

Palantir 本体论、商业化和 Ontology Agent产学研实践

从哲学到行动：本体论驱动的 AI 原生企业构建

柯基数据CEO 吴刚

2026.04.17 南京





30年学研产-1997-2026

永 KG DATA



情报学



人机交互与
智能信息处理



协同办公
GKStar知识库



全球情报
Eikon
知识图谱



知识图谱
大模型
AI知识库

自然语言处理
NLP

知识图谱
KG

大模型
LLM

传统知识库
关键词检索

语义知识库
关联检索及问答

AI知识库
对话式



实践经验：AI赋能数十家科研院所、医院及药械营养企业落地合作

永 KG DATA



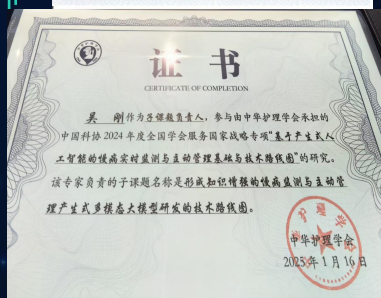
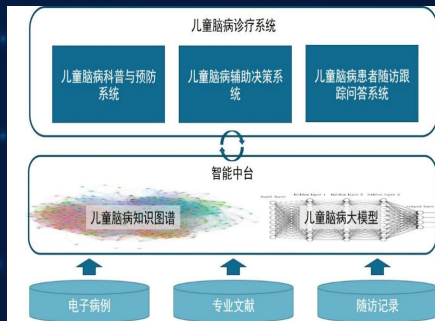


科研成果(围绕本体大模型，国家级、省市级、医院级科研项目支持)

永 KG DATA



资助类别:	面上项目	资助说明:	
备注说明:			
项目编号:		申请代码1:	申请代码2:
项目名称:	知识图谱增强的生成式智能化ICU决策支持系统		
英文名称:	Knowledge Graph-Enhanced Generative Intelligent Decision Support System for ICU Delirium Management		
单位名称:	首都医科大学		
项目批准文号:			
院所:	首都医科大学护理学院		
合作研究单位:	南京柯基数据科技有限公司		



国家级卫健委/科技部/科协&国自然
阜外国家心血管中心&首都医科大学

省市、医院级科研项目
江苏省人民&中医院、儿童医院

中华护理学会重点大模型项目
居家慢病实时监测&主动管理

知识图谱+大模型
评测认证



Agenda

永 KGDATA



**企业级AI知识平台
挑战及进展**



**Palantir本体论
及商业化**



**Ontology Agent
产学研实践**



**落地建议和
AI原生知识平台
未来展望**



AI技术及应用发展

KG DATA



我是 DeepSeek, 很高兴见到你!

我可以帮你写代码、读文件、写作各种创意内容, 请把任务交给我吧~

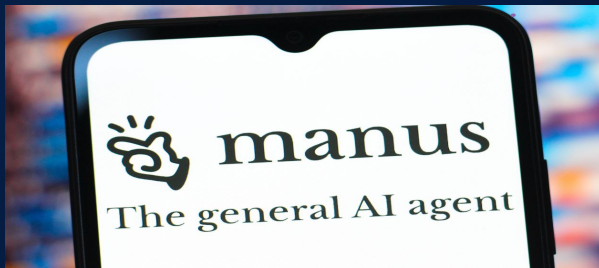
给 DeepSeek 发送消息



深度思考



联网搜索 NEW



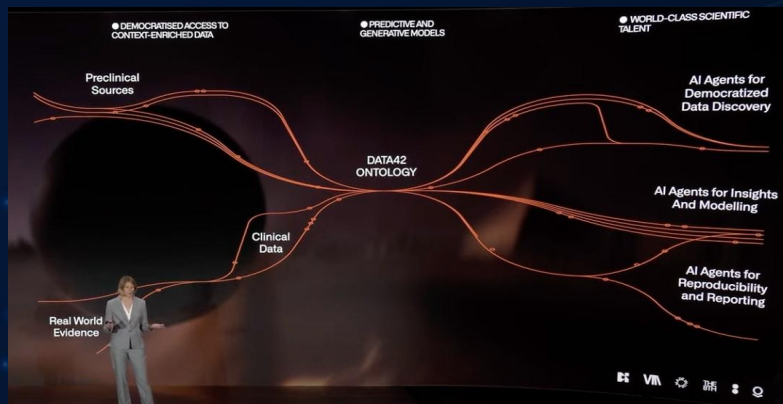
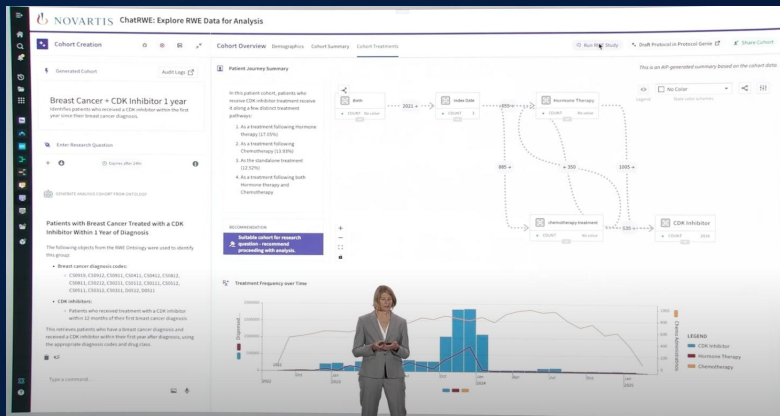
OpenClaw

THE AI THAT ACTUALLY DOES THINGS.

