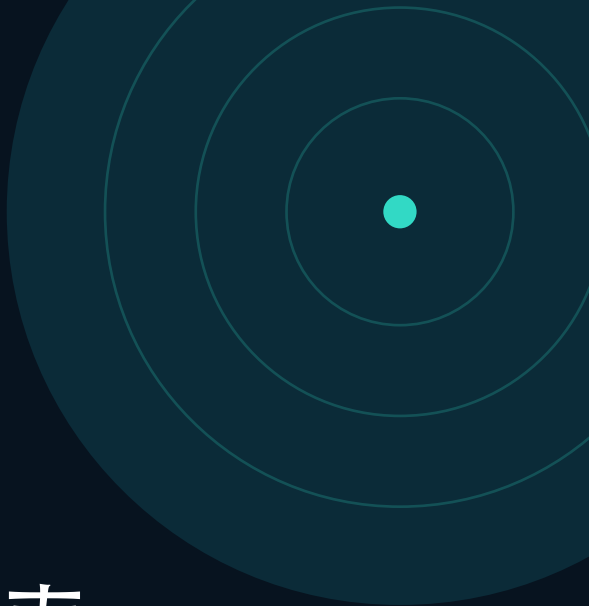




BUILDTABLE.AI / AI DATA ACADEMY

企业业务语义层建设指南

让 AI 真正看懂企业自己的业务语言

把指标口径、字段含义、业务规则、术语、文档、偏好和专家经验组织成 AI 可检索、可引用、可治理的企业上下文。

VERSION 1.0 / 2026-08-11

目录

执行摘要：AI 能看见数据，为什么仍然不懂业务

3

业务语义层的边界

6

语义对象模型

8

术语、同义词和企业黑话

11

业务规则与例外条件

14

业务文档与专家经验

15

分析偏好与回答风格

16

表关系、Join 路径与数据粒度

17

权限、范围与时间有效性

18

上下文检索与组装

19

标准问题集与语义评测

20

BuildTable 实施 workflow

21

从语义层走向 AI 数据智能体

22

30 天落地路线图

24

业务术语卡片

25

业务规则卡片

26

标准问题集

27

发布前自查

28

执行摘要：AI 能看见数据，为什么仍然不懂业务

大模型可以识别自然语言，数据库可以提供结构化数据，但二者之间仍缺少企业自己的业务上下文。老板说“有效销售”，运营说“高潜会员”，供应链说“健康库存”，这些词在不同企业、部门和时期可能有完全不同的定义。

业务语义层把问题中的业务语言连接到指标、字段、关系、规则、权限和文档。它不是一段固定提示词，而是一套可持续维护的上下文系统：在正确时间，为正确问题选择正确知识，并保留来源和版本。

很多企业第一次做 AI 问数时，最容易低估一个问题：AI 能连上数据库，不代表 AI 能理解业务。

数据库里有表、有字段、有数据，但业务问题往往不是按数据库结构提出的。老板问“这个月销售额为什么下降”，运营问“上周高潜客户转化怎么样”，区域经理问“华南门店库存风险在哪里”，这些问题背后都包含大量默认上下文。

AI 如果只看到字段名，很容易把问题理解错。

数据库结构不等于业务语言

企业数据库中的字段通常来自系统实现，而不是业务表达。

例如：

- amount 可能是下单金额、支付金额，也可能是退款后净额。
- status=3 可能代表已完成，也可能代表已取消。
- created_at 可能是订单创建时间，业务复盘却需要按支付时间统计。
- shop_id、store_id、org_id 可能都和门店有关，但含义并不相同。

业务人员不会按这些字段名提问。他们会说“本月 GMV”“有效订单”“高潜客户”“华南大区”“重点门店”。如果 AI 不知道这些词和数据结构之间的关系，就只能猜。

同一个指标背后有很多口径选择

以“销售额”为例，真正进入分析前，至少要明确：

- 是否扣除退款？

- 是否包含优惠券、满减、平台补贴？
- 是按下单时间、支付时间，还是发货时间统计？
- 是否排除测试订单、异常订单、内部订单？
- 多渠道、多系统数据重复时，以哪个系统为准？

