 项目投资回收倍数
四、制作交付效率	按期交付率
五、现金回收能力	平台结算回款及时率

表 附 1-30：短剧制作——收益维度指标与因子

风险维度 6 个指标 7 项因子

指标	因子
一、投流风险	投流成本占比
二、制作成本与履约风险	单项目预算超支率
三、收入确认与跨期风险	未上线已确认收入占比
四、减值与存货风险	完片但无平台排播项目占比
五、关联交易风险	关联方交易收入占比 分账回款资金回流疑点
六、合规与版权风险	授权链条完整率

表 附 1-31：短剧制作——风险维度指标与因子

16 智能计算中心托管

为客户提供 GPU 算力、机柜、电力与运维托管的数据中心服务业态，收益来自长期合同、算力计费利用率与增值服务，资本投入大、回收期长，需关注电力成本、设备代际迭代、债务偿付及客户集中度。

收益维度 6 个指标 8 项因子

指标	因子
一、合同现金流质量与锁定	含最低承诺/照付不议条款合同金额占比 静态投资回收期
二、算力资源变现与利用率	GPU 计费利用率 机柜有效出租率
三、定价与单位经济性	单位贡献率
四、回款、现金实现与资本回收	应收账款周转天数
五、客户结构与续约	到期合同续约率
六、交付与服务兑现	交付达标率

表 附 1-32：智能计算中心托管——收益维度指标与因子

风险维度 6 个指标 8 项因子

指标	因子
一、电力承载与能耗成本	电力成本占收入比
二、设备可靠性与技术迭代	滞后代际卡算力占比 GPU 故障率
三、资金链与偿债风险	融资租赁设备占比
四、客户信用与回款风险	客户收入占比
五、合规与资质底线	关键许可完备性
六、收入真实性与反操纵	履约率 账单-回款勾稽匹配率

表 附 1-33：智能计算中心托管——风险维度指标与因子

17 人工智能模型部署与运维

为企业客户提供 AI 应用开发、垂类大模型微调、部署上线与持续运维的技术服务业态，收益来自项目交付与按调用计费，需关注调用计费真实性、算力成本承载、客户集中度、交付验收及关联交易风险。

收益维度 4 个指标 9 项因子

指标	因子
一、获客规模与转化效率	有效付费客户数 试用转付费率
二、调用变现与单位经济	调用成功率 单千次调用收入
三、收入留存与盈利质量	净收入留存率 年度续费率 毛利率
四、现金回收与合同质量	合同约定期限内回款率 应收账款周转天数

表 附 1-34：人工智能模型部署与运维——收益维度指标与因子

风险维度 4 个指标 7 项因子

指标	因子
一、履约真实性与数据合规	调用日志与账单可勾稽率 重复归属/重复订单率
二、成本承载与供应商风险	算力/资源承载率 第三方模型供应商依赖度
三、客户集中与交付质量	前五大客户收入占比 验收后退款/扣款率
四、关联交易与异常回款	关联方收入占比

表 附 1-35：人工智能模型部署与运维——风险维度指标与因子

附录二 十二个足够总表

“十二个足够”是帽子军团得以成立的一组共同条件。它们分别约束任务怎样拆分、模型底座怎样配置、组织编制怎样形成制衡，以及产出怎样接受检验。四组条件彼此连接：任务不够清楚，后续编制和检验就失去对象；底座缺少多元与对冲，多个结果可能只是重复同一种偏差；编制只有数量而没有细分和制衡，规模越大，错误越可能沿协作链扩散；检验不够密、不够长或不够深，短期正确也难以转化为持续可信。

因此，十二项条件不能被当作可以任选的优点，也不能用某一项的突出表现抵消另一项的缺失。“缺一即警”不是把已有记录和成果全部归零，而是识别当前依赖的条件是否成立；一旦条件无法回答，就应停止依赖尚未成立的部分，查明缺口，完成校准后再验证。