Clears your inbox, sends emails, manages your calendar, checks you in for flights.

All from WhatsApp, Telegram, or any chat app you already use.

Prompt - Agent - Context - Harness



诺华制药基于Palantir Ontology的药研平台



大模型在企业端落地的挑战和破局点

KG DATA

❗ 现状

大模型在 *Demo* 以及 *C* 端市场中表现惊艳，但在企业真实场景中常因幻觉、逻辑断裂、动作失控而受阻。

🎯 核心挑战

如何让 *AI* 真正理解复杂的企业业务逻辑，并在授权的约束空间内精准执行？

🔑 破局点

不再单纯拼模型参数，而是构建企业级的本体论 (*Ontology*) 内核？



成本(构建和运维更新) - 合规循证 - 端到端效率提升(ROI)



医药行业的“深水区”挑战



高复杂度：药物研发涉及海量非结构化数据与跨学科知识。



严合规性：受FDA/EMA/NMPA严格监管，流程必须可审计、可追溯。



知识孤岛：研发、临床、生产、营销各环节数据格式不一，难以互通。



技术痛点：传统AI易产生“幻觉”，且黑盒模型难以满足监管要求的解释性。





多源异构挑战

大型企业数据分散在数千个异构系统中：

- RDBMS、NoSQL、数据湖并存
- 命名规范与数据类型千差万别
- 非结构化文档散落在各处



语义表达鸿沟

传统模型关注“如何存储”而非“业务含义”：

- 外键仅表达关联，无法定义语义
- 缺乏统一的业务词汇表
- 难以弥合部门间的数据定义歧义



维护与演进困境

硬编码的 ETL 管道极其脆弱：

- 对源系统 Schema 变化高度敏感
- “模型上线即过时”的刚性宿命
- 高昂的运维成本与缓慢的响应



本体论：重新定义AI时代的知识管理

永KG DATA

核心定义

本体论是一个形式化的知识框架，由规则、关系和共识词汇组成，使领域知识能够被机器读取和理解。



知识转化

将非结构化、碎片化的领域知识转化为精确的逻辑结构。



统一语言

为多学科、跨国团队建立协作基础，消除专业术语的理解差异。



行业标准

实现不同系统间的互操作性，构建行业级的数据协调机制。



本体 (Ontology)

对特定领域知识的形式化、显式化规范说明。定义概念、属性和关系，是共享的语义模式。

企业作用：

提供统一业务词汇表，消除歧义，是构建知识图谱的“骨架”。



动态本体

随业务环境和实时状态自动演化的本体结构。引入版本控制和LLM辅助建模。

企业作用：

解决传统模型的刚性问题，支持持续集成，应对敏捷业务变化。



知识图谱

以图结构存储的结构化知识库。包含模式层（本体）和数据层（具体实体与关系）。

企业作用：

实现跨源数据关联发现，支持复杂路径查询与推理，增强 AI 可解释性。



LLM&本体知识图谱 - 生成式&判别式

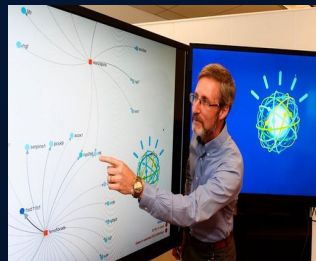
生成式-大模型
参数化知识, 概率生成模型



VS

KG DATA

判别式-本体论、知识图谱
结构化知识、预先生成及决策模型



LLM 的短板

幻觉问题 (Hallucination)

概率模型倾向于生成听起来合理但事实虚假的信息, 缺乏对事实的严格理解。

知识时效性与黑盒化

知识截止于训练数据, 无法获取实时信息, 且推理过程难以追溯和解释。

复杂逻辑推理不足

在涉及多跳、长链条及需要精确领域知识的任务上, 表现往往不如预期。

多源异构数据集成处理不足

对于数据库、Excel等结构化半结构化数据的集成处理

知识图谱的增强

提供确定性事实 (Grounding)

