

奥腾科技寻数问数智能体产品白皮书

产品名称：寻数问数智能体（dataAssetAgent）

文档版本：V1.0

编制单位：上海奥腾计算机科技有限公司

日期：2026 年 8 月



目 录

第一章 产品概述.....	3
第二章 产品架构.....	4
第三章 核心能力.....	5
第四章 安全与合规.....	6
第五章 工程特性.....	7
第六章 典型应用场景.....	7
第七章 服务保障.....	8
第八章 结语	8
第九章 联系我们.....	8

第一章 产品概述

1.1 产品定位

寻数问数智能体是基于 Ottomi-Nexus 多模态 AI 数据中台智能中心打造的对话式数据智能助手。用户以自然语言提问，智能体自动完成意图解析，分派至对应领域的子智能体，通过系统内置工具实时查询数据目录、元数据、质量结果与资产画像，最终汇总为一份包含答案、图表与建议的结构化回复。

产品综合运用数据目录编制、数据质检、多维资产评估三大能力底座，并结合大模型通识、向量知识库与平台实时数据，实现“找数—问数—分析—建议”全链路对话式服务：

- **找数**：语义检索五类数据资产（数据源表、数仓表、指标、用户标签、数据 API），候选清单确认后自动输出完整表结构与元数据；
- **问数**：在线分析数据，自动生成 DAMA 六维度质量评分、质量详情图与 ECharts 图表；
- **分析**：以九维资产评估雷达图刻画数据资产健康度，结合指标与标签语义底座给出口径清晰的分析结论；
- **建议**：给出数据开发与数据质检的组件化方案，命令可直接复制到 AI 画布执行。

1.2 建设背景

传统数据用数方式存在四个突出问题：

找数难：业务人员记不清表名与口径，跨系统翻找数据资产耗时费力；

口径乱：同名不同义、同义不同名，报表数字对不上，跨部门对齐靠人工核对；

分析门槛高：数据质量评估与分析依赖 SQL 技能与专业工具，业务人员难以自助完成；

AI 外挂风险大：通用大模型直接接数据存在幻觉与泄密风险，回答缺乏实时数据支撑。

智能问答的准确性依赖两层底座：**权限决定可见范围，指标与标签体系决定命中准确度**。寻数问数智能体内置于数据中台，与数据权限、用户安全等级同源同底座，问答范围只覆盖已上架数据资产，未治理数据不进入回答，从机制上规避了上述问题。

1.3 核心价值

找得准：语义检索 + 指标标签语义底座 + 上架资产口径，候选结果可靠、口径清晰

答得实：回答严格基于内置工具实时查询的真实数据，防幻觉机制兜底，缺失时如实告知

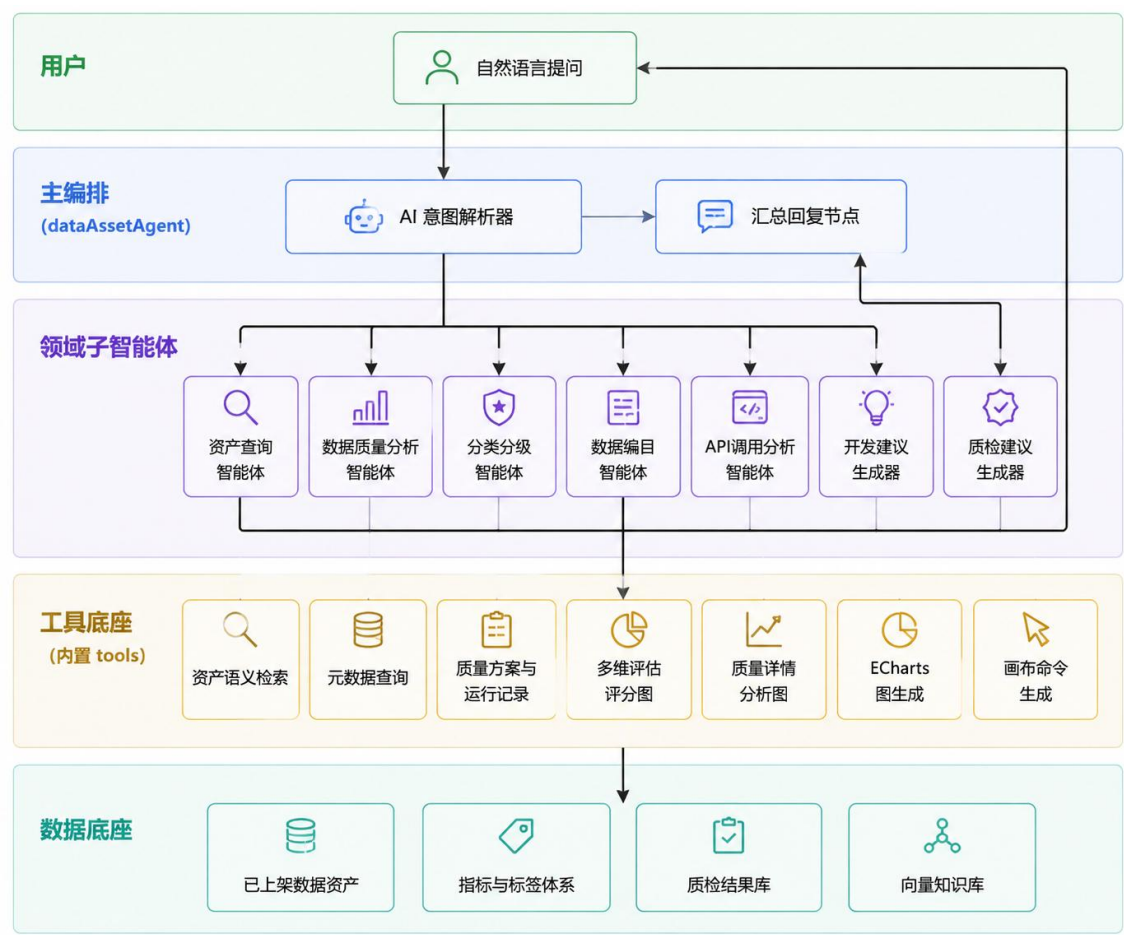
用得合规：四级数据权限与用户安全等级直达问答结果，个人信息全角色强制脱敏，全程留痕可审计