这些判断不是 SQL 技巧，而是企业自己的业务规则。没有语义层，AI 很可能生成一条“看起来能跑”的 SQL，但结果并不是企业真正认可的业务答案。

AI 问数还必须理解权限和身份

企业数据分析不是所有人看同一张总表。

同一个问题“这个月销售额是多少”，总部、区域经理、门店店长、品牌负责人看到的范围可能完全不同。有人能看全国，有人只能看自己区域；有人能看金额，有人只能看趋势；有人能看客户明细，有人只能看聚合结果。

如果 AI 不知道提问者是谁、属于哪个组织、能访问哪些字段和行，就不能稳定给出合规答案。

业务语义层解决什么

业务语义层的作用，是把企业自己的业务语言翻译成 AI 可以使用的上下文。

它至少要告诉 AI：

- 表和字段在业务上表示什么。
- 核心指标如何定义和计算。
- 字段值、状态码、内部简称分别代表什么。
- 哪些维度可以和哪些指标一起分析。
- 哪些业务文档、分析习惯和专家经验需要被复用。
- 不同用户在什么权限边界内使用数据。

有了这些上下文，AI 才不是在数据库里盲猜，而是在企业已定义的业务语言里回答问题。

AskTable 如何承接

AskTable 的价值不只是把自然语言转成 SQL，而是让 AI 在企业真实数据、业务语义和权限边界之上工作。

在 AskTable 产品体系里，BuildTable 负责把数据采集回来并建成 AI 可用的数据基础；AskTable 通过表备注、字段备注、字段值索引、业务文档、指标口径、分析偏好和权限上下文，让 AI 更稳定地理解业务问题、生成查询、解释结果。

所以，AI 问数的第一步不是“连数据库”，而是先让 AI 看懂企业自己的业务语言。

02

业务语义层的边界

很多企业已经有数据字典，也有指标口径文档。为什么到了 AI 问数阶段，还要讨论“业务语义层”？

原因在于：数据字典和指标体系解决的是数据治理的一部分问题，但 AI 要稳定参与业务分析，需要更完整的上下文。

数据字典主要解释字段

数据字典通常回答这些问题：

- 这张表叫什么？
- 这张表来自哪个系统？
- 字段名是什么？
- 字段类型是什么？
- 字段大概代表什么含义？

它的对象主要是表和字段。对数据开发、报表开发、系统集成很有帮助，但它通常不足以支撑 AI 独立理解业务问题。

因为业务人员提问时不会说“查询 pay_amount 字段按 province_code 分组”，而是会问“各区域本月销售额同比怎么样”。这中间缺少的是业务概念、指标规则、维度关系和分析习惯。

指标体系主要统一看什么数、怎么算

指标体系解决的是另一类问题：

- 企业要统一管理哪些核心指标？
- 指标从哪些业务目标拆出来？
- 指标的定义、公式、过滤条件是什么？
- 指标按哪些维度观察？
- 指标由谁负责维护？

指标体系是 AI 问数的重要前提。没有统一指标口径，AI 即使生成了 SQL，也可能让不同部门看到不同版本的“销售额”“转化率”“复购率”。

但指标体系仍然不是完整语义层。它重点管理指标，不一定覆盖字段值、状态码、业务黑话、业务文档、用户偏好和权限上下文。

业务语义层是 AI 可用的上下文

企业业务语义层不是简单的数据字典升级，而是把指标口径、字段含义、业务规则、权限范围和专家经验，组织成 AI 可以稳定理解和复用的上下文。

它覆盖六类信息：

1. 数据结构语义：表、字段、字段类型、字段业务含义。2. 指标口径语义：指标定义、计算规则、过滤条件、时间口径。3. 业务实体语义：客户、门店、商品、渠道、组织、区域之间的关系。4. 字段值语义：枚举值、状态码、别名、简称、内部黑话。5. 业务经验语义：业务文档、分析偏好、专家判断规则。6. 权限上下文：谁能看什么数据、按什么组织范围看、哪些字段要脱敏。

一句话：数据字典是给人和系统看字段，指标体系是给企业统一业务数字，业务语义层是给 AI 看懂企业业务上下文。

三者应该怎样协同

比较合理的建设顺序是：

1. 先建设指标体系，明确企业要统一看什么数、怎么算。2. 再补齐业务语义层，把指标、字段、业务规则和专家经验组织成 AI 上下文。3. 再叠加权限治理，让 AI 在不同用户身份下返回正确范围的答案。4. 最后把稳定分析流程沉淀成 AI 数字员工，自动跑日报、复盘、预警和触达。