作为“事实锚点”, 通过结构化验证减少幻觉, 确保输出的准确性。

实时知识更新

与企业实时数据源集成, 动态更新知识, 解决训练数据滞后问题。

增强复杂逻辑推理

利用图结构处理复杂关系, 支持多跳推理, 提供可解释的推理路径。

多源异构数据集成

基于本体集成多源异构数据, 形成企业级知识大脑



LLM与本体论：两种不同的知识路径

永KGDATA



LLM：隐性模式学习

核心架构

通过海量数据训练，学习统计概率与语言模式。

关系捕捉

隐性捕捉概念关联，缺乏显式的逻辑规则。

主要风险

容易产生“幻觉”，在受监管领域可能导致严重后果。



本体论：显式逻辑框架

核心架构

基于形式化的规则、关系和共识词汇构建。

关系捕捉

显式定义实体间的逻辑关系，支持确定性推理。

核心优势

提供严谨性、互操作性和透明的可审计推理能力。

结论： 在医疗等高准确性行业，本体论带来的严谨性远比灵活性更具商业价值。



企业级知识库的演化

永KGDATA

Ontology驱动的企业级AI知识平台致力于多源异构数据处理更加自动化、
用户交互更加智能化且避免幻觉，满足企业合规循证要求



RAG驱动的知识库

- 存在忽略上下文语义关联的问题
RAG知识处理主要基于大模型，结合向量数据库构建，可将非结构化文本转化为向量形式存储，依赖于向量切片方法。
- 检索召回不够精准且存在幻觉挑战
支持通过语义检索匹配相关信息并生成回答，交互上智能交互提升，但很难在实际场景中落地使用。主要是RAG技术存在检索召回率低、无法避免大模型幻觉问题，且缺乏意图识别和难以应对复杂查询等问题。

传统知识库

- 知识处理依赖人工，构建成本高效率低。人工上传与分类，若需要处理图片、视频等非文本内容，依赖手动标注关联，效率极低。
- 知识应用还依赖关键词匹配，不够智能。无智能问答能力，用户需手动翻阅文档寻找答案。

Ontology驱动 AI知识库

- 构建本体驱动的多源异构数据建模
Ontology驱动的企业多源异构数据建模，实现数据标准化整合和动态更新。组织和关联企业内外部结构化、半结构化与非结构化数据，打破数据孤岛
- 突破多模态知识构建的瓶颈
实现“文本 + 图表 + 音频 + 视频”的全自动解析与存储，无需人工过多参与。
- 升级更加智能化的交互应用
需要支持跨域跨模态的智能搜索和问答、个性化推荐、写作和决策等应用，覆盖更复杂的业务场景。提升问题的语义理解能力，精准召回的同时，满足循证合规的要求。

成本(构建和运维更新) - 合规循证 - 应用端到端效率提升(ROI)





Agenda

永 KGDATA



**企业级AI知识平台
挑战及进展**



**Palantir本体论
及商业化**



**Ontology Agent
产学研实践**



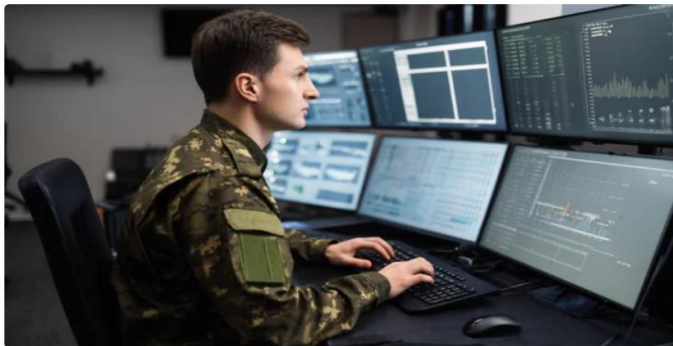
**落地建议和
AI原生知识平台
未来展望**



Palantir - 企业级AI巨头

KGDATA

美国陆军与Palantir签订十年百亿美元合同，构建软件和数据框架

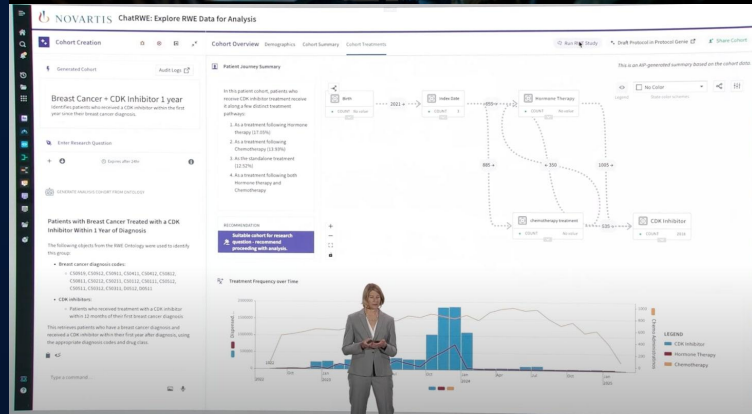


美国陆军通过合同合并取代分散协议，将未来十年软件与数据服务需求纳入单一框架，Palantir获得最高可达100亿美元合作机会。此举标志双方合作进入更高层次，涉及多项关键技术项目。