成本可控：工具化调用完成取数与计算，大批量数据由平台引擎处理、不经过大模型，token 消耗少

建议可执行：开发与质检建议以 142 个画布组件形式输出，兼容多种数据库，命令可直接复制执行

第二章 产品架构

2.1 总体架构



2.2 智能体编排体系

层级	名称	职责
主编排	dataAssetAgent	意图解析、子智能体分派、结果汇总回复
子智能体	资产查询智能体	资产检索、元数据查询、明细与可视化
子智能体	数据质量分析智能体	DAMA 六维度评估、根因分析、质量报告
子智能体	分类分级智能体	敏感数据识别、分级管理、脱敏建议
子智能体	数据编目智能体	注释补全、重复校验、编目回写
子智能体	API 调用分析智能体	API 使用情况统计与分析
子智能体	开发建议生成器	组件化开发方案与画布命令生成
子智能体	质检建议生成器	质检方案与质检画布命令生成

意图解析器将用户问题映射为七个领域中的一个或多个；各子智能体独立取数；汇总节点严格基于各子智能体返回结果整合回答，禁止编造。

2.3 工具底座

工具底座全部内置于数据中台，与平台数据同源直连，问答不依赖外部数据通道：

工具	用途
----	----

工具	用途
资产语义检索	按关键词/模糊语义检索五类数据资产，返回候选清单
元数据查询	拉取表结构、字段、类型、注释、主键、分区等完整元数据
质量方案与运行记录	生成质检方案、查询历史运行记录与规则结果
脏数据预览	抽样查看脏数据内容，定位错误模式
多维评估评分图	生成九维资产评估雷达图
质量详情分析图	生成单表质量详情图（规则通过率、脏数据分布、六维得分）
ECharts 图生成	将结构化分析结果渲染为图表
画布命令生成	生成可复制到 AI 数据开发/质检画布的命令式指令

第三章 核心能力

3.1 对话式找数

- **五类资产统一检索**：数据源表、数仓表、指标、用户标签、数据 API 一个入口检索；
- **模糊匹配 + 候选确认**：记不清完整名称也能检索，先返回候选清单供确认，再查询详情；
- **元数据强制查询**：确认数据表后自动输出完整表结构（字段、类型、注释、主键、分区），不可省略；
- **上架资产口径**：问答范围只覆盖已上架数据资产，未治理数据不进入回答，保障答案质量。

3.2 在线问数与图表

以 DAMA-DMBOK 国际数据管理知识体系为评估框架：

维度	含义	常见度量指标
准确性	数据正确反映真实世界的程度	值域合规率、与权威参考源一致率
完整性	所需数据的存在程度	非空率、必填字段填充率
一致性	同一数据在不同系统的等价程度	跨表字段值匹配率
及时性	数据反映当前时点的程度	更新延迟时长、SLA 达标率
有效性	数据符合格式与取值范围的程序	格式合规率、枚举值命中率
唯一性	实体在数据集中仅出现一次的程度	主键重复率、实体去重率

- 质量分析输出综合评分与等级（90 分以上优秀、70 - 89 良好、60 - 69 一般、60 以下较差）；
- 自动生成质量详情分析图与趋势图、通过率饼图等 ECharts 图表；
- 支持脏数据抽样预览，按 RCA 四层模型（表层—模式层—流程层—系统层）分析根因。