这也是“企业 AI 数据分析基础设施系列”的核心逻辑。

AskTable 如何承接

AskTable 的语义配置不只是维护字段说明，而是把 AI 问数需要的上下文集中组织起来：表字段备注、指标口径、字段值索引、业务文档、分析偏好、用户身份和权限范围。

这使 AI 不只是“会写 SQL”，而是更接近企业内部懂业务规则的数据分析助手。

数据字典主要解释表和字段，指标体系统一经营数字，知识库存放更广泛的文档，RAG 是检索技术。业务语义层位于它们之间，负责把与分析有关的对象、规则和上下文连接起来，并决定哪些信息在一次问题中应被使用。

03

语义对象模型

AI 问数要准确，不能只依赖模型自己猜。企业需要主动把关键业务知识告诉 AI。

这些知识不一定都很复杂，但必须足够明确、可复用、可治理。否则 AI 每次回答都像临时理解一次业务，很难稳定。

第一类：表和字段的业务含义

最基础的是表备注和字段备注。

AI 需要知道：

- 这张表记录什么业务过程？
- 一行数据代表一个订单、一个客户、一次访问，还是一个库存快照？
- 字段在业务上表示什么？
- 字段单位是什么？
- 字段能否为空，空值通常代表什么？
- 字段是否可用于分析，还是只是系统内部标识？

例如库存表最关键的不是“有库存数量字段”，而是这张表是实时库存、日末库存，还是月末快照。如果 AI 不知道这一点，就可能把多个日期的库存相加，得到完全错误的答案。

第二类：指标定义和计算规则

业务指标不能只靠字段名推断。

一个合格的指标口径至少要包含：

- 指标名称：例如 GMV、销售额、复购率、库存周转天数。
- 业务含义：这个指标用来衡量什么。
- 计算公式：分子、分母、聚合方式。
- 过滤条件：是否排除退款、测试订单、异常订单。
- 时间口径：按下单时间、支付时间、发货时间还是确认收货时间。
- 统计周期：日、周、月、自然月、财务月。

- 数据来源：来自哪个系统、哪张表、哪些字段。
- 责任人：谁负责解释和维护这个口径。

这些信息越明确，AI 越不需要在每次提问时重新猜测。

第三类：维度和实体关系

很多分析问题不是只算一个总数，而是要按维度拆解。

企业需要告诉 AI：

- 哪些维度可以用于分析：时间、区域、门店、商品、渠道、客户、活动。
- 维度之间是什么关系：门店属于区域，商品属于类目，客户属于会员等级。
- 哪些指标适合按哪些维度看。
- 哪些维度需要做层级下钻。

例如“销售额下降原因”通常需要按时间、区域、商品、渠道、活动逐层拆解。如果没有维度关系，AI 只能生成单次查询，很难形成分析路径。

第四类：口径优先级和例外规则

企业内部常常存在多个系统、多个口径。语义层需要记录优先级。

例如：

- 财务收入以财务系统为准，经营销售以订单系统为准。
- 会员等级以 CRM 为准，门店归属以组织系统为准。
- 广告花费以平台账单为准，投放效果以内部归因表为准。
- 历史老门店合并到新门店时，需要按组织映射表处理。

这些规则如果不进入语义层，AI 就会在多个相似字段之间摇摆。

第五类：权限相关的业务规则

AI 使用数据时还要知道权限边界。

语义层里应当包含：

- 哪些表和字段对哪些角色可见。
- 哪些用户只能看自己组织范围。

- 哪些字段需要脱敏。
- 哪些指标只能以聚合方式展示。
- AI 数字员工自动执行任务时沿用哪个身份权限。

这类信息不是单纯安全配置，也会影响 AI 如何理解问题、选择字段和呈现答案。

AskTable 如何承接

AskTable 支持通过元数据、字段备注、字段值索引、业务文档、指标口径和权限配置，把这些知识组织成 AI 可使用的上下文。

企业不需要一次性把所有知识整理完。更现实的方式是从高频场景开始，比如经营日报、广告复盘、库存预警，先把这些场景需要的指标、字段、维度、规则和权限补齐，再逐步扩展到更多业务域。