来源

战略合作深化

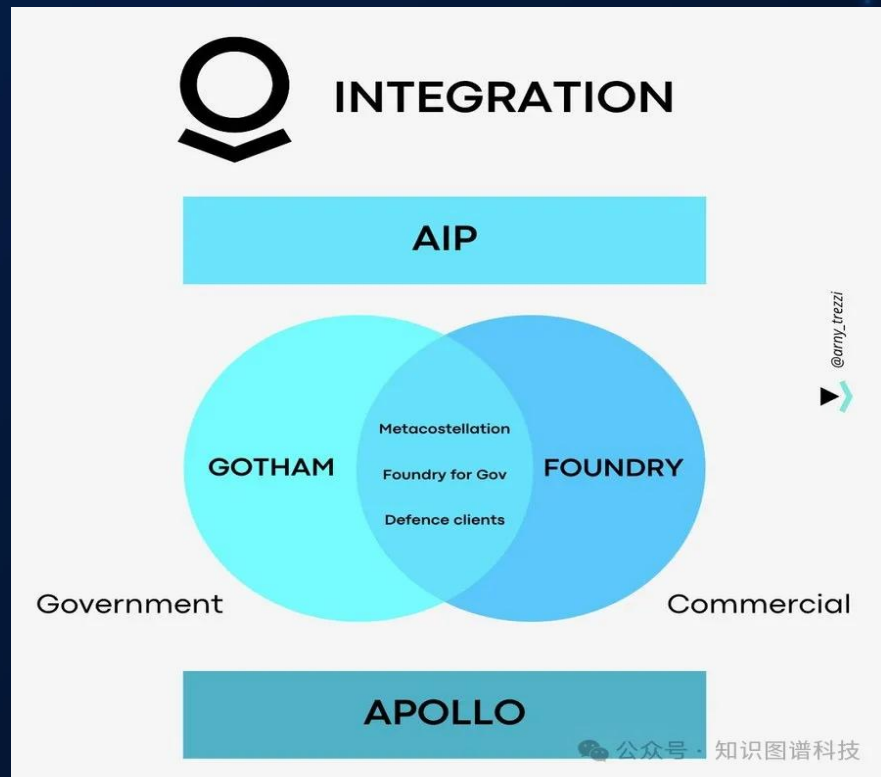
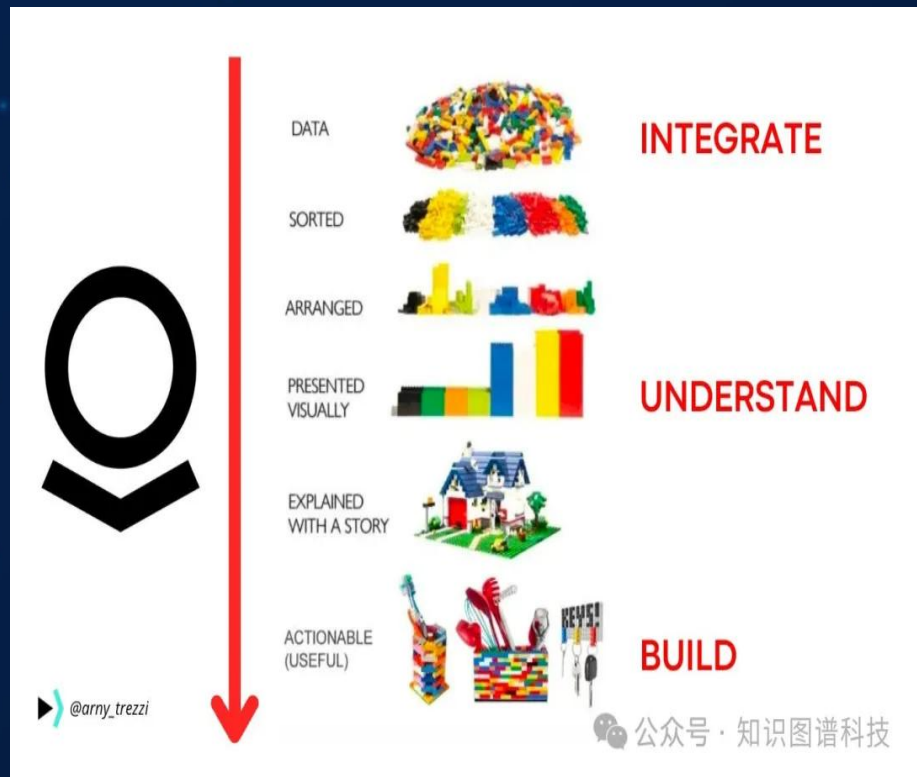
美国陆军宣布将数十份现有合同合并为与Palantir的单一企业协议，通过量化折扣提升采购效率，并允许在未来十年内从该公司购买高达100亿美元服务[1]。该框架协议将覆盖陆军多个软件与数据集成需求，包括研发原型系统、提供数据整合平台及执行特种作战软件更新。分析机构William Blair此前估算，与Anduril合作的NGC2原型项目中约3000万美元价值归属Palantir，未来三年可能持续产生超过1.5亿美元营收[2]。





Palantir产品矩阵

永KGDATA





Palantir本体论“骗局”？

Palantir 的“本体论骗局”

原创 冯若航 老冯云数 2026-02-21 10:47:08 上海 手机阅读 跟

Ontology 就是数据库建模。“本体论”这个词唯一的作用，就是让不懂数据库的人觉得这是个新东西，然后心甘情愿地为旧东西付出一千倍的价格。

一、两种表情

2025 年末，Palantir 市值冲上 4000 亿美元，在两年间翻了二十多倍。这家公司每次路演、每篇白皮书、每个技术博客，都在反复念叨同一个词：**Ontology**（本体论）。

这个词被包装成 Palantir 的核心技术壁垒、护城河与灵魂。投资人听了肃然起敬，五角大楼的将军听了觉得这是信息战的未来，企业高管听了觉得自己如果不买就会被时代淘汰。





在一次面向企业 CXO 的 AI 峰会上，Palantir 的销售副总裁打出一页印着“Ontology”大字的 PPT。台下的高管们微微点头，露出那种“我不太懂但感觉很厉害”的表情。旁边几位被拉来陪会的架构师面面相觑，其中一位低声说了一句：