一、十二个足够总表

类别	条件	核心要求	条件不足时需要核查	正文回查
任务	足够小	一次只做一件事，并明确输入、输出和合格标准。任务虽小，所调用的模型能力并不因此降低。	是否把多个判断混在同一任务中；输入输出是否清楚；结果能否按预设标准验收。	第八章 8.1 节、第八章 8.3 节
任务	足够窄	任务权限和判断边界保持有限，不依赖执行过程中的临时自由裁量；需要裁量时升级给人。	允许与禁止的范围是否明确；任务是否要求在缺少规则时自行作出重大判断；升级条件是否存在。	第八章 8.2 节、第八章 8.3 节
任务	足够碎	每项产出可以独立检验，每一步骤可以重放，使错误能够被隔离和定位。	任务是否仍然只能整体判断成败；中间步骤、输入和产出是否可还原；错误能否定位到具体环节。	第八章 8.3 节、第八章 8.4 节
底座	足够多元	使用不同底座和供应商，降低对单一能力来源和同源偏差的依赖。	多个结果是否实际来自相同或高度相关的底座；能力来源是否过度集中；所谓多元是否只是名称不同。	第八章 8.5 节
底座	足够对冲	对同一任务进行并行判断，让分歧暴露出来，并按规则进入复核或升级。	并行结果是否具有独立判断价值；分歧是否被看见和处理；一致结果是否可能来自共同偏差。	第八章 8.5 节
底座	足够有性价比	根据任务难度配置不同能力和成本的模型，使模型能力、任务要求与资源投入相匹配。	是否所有任务都无差别调用同一能力层级；较低成本配置能否满足合格标准；节省成本是否损害检验要求。	第八章 8.5 节
编制	足够多	数量能够覆盖全部任务，并为并行执行、交叉复核和必要替补提供空间。	是否存在无人承接或无人复核的任务；数量增长是否超过规则、检验与责任的承载能力。	第九章 9.6 节
编制	足够细	分工细化到可以识别具体职责、交接关系和错误环节，避免责任被笼统归入一个整体。	任务失败时能否确定问题发生在哪一步；生产、验收和检验职责是否混在一起；交接对象是否清楚。	第九章 9.4 节、第九章 9.6 节

类别	条件	核心要求	条件不足时需要核查	正文回查
编制	足够冲	让不同立场在组织内部形成制衡，使质疑、复核和异议能够阻断系统性偏差。“冲”指判断立场之间具有必要张力，不是制造无序对立。	负责执行的人是否同时决定验收结论；不同角色是否只是重复同一种判断；异议是否能够被记录和处理。	第九章 9.4 节、第九章 9.6 节
检验	足够密	每项产出都进入验收或检验流程，不设默认免检区。	是否存在因任务量大、结果看似熟悉或历史表现良好而被跳过的产出；检验记录能否与任务对应。	第十章 10.1 节、第十章 10.4 节
检验	足够长	回溯覆盖足够长的历史、不同环境和边界情形，用于区分真实能力、短期运气和环境变化。	观察期是否只包含单一阶段；结论是否依赖短期表现；环境变化后原有答案是否已经过期。	第十章 10.2 节、第十章 10.4 节
检验	足够深	溯源能够从结果回到输入、模型调用、规则版本、执行路径、检验结果和责任记录。	出现错误时能否找到具体输入、规则缺口或流程环节；修订后能否再次验证同类错误是否被阻断。	第十章 10.3 节、第十章 10.4 节

表 附 2-1：十二个足够总表

二、总表的使用边界

这张表适合用于快速检查和正文导航，不代替各章对概念、组织关系和失效边界的完整说明。任务与底座两组条件的详细逻辑见第八章，编制条件见第九章，检验条件见第十章。

使用总表时，应把四组条件放在同一条证据链上理解。例如，一项任务即使足够小，如果边界仍依赖临时自由裁量，就不能据此认为它已经足够窄；多个模型即使同时给出结果，如果其底座高度同源，也不能仅凭数量认定形成了有效对冲；编制数量增加，如果没有同步增加规则维护、证据检查和异常处理能力，也不等于组织承载能力同步扩大。同样，单次验收通过不能替代持续回测，较长的观察期也不能替代对输入、规则 and 责任的深入溯源。