建议至少管理七类语义对象：业务实体、字段、指标、维度、关系、规则和文档。每个对象都应有唯一标识、业务名称、技术映射、负责人、版本和适用范围，并能与其他对象建立明确关联。

业务实体表示客户、订单、门店、商品等现实对象；字段描述属性；指标表达可计算结果；维度提供分析视角；关系说明对象如何连接；规则处理过滤与例外；文档补充数据库之外的背景。AI 回答不是只找一个字段，而是在这些对象之间构建可验证路径。

术语、同义词和企业黑话

企业里大量业务知识并不在数据库结构里，而是在文档、经验、简称和日常沟通中。

如果 AI 只看表名和字段名，很难理解这些隐藏知识。业务语义层要做的，就是把这些分散知识变成 AI 能检索、能引用、能复用的上下文。

业务黑话是 AI 问数的常见误差来源

每个企业都有自己的业务黑话。

例如：

- “高潜在客户”可能是近 30 天访问 3 次以上但未购买的人群。
- “沉睡用户”可能是 180 天未复购会员。
- “有效订单”可能要排除退款、测试、内部采购和风控拦截订单。
- “重点门店”可能来自运营手工维护的门店清单。
- “BG1”可能是某个业务组、渠道组或品牌组。

这些词对内部人员很自然，但对 AI 并不天然可见。没有解释时，AI 只能按字面理解，或者在字段里碰运气。

字段值索引用来解决“值”的理解问题

很多时候，AI 不只是找不到字段，而是找不到字段里的具体值。

例如用户问：

- “生意参谋渠道的转化怎么样？”
- “罗盘里华南大区库存风险如何？”
- “FILA 门店和始祖鸟门店本月表现差异？”

AI 需要知道“生意参谋”“罗盘”“华南大区”“FILA”“始祖鸟”这些词，分别可能对应哪些字段值、别名或映射关系。

字段值索引的价值在于，把重要文本字段里的业务值建立可检索索引，让 AI 能用自然语言找到对应记录，而不是要求用户准确输入数据库里的标准编码。

业务文档补充数据库外的知识

很多分析规则并不适合放进字段备注。

例如：

- 经营日报模板。
- 广告复盘方法。
- 库存风险判断规则。
- 会员分层定义。
- 活动复盘口径。
- 行业平台字段解释。
- 内部分析 SOP。

这些内容更适合以业务文档形式沉淀。AI 在回答问题时，可以先召回相关文档，再结合表字段和指标口径完成分析。

分析偏好决定 AI 如何表达答案

同一个数据问题，不同团队希望看到的答案形式可能不同。

有的管理者希望先看结论，再看数据；有的运营希望先看异常，再看原因；有的销售团队希望按门店排行；有的供应链团队希望按风险等级分组。

分析偏好可以沉淀为：

- 回答结构：先结论、再原因、最后建议。
- 排序习惯：优先看下降最多、风险最高、贡献最大。
- 对比方式：同比、环比、目标达成率、预算消耗率。
- 图表偏好：趋势图、排行表、结构占比、漏斗图。
- 追问路径：先看整体，再按区域、商品、渠道下钻。

这些偏好让 AI 更像企业自己的分析助手，而不是每次都给出通用答案。

AskTable 如何承接

AskTable 的业务文档、字段值索引、语义检索和分析偏好，可以把数据库之外的知识纳入 AI 问数流程。

当用户用内部黑话提问时，AI 可以先通过语义检索找到相关字段、字段值和业务文档，再结合权限边界生成回答。这样，企业不需要把所有业务知识硬编码进 SQL，也不需要每个员工都学会数据库字段名。

同义词表不能只是简单替换。需要记录术语所属领域、适用部门、可能歧义、优先解释和反例。例如“客户”在销售系统中可能指合同主体，在零售分析中可能指会员，在财务系统中可能指应收对象。

字段值索引用于把自然语言中的品牌、区域、活动、产品昵称映射到真实值。索引应支持别名、拼写差异和历史名称，同时避免把相似词错误合并。

业务规则与例外条件

业务规则决定问题如何落到数据上。常见规则包括有效订单判定、客户归属、渠道拆分、库存健康阈值、异常排除和活动周期。规则若只存在于 SQL、Excel 或员工记忆里，AI 很难稳定复用。