“他说的是表和存储过程吗？”

同事看了一眼 PPT 上的架构图，沉默了三秒：“.....是的。”

概念	哲学	数据库	面向对象	Palantir
事物的类型	范畴（Category）	表（Table）	类（Class）	Object Type
事物的特征	属性（Property）	列（Column）	字段（Field）	Property
事物之间的关联	关系（Relation）	外键（FK）	关联（Association）	Link
对事物的操作	—	存储过程（SP）	方法（Method）	Action
一个具体事物	个体（Individual）	行（Row）	对象（Object）	Object

Ontology System Data • Logic • Action • Security [x] Language • Engine • Toolchain

	Data	Logic	Action	Security
 Language Models the decisions of the enterprise (i.e., the data, logic, action & security leveraged by Humans + AI).	Semantic Objects • Multimodal • Dynamic Links • Interfaces...	Business Rules • AI-Driven Functions • Conventional ML Models • Optimization (LP) Models...	User-Authored Actions • Agentic Actions • Sandboxed Simulations • Writeback Orchestration...	Role-based controls • Marking-based controls • Purpose-based controls • End-to-end Lineage...
 Engine Powers integration and secure orchestration across storage, compute, and transactional systems, at scale.	Ontology Metadata Service • Object Set Service • Subscription Service • Funnel • Highbury • Time Series • Geospatial • Media • OSW datastores...	Python Functions • TypeScript Functions • Compute Modules • LMS • AIP Logic • Model Adapter Framework • MAZE...	Actions Service • Validations • Side Effects • Scenarios • Automate • Machinery • Webhooks • CDC • Event Listeners...	Roles • Markings • Granular Permission Service • Checkpoints • Approvals • Property Security Groups...
 Toolchain Primitives that enable Developers and Agents to treat the Ontology as a backend.	 OSDK • PSDK • Branching • Marketplace • DevOps			