“十二个足够”提供的是判断条件，不是固定的数量门槛、计算公式或认证标签。不同任务的复杂程度、风险水平和证据条件不同，具体怎样达到“足够”，应以相应规则、验收标准、运行记录和回测结果为依据。总表本身不能证明某项任务、某个组织或某一成果已经满足全部条件，也不能把理论上的制衡关系写成未经检验的可靠性保证。

当某一条件无法成立时，应先确认影响范围：问题来自任务设计、底座配置、编制分工，还是检验深度；随后停止依赖受影响的判断，修订相关规则或安排，并用新的证据重新验证。已有记录仍然保留，因为它们正是识别偏差、比较版本和判断校准是否有效的基础。信任不依靠过去成绩自动累积，而要在每一次任务、每一次检验和每一次修订中重新形成。

附录三 名词释义与中英文对照

本表按英文名称 A-Z 排列，说明各名词在本书中的含义。登记与度量仪器的功能、组织运行、检验流程及罗盘应用的具体规则，仍以正文相应章节为准。

中文名词	英文全称	释义
验收	Acceptance Review	指依据预先确定的任务目标、适用规则和交付标准，检查产出及执行过程是否合格，并明确其可进入后续哪一环节的判断。
调度管理	Administration	指四大职能中的 A 类职能，负责将目标组织为可执行任务，管理任务拆解、依赖、选岗和协作衔接，不替代专业判断或独立验收。
骗子视角	Adversarial Manipulation Perspective	指主动寻找数据、指标或评分被包装和操纵路径的对抗性观察立场，不指认任何具体经营者具有欺诈身份。
智能体	Agent	指能够围绕给定目标，调用模型、知识、技能和工具执行工作的技术主体。在帽子军团中，其活动受当前任务、帽子和授权范围约束。
智能体能力表现	Agent Capability and Performance Profile	指用于岗位检索与调度的能力及实际表现记录，包括能力来源、版本、适用条件和任务表现。具备能力不等于获得相应任务权限。
身份编码	Agent Identifier (Agent ID)	指 AI 岗位的稳定组织标识，用于关联岗位职责、状态、认领关系和历史记录。同一岗位参与不同任务时，身份编码保持连续；运行编码则标识某次实例执行。
智能体实例	Agent Instance	指为执行具体任务而启动的一次技术执行实体，承接当次目标、上下文、帽子、权限、能力配置及交付要求。
兵对兵	Agent-to-Agent Collaboration	指智能体在既定规则内进行任务分派、执行、交接和复核的协作通道，不具有自行变更治理规则或扩大授权的权力。
AI 员工	AI Employee	指帽子军团对纳入岗位、责任和协作体系的 AI 使用的组织称呼。内部管理须进一步区分稳定岗位与具体运行实例。

中文名词	英文全称	释义
AI 岗位	AI Organizational Position	指组织内部相对稳定的身份及长期记录载体，承载职责、状态、具名真人认领关系和历史表现，不因单次任务或实例运行结束而消失。
岗位台账	AI Position Register	指持续记录 AI 岗位身份、职责、状态、认领关系和历史表现的管理记录，与单次实例运行信息分层保存。
业务执行	Business	指四大职能中的 B 类职能，负责形成与任务目标对应的分析、处理和交付，并保留支持结果的输入、依据和过程记录。
问业	Business Relevance Review	指检查候选因子是否对应有意义的经营问题，以及其影响机制和成立条件的检验视角，形成经营问题与机制图谱。
量尺	Calculation and Scoring Specification Review	指检查定义、参数、单位、观察期间及边界处理能否支持唯一计算和评分的检验视角，形成定义、参数与边界表。
候选资格状态	Candidate Eligibility Status	指候选因子经审查后所处的排除、待澄清、待修订或可选四类状态。可选表示具备参与组合筛选的资格，不等于最终入选。
候选因子	Candidate Factor	指已经提出、尚需接受设计审查和资格判断的因子。成为候选不代表已进入最终方案或已通过企业验证。
本文案例编号	Case Number in This White Paper	指附录一为 17 个展示案例设置的顺序编号，如案例 01。该编号仅用于附录一内检索，与行业目录中的行业编号分别使用。
首席执行官 (CEO)	Chief Executive Officer	在图 14-1 中指负责目标、优先级及资源边界统筹的组织角色。
同源偏差	Common-Source Bias	指多个判断因共享相同或高度相关的模型底座、资料或方法而产生的共同偏差。多个结果一致不能单独证明判断正确。
罗盘读数	Compass Reading	对已登记合同按日更新的因子数据，持续记录经营、履约与证据变化，并保留时点、期间及规则版本；缺失或延迟须标示，不构成投资行动建议。