规则卡应记录触发条件、处理逻辑、优先级、适用对象、例外、生效时间、负责人和示例。多条规则冲突时，需要明确是更具体规则优先、更新版本优先，还是由人工确认。

临时政策必须有失效条件。例如促销期特殊折扣、某区域临时归属、系统迁移期间双口径并行，都不应永久污染默认语义。AI 在引用临时规则时，应能说明规则名称和有效期。

规则验收问题

- 同一问题在规则生效前后是否得到正确结果。
- 规则冲突时是否按照既定优先级处理。
- 例外对象是否被准确包含或排除。
- 答案能否显示引用了哪条规则。

业务文档与专家经验

企业知识大量存在于 SOP、制度、产品手册、活动方案、检修手册和复盘文档中。语义层不应把文档全文无差别塞进上下文，而应抽取与分析有关的对象、规则、阈值、步骤和判断依据。

每份文档需要记录来源、发布部门、版本、有效期、适用范围和可信等级。已废弃制度、会议草稿和未经确认的经验不能与正式规则等价。引用文档时，应优先使用有效、权威、与当前问题最相关的版本。

专家经验可以沉淀为判断卡：当哪些指标出现什么组合时，通常意味着什么原因；下一步应查看哪些维度；哪些结论只能作为线索，不能直接视为因果。这样 AI 才能提供有边界的分析建议。

分析偏好与回答风格

不同角色需要不同答案。高管关注结论、影响和建议；运营需要拆解维度与可执行动作；财务关注口径、凭证和差异；数据团队关注 SQL、来源和质量。语义层应允许按角色配置默认分析深度和表达方式。

分析偏好包括默认时间范围、对比基准、排序方式、图表类型、精度、单位和是否展示明细。偏好不是硬编码答案，而是在不改变事实和口径的前提下，让结果更符合用户工作方式。

当问题缺少关键条件时，应定义追问策略。例如“销售怎么样”至少可能缺少时间、范围和指标；AI 可以根据用户角色补默认值，也可以追问。默认值必须可见，不能静默猜测。

表关系、Join 路径与数据粒度

数据库可连接不代表关系可安全使用。同一客户可能在会员、订单和售后系统中有不同主键；订单与商品是一对多，订单与支付也可能一对多。错误 Join 会造成重复计算，结果看似合理却被放大。

语义层需要记录实体关系、主外键、基数、推荐 Join 路径和适用指标。存在多条路径时，应说明哪条用于交易分析、哪条用于组织归属、哪条只适用于当前快照。

每个指标都要声明粒度。订单级指标与订单明细级维度组合时，需要先聚合或去重；月度快照不能直接与日流水相加。AI 在生成查询前，应检查粒度兼容性和扇出风险。

关系测试

- 主键是否唯一、空值是否可接受。
- Join 后行数和指标是否异常放大。
- 历史归属与当前归属是否区分。
- 多条路径是否有明确选择规则。

权限、范围与时间有效性

语义层本身也受权限控制。某些指标定义可以公开，但字段、规则或文档只对特定角色可见。AI 选择上下文时必须同时满足数据权限与知识权限，避免通过解释泄露敏感信息。

语义对象应支持生效时间和版本。组织架构、商品分类、政策规则和术语定义都会变化。回答历史问题时，系统需要知道应使用历史语义还是当前语义，并在必要时向用户说明。

权限与版本要可审计：谁在何时发布了哪一条语义，哪些问题使用了它，修改后影响哪些测试和下游 Agent。没有审计链，语义层会逐渐变成难以信任的提示词集合。

上下文检索与组装

语义层的关键不是“存了多少知识”，而是“每次问题选对多少知识”。把全部表结构、指标和文档一次性发送给模型，会增加成本、噪声和冲突，反而降低准确性。

推荐先识别问题中的实体、指标、时间和意图，再分别检索相关字段、关系、规则与文档，最后按权威性、相关性、版本和权限排序。上下文应尽量小而完整，并保留引用标识。

当多个对象冲突时，系统应优先使用正式发布、当前生效、范围更具体的定义；无法自动判断时，向用户追问，而不是选择一个看似合理的解释。

组装顺序示例

用户与权限 → 问题意图 → 指标和实体 → 维度与关系 → 规则与例外 → 文档与偏好 → 查询与回答。

标准问题集与语义评测

语义层不能只靠人工浏览验收。企业需要一组真实问题，记录预期理解、预期指标、关键过滤条件、正确数据范围和可接受答案。每次语义或模型变化后自动回归。

评测至少分四层：理解层判断是否识别正确实体与意图；查询层判断字段、Join、过滤和聚合是否正确；答案层判断数字、图表和解释是否满足问题；追溯层判断是否能引用定义、规则和来源。