Palantir本体论“骗局”？

🔑 核心背景

Palantir 凭借“本体论”（*Ontology*）概念在市场上获得巨大成功，成为企业级 AI 的标杆，估值与营收双增。

⚠️ 技术质疑

有观点认为这只是传统数据库建模的“旧瓶装新酒”，旨在通过复杂的哲学包装获取高额商业利润。

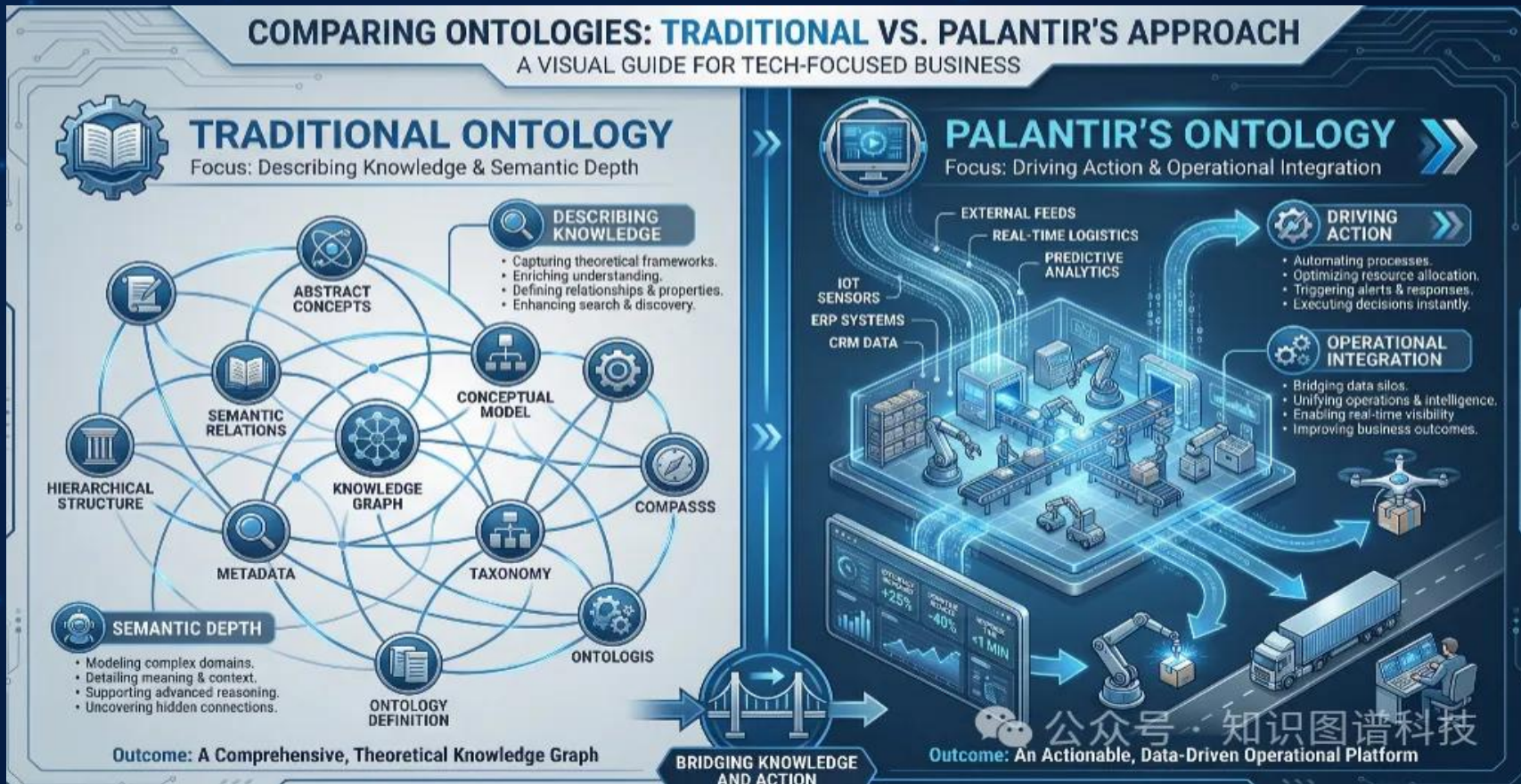
❓ 核心议题

在 LLM 狂飙突进的今天，企业是否真的需要这套昂贵的“本体论”？它在 AI 落地中究竟扮演什么角色？

Palantir 本体概念	传统数据库概念	业务含义
<i>Object Type</i> (对象类型)	Table (数据表)	定义实体，如“客户”、“订单”
<i>Property</i> (属性)	Column (字段/列)	描述实体的特征，如“姓名”、“金额”
<i>Link</i> (关联)	Foreign Key (外键)	定义实体间的关系，如“谁下了单”
<i>Action</i> (操作)	Stored Procedure (存储过程)	定义业务逻辑，如“创建订单”、“取消发货”



Palantir本体论和传统本体、知识图谱差异





Palantir本体论和传统本体、知识图谱差异

特性	传统知识图谱 (RDF/OWL)	传统知识图谱 (LPG/Neo4j)	Palantir 本体论 (Ontology)
基本单元	三元组 (Subject-Predicate-Object)	节点、边、属性	对象、属性、链接、动作
主要用途	语义互操作、逻辑推理	复杂关系挖掘、图遍历	业务运营、决策执行、AI 落地
查询语言	SPARQL	Cypher / Gremlin	OSDK / 自然语言 (AIP)
数据一致性	最终一致性	事务一致性 (ACID)	强一致性 (Phonograph 事务)
回写能力	极弱 (通常需外部集成)	弱 (需自定义代码)	强 (内置 Action Types & Webhooks)
安全模型	粗粒度 (图/类级别)	中粒度 (节点/边级别)	极细粒度 (细胞级 + 动态策略)
与原始数据关系	往往是脱离源系统的副本	往往是脱离源系统的副本	实时映射, 保留完整血缘 (Lineage)



数据治理层

国外大型企业在 ERP、CRM 等核心系统上的**标准化程度较高**，为本体论的对象化映射提供了天然的结构基础



技术产品层

本体论驱动的 AI 原生架构，实现业务逻辑的数字资产化。同时产品支持FDE的快速交付和迭代，其次通过 bootcamp模式验证业务价值



商务政治层

深耕政商关系，将商业利益与国家安全叙事深度绑定。融资能力强，从政府上亿美金大单到企业级大单



组织交付层

以 FDE 为核心的特种作战模式，跨越产品与业务的鸿沟。

核心启示

“复制一款产品容易，但复制一套‘数据+技术+商务政治+组织’四位一体的体系极难。这才是 Palantir 真正的护城河。”



Palantir成功因素 - 美国政企数据治理基础：本体论的“肥沃土壤”

永KGDATA

≡ 标准化程度高

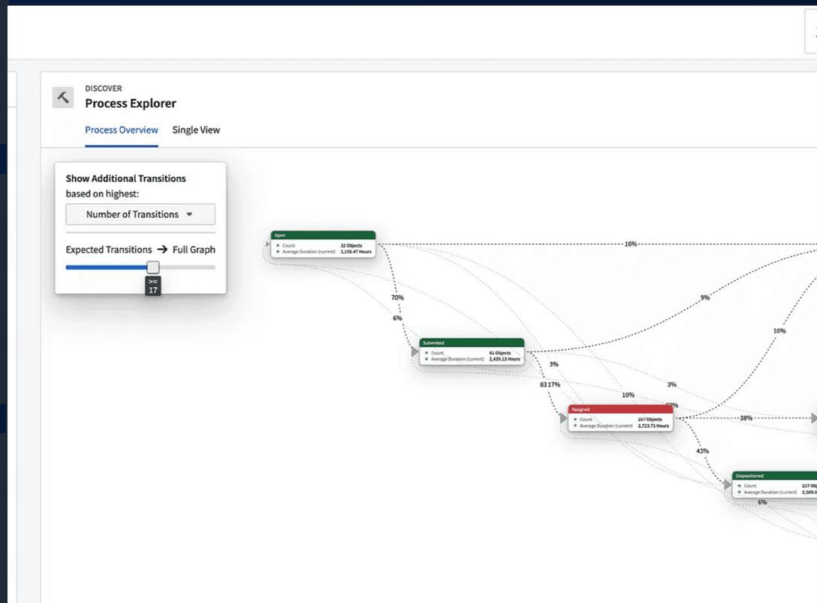
国外大型企业在 ERP、CRM 等核心系统上的**标准化程度较高**，为本体论的对象化映射提供了天然的结构基础。