中文名词	英文全称	释义
罗盘终端	Compass Terminal	供具备相应资格与授权的持牌机构/境外合格投资者读取登记资料、规则和读数的工作界面，由使用者按自身方法论形成判断，与公众行业研究入口分别管理。
合同罗盘	Contract Compass	对登记在册的合同提供度量的框架，由通用罗盘与行业罗盘组成，分别观察合同结构和具体生意；两者各有收益与风险维度，分别呈现、不相加、不互相抵消。
组合登记簿	Contract Portfolio Registry	投资者通过持牌机构完成交易和交收后，逐日记录合同与组合的关系及事实，并关联原始成交和持仓记录；不接触客户资产。
合同登记簿	Contract Registry	保存合同身份、条款、初始画像、运行状态、底层证据及历史版本的记录。登记不产生资产、不改变原合同权利义务，也不构成背书。
合同终端	Contract Terminal	与发行人或其 AI 代理持续对话的交互式系统，主动索取、核验及追问材料，把合同整理为符合准入规则的结构化记录；进场不等于投资推荐。
管控合规	Control	本书所称“管控合规”，是 AI 在授权内取证、核对、提示异常、开展设计审查和质量检查的职能。涉及合规批准、投资判断、适当性判断、对外生效的企业最终评分及法律意见，由具备相应授权与资质的人作出或确认并留痕。AI 仍可按规则计算和形成评分建议；内部计算、独立检查与最终批准分别承担不同责任。
通用码	Core Codes	指通用罗盘采用的指标及其评分表达，用于呈现合同结构层面的收益支撑和风险负载。具体项目读数由适用规则与资料形成，不等同于通用罗盘框架本身。
通用罗盘	Core Compass	指观察合同结构及跨行业共同要素的评分框架，分别设置收益维度和风险维度。
跨行业校准	Cross-Industry Calibration	指依据可解释的外部基准、数据及版本条件，对不同行业结果建立比较基础的过程，不意味着消除行业差异。

中文名词	英文全称	释义
模型对冲	Cross-Model Cross-Checking	指在可比较的任务条件下，由不同模型分别形成判断，再核对其假设、依据和分歧的检验安排，不因模型数量增加而自动成立。
寻源	Data Source and Access Review	指检查数据来源、取得条件、覆盖范围和限制是否可信且明确的检验视角，形成数据路径与条件表。
收口	Deliverable Consolidation	指汇合已通过验证的分项产出、处理未决事项并形成完整交付内容的过程，不替代独立检验或消除尚未解决的分歧。
部署支撑	Deployment	指四大职能中的 D 类职能，负责准备和维护模型、数据、技能、工具、系统及运行环境，保障权限配置和版本受控。
维度	Dimension	指罗盘的上层观察方向，包括收益维度和风险维度。维度下设置指标，指标由具体因子支持判断；通用罗盘与行业罗盘各自保留这两个维度。
分水	Discriminative Validity Review	指检查因子差异是否能够支持对不同经营状态作出有意义区分的检验视角，形成区分假设与证据表。
动态组队	Dynamic Team Formation	指依据具体任务的工作单元、能力需求、权限和可用状态，选择岗位并启动实例形成临时协作的过程。
息税折旧摊销前利润 (EBITDA)	Earnings Before Interest, Taxes, Depreciation and Amortization	本书用于即时零售前置仓案例；单仓 EBITDA 率的计算范围与分母须在正式规则中明确。
执衡	Eligibility and Composition Review	指接收六鉴完整意见，判断候选资格并从可选池中选择组合的职责位置。执衡不替代六项独立检验，也不代表企业验证已经完成。
资格台账	Eligibility Register	指记录候选因子版本、资格状态、判断依据、分歧及更新情况的记录，用于支持后续组合筛选和追溯。
全程溯源	End-to-End Traceability	指从结果回查输入资料、模型与工具调用、规则版本、执行路径、检验意见和责任主体的能力，其实际深度取决于证据是否完整留存。