3.3 多维资产评估

资产多维评估评分图以雷达图形式综合刻画数据资产健康度，覆盖九个维度：

资产综合得分、编目完整性、质检健康率、敏感等级、调用次数、引用次数、授权数、主题覆盖度、更新频率。

评估结果帮助业务方快速判断数据资产可用程度，结合数据目录与质检结果，给出建议使用的数据表与数据主题（明细表、宽表、汇总表、标签表、结果表）。

3.4 数据开发与质检建议

- **组件化方案：**开发建议基于 142 个画布组件生成，覆盖数据输入、融合、转换、输出、脚本、运算等类型，不绑定数据库类型；
- **性能保障：**组件自动调用并行计算引擎处理大批量数据，性能更有保障；
- **可复制命令：**建议以命令式指令输出，可直接复制到 AI 数据开发画布或质检画布执行；
- **质检方案：**将业务需求翻译为质检规则配置与画布命令，区分单表质检与多表一致性检查。

3.5 通识、知识库与实时数据

- **大模型通识与向量知识库结合，**为问题理解与方案建议提供背景知识；
- **平台实时数据通过内置工具获取，**回答基于当下真实状态；
- **防幻觉机制：**回答严格基于工具查询结果，数据缺失时如实告知并反问补充，绝不编造表名、字段与数值；
- **异常熔断：**工具调用异常时明确回复异常状态，不伪造数据、不沿用历史结果。

3.6 指标与标签语义底座

- **指标体系：**原子指标、派生指标、复合指标三级体系，业务口径一处定义——问“营收”“活跃用户”等业务词，直接命中对应指标与口径，避免同名不同义；
- **标签体系：**标签、群组、画像三层体系，业务语义直接参与资产检索，模糊找数与人群圈选用业务语言即可完成；
- **联动机制：**指标与标签作为资产语义层，与数据目录、质检、多维评估共同支撑问答准确性。

第四章 安全与合规

4.1 权限贯通

智能中心内建于数据中台，与数据权限、用户安全等级同源同底座：

- **四级数据权限：**数据源、表、列、行四级权限直达问答结果——无权限的数据，检索不到、查看不了、分析不得；
- **安全等级差异化：**用户安全等级与数据访问联动，同一字段对不同角色差异化呈现；
- **个人信息强制脱敏：**姓名、手机号、身份证等对全角色强制脱敏，问答与图表中同样遮盖；

- **可见性受控**：资产检索只返回当前登录用户可见的数据集与指标，候选清单本身就是权限过滤后的结果。

4.2 全程审计留痕

问答、查询、图表生成、建议输出全程记录：谁问了什么、看到了什么、何时操作，审计可追溯。使用情况可统计，监管报送材料随取随用。

4.3 工具只读与写入确认

查询类工具全部只读；涉及回写操作（如编目注释更新）的工具，必须经用户明确确认后执行，防止误操作。

第五章 工程特性

5.1 工具化调用，token 消耗少

问答通过系统内置 tools 完成取数与计算，大批量数据由平台引擎处理、不经过大模型。不用让大模型扛着全量数据跑，token 消耗少、响应更快、成本可控。

5.2 组件化开发，兼容多种数据库

开发建议以画布组件形式输出，不依赖特定数据库类型；组件自动调用并行计算引擎处理大批量数据，性能有保障。

5.3 用量统计，运营可度量

平台统计智能体的活跃用户数与分时段 token 消耗，以折线图呈现使用趋势——哪些部门在用、什么时间段消耗多，一目了然，为运营优化与资源规划提供依据。

5.4 内置而非外挂

智能体运行于数据中台智能中心，与数据目录、权限、审计、指标、标签体系同底座，不是独立部署的外挂工具，不存在数据外流通道。

第六章 典型应用场景

应用场景	能力对应
政务数据目录查询	委办局人员对话式检索可共享数据资产，权限过滤保障合规
数据质量月报	对话发起质检方案，六维度评分、脏数据分布与改进建议一键生成
数据资产盘点	九维评估雷达图批量刻画资产健康度，为资产入表与运营评价提供依据
数据开发提速	业务需求经开发建议生成器转为组件化画布命令，开发周期缩短
知识问答与口径查询	指标口径、编目规范等制度知识一问即答，降低跨部门沟通成本

第七章 服务保障

- **技术支持：**7×24 小时技术支持（电话/邮件/远程协助），支持邮箱 support@oceandatum.com
- **常态化运维：**定期巡检、性能优化、安全漏洞修复、版本升级
- **培训服务：**分层培训（系统管理员/技术人员/业务人员），考核发证
- **补丁与版本：**补丁实行发布→灰度验证→通知→正式发放流程，保留回滚方案
- **远程演示：**提供远程演示环境，可远程操作验证问答全流程

第八章 结语

寻数问数智能体让数据问答回归“权限内、底座上”：内置于数据中台，权限贯通、只读工具、防幻觉机制三重保障数据安全；指标与标签语义底座、上架资产口径、九维评估三重机制保障问答准确性。对话即查询、回答即分析、建议即画布命令，为组织的数据用数体验带来质的变化。

第九章 联系我们

上海奥腾计算机科技有限公司

- 官网：www.oceandatum.com
- 信息咨询：info@oceandatum.com
- 技术支持：support@oceandatum.com（项目交付后使用）