≡ MDM 成熟度

相对成熟的主数据管理 (MDM) 意识，使得“将现实世界映射为数字对象”不再是空中楼阁，而是可落地的工程实践。

合规性红利

严格的数据隐私与合规要求，反而凸显了 Palantir 在**权限控制**与**审计追踪**上的技术优势，降低了信任门槛。





Palantir成功因素 - FDE 模式：深入“泥潭”的超级个体

永KGDATA

👤 定义与职责

Forward Deployed Engineer 是驻扎在客户一线的“特种兵”工程师，直接在混乱的原始数据中寻找业务逻辑。

</> 双重能力

要求顶尖的**工程实现能力**（写代码）与卓越的**业务洞察力**（懂政治、懂流程），能够与高层直接对话。

⚡ 去中心化决策

FDE 拥有极大的现场决策权，能够根据业务真实痛点快速调整产品逻辑，而非等待总部的漫长审批。



公众号 · AI重塑未来



Palantir成功因素 - Bootcamp：所见即所得的交付革命

KG DATA

⚡ 现场作战

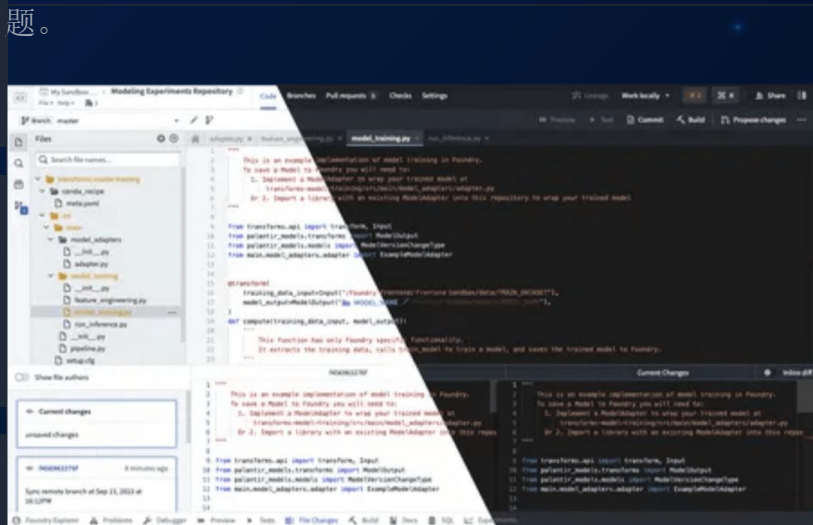
1-5 天的训练营，直接使用客户的**真实数据**解决真实业务问题。

📈 效率飞跃

2025 年美国商业收入增长 **121%**，核心驱动力正是这种“快速证明价值”的模式。

👁️ 心理攻势

当客户看到自己的数据在几小时内被理顺并产生洞察时，怀疑会迅速消融。





**企业级AI知识平台
挑战及进展**



**Palantir本体论
及商业化**



**Ontology Agent
产学研实践**



**落地建议和
AI原生知识平台
未来展望**



本体推理AI大脑驱动的企业级新一代AI知识平台

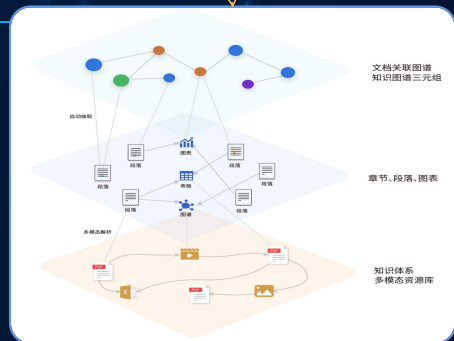
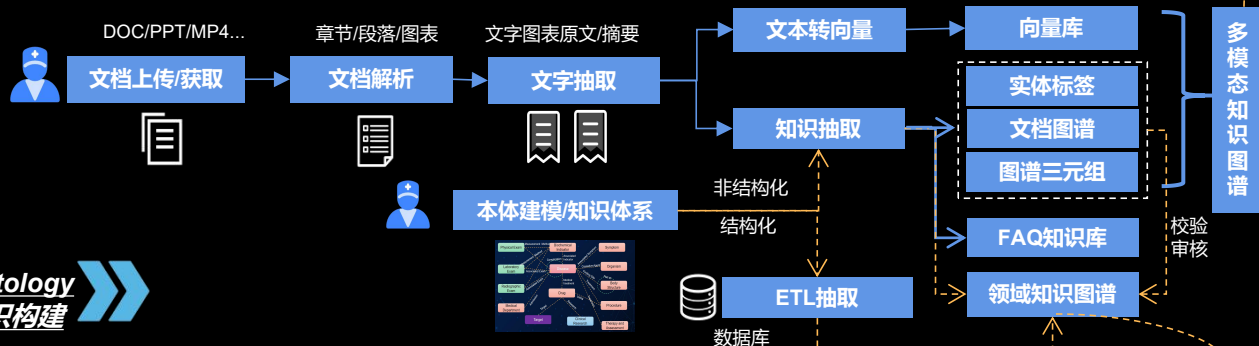
永KG DATA





基于Ontology的AI知识库和Agent应用

KGDATA



语义图谱 (自动化构建更新, 辅助大模型精准召回)