中文名词	英文全称	释义
人工升级	Escalation to an Authorized Human	指任务遇到规则空白、权限例外或需要自由裁量的事项时，将有关事实、证据和执行记录提交具名真人处理的机制。
因子	Factor	指支持指标判断的可核验经营变量，须明确业务含义、数据来源、计算口径、单位、观察期间及评分规则。行业罗盘每个维度设置 3 至 9 项因子，每个指标至少对应一项，每项因子只归入一个指标。
因子全池	Full Factor Pool	指用于因子关系比较的既有因子集合，是检查候选因子重复性、依赖关系和新增信息的参照范围。
完整六审	Full Six-Perspective Review	指候选因子接受全部六鉴检验的完整审查。定义、用途或评分规则发生实质变化后，应形成新版本并重新接受完整六审。
全检	Full-Coverage Inspection	指每项产出均进入与其性质相匹配的验收或检验流程，不设置默认免检区。全检不等于每项均由人工检查，也不等于结果必然正确。
帽子	Governance Hat	指智能体在特定任务中适用的规则、职责、权限及升级安排的集合，用于规定允许事项、禁止事项和越界处理方式。帽子可以随任务变化，不等同于模型或行政职级。
帽对兵	Governance-to-Agent Control	指人与智能体之间以规则和授权为核心的治理通道，用于确定允许范围、禁止事项、监督要求及升级条件。
图形处理器 (GPU)	Graphics Processing Unit	本书用于算力服务场景；GPU 计费利用率及故障率的统计范围以相应正式规则为准。
国内生产总值 (GDP)	Gross Domestic Product	本书参考季度 GDP 核算中的行业增加值数据；行业增加值与 GDP 总量应分别理解。
商品交易总额 (GMV)	Gross Merchandise Value	反映规定期间和范围内的商品交易规模；退款、取消及统计范围须由正式规则明确，不直接等同于营业收入。

中文名词	英文全称	释义
帽子军团	HAT over MASS	指以人为规则、授权和最终责任的承担者，通过明确的任务边界组织大量智能体协作的 AI 组织方法，由 HAT 信任机制与 MASS 微任务协作方式共同构成。
HAT	Human-Commanded, AI-Operated, Trust	人管控、AI 运作的信任机制：人制定规则、决定授权并承担最终责任，AI 在边界内执行。人的责任与 AI 潜力须同时成立。
编号层	Identification Layer	指通过身份编码、运行编码及其关联记录，将稳定岗位、能力配置、任务和执行证据连接起来的层次。
照影	Incremental Information Review	指检查候选因子相对于既有因子全池的重复、包含、依赖、互补和新增信息的检验视角，形成全池关系与映射表。
指标	Indicator	指维度下的评价项目，规定该维度要回答什么问题，由具体因子支持判断和评分。行业罗盘在一个品类的每个维度内设置 3 至 9 个指标，内容按品类确定，不采用跨品类固定清单。
行业码	Industry Codes	指行业罗盘采用的指标、因子及其评分表达。具体项目读数须结合行业范围、观察期间、规则版本和适用条件解释，不等同于行业编号。
行业罗盘	Industry Compass	指牛码 SHARP 双罗盘中针对单一资产品类的专属评分规则，与全市场统一的通用罗盘并列；观察具体生意的收益来源、风险暴露及品类锚点，采用“维度—指标—因子”三层结构。
行业编号	Industry Identifier	指用于在行业目录中定位相应行业及其规则的标识，不代表评分，也不等同于行业码。
老炮视角	Industry Veteran Perspective	指利用长期行业经验识别纸面规则与实际经营之间的差异，以及行业周期、惯例和隐性约束的观察立场。