错误归因比单一准确率更重要。是术语歧义、指标缺失、关系错误、权限不足、数据质量问题，还是模型生成问题？只有归因清楚，团队才知道应该修改语义、数据模型还是提示策略。

推荐指标

- 意图识别正确率与澄清率。
- 指标、维度和过滤条件命中率。
- SQL 执行成功率与结果一致率。
- 权限违规数和敏感信息泄露数。
- 答案引用覆盖率与用户采纳率。

BuildTable 实施 workflow

BuildTable 先读取表、字段、类型、主键和候选关系，建立技术元数据基础；再由业务与数据团队补充实体、指标、维度、术语、规则和文档。自动识别提高效率，人工确认保证业务含义。

配置完成后，用标准问题集验证字段选择、Join 路径、口径和权限。问题失败时记录错误类型并修正对应对象，而不是不断堆叠全局提示词。通过评测的语义版本才能发布给 AskTable 等下游 Agent。

发布后持续收集真实问题、未命中术语、失败查询和用户反馈，形成候选改进项。每次改动都要经过影响分析和回归测试，使语义层成为可运营产品，而不是一次性项目文档。

从语义层走向 AI 数据智能体

业务语义层的价值，不只是让 AI 问数更准。更进一步，它可以把企业高频数据分析流程沉淀成 AI 数据智能体，让日报、复盘、预警和触达自动运行。

前提是：AI 必须先有稳定的指标、语义和权限基础。

没有语义层，数字员工只能做表面自动化

很多企业一开始会把数字员工理解成“定时跑报表”。

但真正有价值的自动化分析，不只是把一张表按时发出去，而是能持续回答：

- 今天哪些指标异常？
- 异常发生在哪些区域、商品、渠道或门店？
- 和目标、预算、上周、去年同期相比差多少？
- 可能原因是什么？
- 哪些对象需要跟进？
- 下一步应该通知谁、生成什么动作？

这些判断都依赖业务语义层。如果 AI 不知道指标口径、字段含义、业务规则、字段值和权限范围，就无法稳定执行分析任务。

语义层让高频分析流程可复用

有了语义层之后，企业可以把高频分析流程拆成可复用步骤。

例如经营日报：

1. 读取当天核心指标。
2. 对比目标、昨日、上周同期和去年同期。
3. 找出异常指标。
4. 按区域、门店、商品、渠道下钻。
5. 给出原因线索和跟进建议。
6. 按组织权限发送给对应负责人。

这套流程里，指标口径、维度关系、异常阈值、分析偏好、组织权限都可以沉淀下来。下一次运行时，AI 不需要从零理解业务。

适合沉淀成 AI 数据智能体的场景

优先选择高频、规则相对稳定、结果需要持续跟进的场景。

典型场景包括：

- 经营日报：自动汇总收入、利润、订单、转化、客单价和异常波动。
- 广告复盘：自动分析花费、曝光、点击、转化、ROI 和预算消耗。
- 库存预警：结合库存快照、近 7/30/90 天销量和在途库存识别风险。
- 会员运营：按会员等级、活跃度、复购周期识别触达人群。
- 活动复盘：分析活动前后指标变化、商品贡献和渠道效果。
- 门店巡检：定期发现门店销售、客流、库存、履约异常。

这些场景都有共同点：业务人员经常问，分析路径相对固定，但每次人工做都很耗时。

从语义层到数字员工的建设路径

建议按四步推进：

1. 选一个高价值场景：不要一开始追求覆盖全公司。 2. 补齐语义上下文：指标、字段、维度、业务文档、字段值、权限。 3. 固化分析流程：把专家的分析步骤、判断规则和输出格式沉淀下来。 4. 配置定时和分发：按岗位、组织和权限，把结果发送给对应人员。

这样建设出来的 AI 数据智能体，才不是简单“定时问一次 AI”，而是基于企业自己的语义和规则稳定执行。

AskTable 如何承接

在 AskTable 产品体系中，BuildTable 负责把数据采集、存储和建模为 AI 可用的数据基础；AskTable 负责语义配置、权限内问数和分析；AI 数字员工则把已经验证过的分析流程沉淀为可调度的 Skill / Solution。

这条链路的关键是：先让 AI 看懂企业业务语言，再让 AI 复用专家分析方法，最后让 AI 自动执行高频数据工作。

AskTable 可以作为语义层的下游应用示例：它消费 BuildTable 发布的数据模型、指标、语义和权限，为业务用户提供问数、报告、复盘与预警。BuildTable 负责建设可复用的数据基础，下游 Agent 负责在具体场景中交付体验和动作。

30 天落地路线图

第一周选择一个高价值场景，收集 20-50 个真实问题，确认用户、数据范围和验收人。不要从全公司数据字典开始，而要从业务问题反推所需实体、指标和规则。

第二周建立核心模型和语义对象：表字段说明、实体关系、指标维度、术语、规则和权限。优先覆盖高频问题，未使用的知识暂不纳入。

第三周运行标准问题集，按理解、查询、答案、追溯四层评测；修正歧义、Join、粒度和规则冲突。邀请真实用户试用并记录追问与未命中表达。

第四周发布首个语义版本，接入下游 Agent，建立反馈、变更、回归和月度复盘机制。完成标准不是“知识录入完毕”，而是高价值问题能稳定得到可信答案。

附录 A

业务术语卡片

建议字段：术语、别名、业务定义、所属领域、对应实体/指标/字段、适用部门、歧义说明、反例、负责人、生效日期和版本。

字段	示例	填写要求
术语	高潜会员	使用业务现场真实表达
别名	潜客、高价值潜客	收录简称、别名和常见错写
业务定义	最近 30 天有互动且预测购买概率高于阈值的会员	定义必须可判断或可计算
所属领域	会员运营	用于缩小检索范围
对应对象	member、purchase_probability	关联实体、指标和字段
适用部门	会员运营、门店导购	防止跨部门误用
歧义说明	销售部门“潜客”指尚未签约企业客户	写明同词异义
反例	已流失且无有效联系方式的会员	帮助模型识别边界
负责人	会员运营负责人	对含义和变更负责
版本	V1.0，2026-08-01 生效	支持历史追溯

附录 B

业务规则卡片

建议字段：规则名称、触发条件、处理逻辑、优先级、适用范围、例外、引用文档、负责人、生效与失效时间、测试问题。

字段	示例
规则名称	有效订单判定
触发条件	查询订单、销售额、客单价等交易指标
处理逻辑	支付成功，排除测试单与已全额退款订单
优先级	正式财务口径高于活动临时口径
适用范围	中国大陆直营网店
例外	预售订单按尾款完成日计入
引用文档	《销售核算制度》V3.2
生效时间	2026-01-01 至新版本发布
负责人	财务 BP / 销售运营
测试问题	本月有效订单数是多少？

附录 C

标准问题集

每条问题记录：用户角色、自然语言问题、预期意图、预期指标、维度、过滤、权限范围、预期答案、允许误差、引用要求和失败类型。

项目	示例
用户角色	华东区域经理
问题	上周哪些门店销售下降最明显，主要原因是什么？
预期指标	实付销售额、订单数、客单价
预期维度	门店、日期、商品品类
默认过滤	华东区域、有效订单
权限范围	仅所辖门店
预期行为	先排名，再拆解订单数与客单价
引用要求	展示指标口径、时间范围和数据更新时间
失败类型	意图错误 / 指标错误 / Join 错误 / 权限错误

附录 D

发布前自查

- 核心术语是否存在歧义与同义词。
- 指标、实体、维度和关系是否连通。
- 规则是否有范围、优先级和有效期。
- 文档是否有来源、版本和可信等级。
- 权限与知识范围是否同时生效。
- 标准问题集是否完成回归测试。

发布评审应同时由业务负责人、数据负责人和权限负责人参加。任何影响指标口径、组织范围或敏感数据的未解决问题，都不应通过“增加提示词”绕过。

关于 BuildTable

BuildTable 帮助企业采集、理解、建模并发布 AI 友好的数据模型，把表、字段、关系、指标、语义和权限沉淀为可持续运营的数据基础。

了解更多：buildtable.ai / 联系我们：扫码网站右下角企业微信