中文名词	英文全称	释义
破局	Manipulation Resistance Review	指检查真实经营状态未改善时能否通过改变口径、输入或记录方式提高分数的检验视角，形成攻击、防御与依赖表。
微任务智能体	Micro Agent	指在一次运行中承担简单、短时、边界有限任务的智能体。“微”描述其当次任务范围，不表示模型能力较低，也不是永久岗位类别。
MASS	Micro Agents; Simple & Short Tasks	微任务智能体与简单且短时的任务。完整模型能力进入有限任务范围，通过可检验的微任务协作形成规模。
微任务	Microtask	指从复杂目标中拆分出的、保留完整业务意义的工作单元，具有明确输入、单一交付目标、有限权限、可识别输出及独立验收条件，并可被单独定位、修正或重做。
多视角席位	Multi-Perspective Role	指为不同观察、判断和检验责任设置的任务位置，规定关注的问题及承担的责任，不等同于模型、工具或能力来源。
具名真人认领	Named Human Accountability Assignment	指将 AI 岗位与可识别的真人责任主体建立持续关联的组织安排。认领者在授权范围内承担规则、授权、监督和异常处理责任，不以逐项代替 AI 执行为必要条件。
牛码调整系数	NUMA Adjustment Factor	用于为行业侧判断补充经营环境参照的调整系数，不限定于单一数据来源。本版主要参照中国国家统计局相关数据，以 1 为中性基准。具体方法、适用范围及版本见第五章 5.3 节；结果须持续复核与回测。
牛码新市帽子军团	NUMA HAT over MASS	指牛码新市依照“帽子军团（HAT over MASS）”方法建立的 AI 组织实践，包括岗位设置、任务编组与检验安排。NUMA 指牛码新市；正文中的组织名称统一为“牛码新市帽子军团”，通用方法仍称“帽子军团”。

中文名词	英文全称	释义
牛码 SHARP	NUMA SHARP	指牛码新市用于表达和比较合同现金流风险收益结构的中心指标与统一量化语言。因子按公开权重形成指标及维度得分，归一后经牛码调整系数完成跨品类校准，形成行业 SHARP，再与通用罗盘共同形成单合同牛码 SHARP；具体计算须依适用版本的规则执行。
观察期间	Observation Period	指因子计算和判断所覆盖的时间范围，应与行业经营周期相适应，并随结果保留。
持续回测	Ongoing Backtesting	指对明确版本的判断持续进行事实检验，以识别稳定能力、短期运气和环境变化。使用历史资料的检验与后续新事实的验证，应分别保留所用版本、样本和期间。
线上线下融合 (O2O)	Online to Offline	指线上渠道与线下商品或服务履约相衔接的业务方式；具体到店、到家场景按业务范围界定。
在线旅行平台 (OTA)	Online Travel Agency	提供住宿等旅行产品预订或代理服务的平台；本书在酒店案例中用于说明渠道及佣金成本。
订单登记簿	Order Register	9-03 白皮书中记录已完成交易的成交条件、持仓、分派及退出状态的登记簿，与合同登记簿及组合关系保持连接。
正常经营者视角	Ordinary Operator Perspective	指从正常经营活动出发，分析因素如何影响收入、成本、效率和持续经营的观察立场。
销售点系统 (POS)	Point of Sale	本书指收银与销售记录系统，是核对交易记录及支付入账的数据来源之一。
登记与度量基础设施	Registration and Measurement Infrastructure	合同证券化市场中连接合同事实、证据、规则、读数及组合关系记录的功能体系；供持牌机构使用，并对自身数据、系统、服务与治理职责负责。
职责层	Responsibilities Layer	指组织三层设计中确定任务由谁承担、岗位如何分工及如何相互制衡的层次。

中文名词	英文全称	释义
收益维度	Return Dimension	指罗盘中观察收益形成与兑现支撑的维度。行业罗盘的收益因子，收益支撑越强，得分越高；该得分不等同于实际收益率或收益承诺。
投资回报率 (ROI)	Return on Investment	本书指短剧案例中的投流回报指标；其收入、成本和期间口径以正式规则为准，不能直接视为投资者可得回报。
风险维度	Risk Dimension	指罗盘中观察风险来源与风险暴露的维度。行业罗盘的风险因子，风险越高，得分越低；该得分不直接代表违约概率或预计损失率。
规则匹配	Rule Matching	指任务执行前确认适用规则、版本、权限及升级条件的过程。发现规则缺口或权限冲突时，应提交处理。
规则版本	Rule Version	指对某一时期适用的规则内容及修订状态的可识别标记，用于说明一次执行采用的约束、定义或评分依据。
规则层	Rules Layer	指组织三层设计中确定允许事项、禁止事项、权限条件、交付标准和升级要求的层次。
运行编码	Run Identifier	指一次智能体实例运行的标识，用于关联所属任务、输入输出、权限、调用、版本和检查记录。重做或重新执行应分别保留运行记录。
评分规则	Scoring Rules	指将观察结果转换为分值的规则，包括评分方向、分档区间、端点归属及缺失值、异常值的处理方式。
评分口径	Scoring Specification	指对评分所依据的对象、统计范围、计算方式、单位和期间的统一规定，使相同事实能够被一致计算。
七站作业流程	Seven-Stage Task Workflow	指任务依次或按依赖关系经过接单、规则匹配、组队、执行、验证、收口和归档的运行流程，用于明确各环节输入、输出、继续条件和异常去向。
说明性模拟编码 (SIM)	SIM / SIM-01	本书用于标识说明性模拟任务；SIM-01 不是后台真实运行编号，V0.1、V0.2 亦为模拟版本。

中文名词	英文全称	释义
六审	Six Design Reviews	六鉴对应的六项设计审查的合称；六鉴是视角名称，六审是工作合称。执衡另行承担资格判断与组合筛选。
六鉴	Six Review Perspectives	指问业、量尺、寻源、分水、照影和破局六类独立检验视角，分别审查候选因子的经营意义、可计算性、数据条件、区分依据、信息增量和抗操纵性。
库存单位 (SKU)	Stock Keeping Unit	用于区分具体商品品项的库存管理单位；SKU 动销率的商品范围、期间及计算口径以正式规则为准。
专项小队	Task-Specific Team	指围绕特定任务目标组成的临时协作单元。任务完成后可以解散，成员的岗位身份、真人认领关系及执行记录继续保留。
牛码 SHARP 双罗盘	The NUMA SHARP Twin Compasses	指由通用罗盘和行业罗盘组成的并列观察框架，分别呈现合同结构与具体经营的收益支撑和风险负载。
三类行业视角	Three Industry Perspectives	指正常经营者、老炮和骗子三类观察立场，分别关注日常经营机制、行业经验与隐性约束、指标包装及规则利用路径。
三个投资者视角	Three Investor Perspectives	指分别从生意、价格和风险出发，分析价值创造及持续性、投资价格条件和潜在不利结果的三类判断立场。
十二个足够	Twelve Sufficiency Conditions	指帽子军团成立所需的四组条件：任务足够小、窄、碎；底座足够多元、对冲、有性价比；编制足够多、细、冲；检验足够密、长、深。具体解释见附录二，不代表统一数值门槛。

表 附 3-1：名词释义与中英文对照

香港中环康乐广场 8 号
交易广场 2 期
2105-2108 室

滴灌通集团网站: www.microconnect.com
邮箱: info@microconnect.com