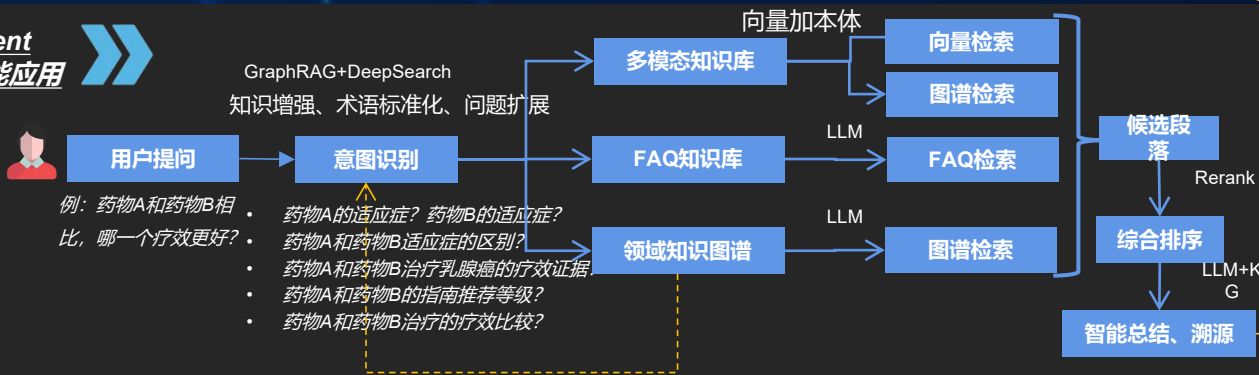


高质量业务图谱 (人工校验及维护, 作为答案来源)



推理图谱 (基于AI知识库和Agent实时推理决策)

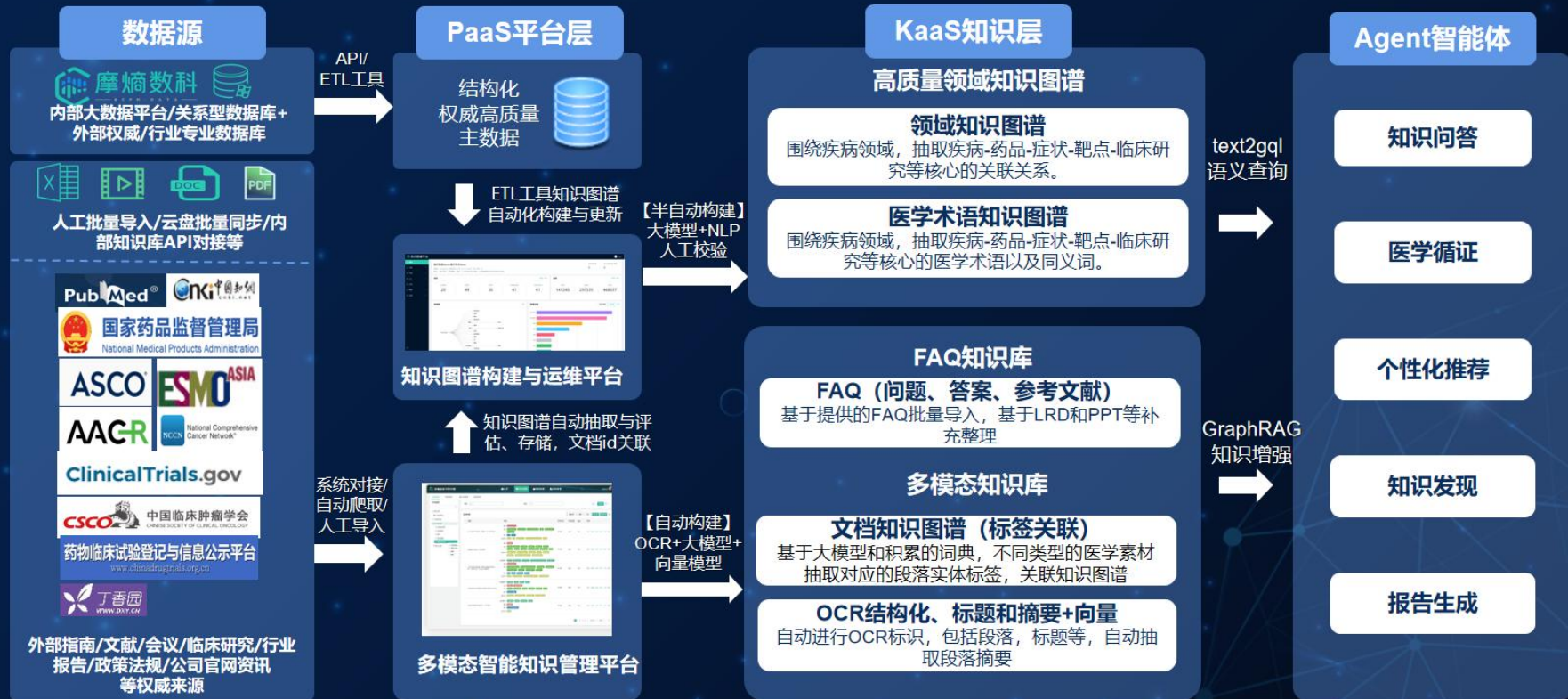
Agent 智能应用





基于Ontology的AI知识库和Agent应用

永KGDATA

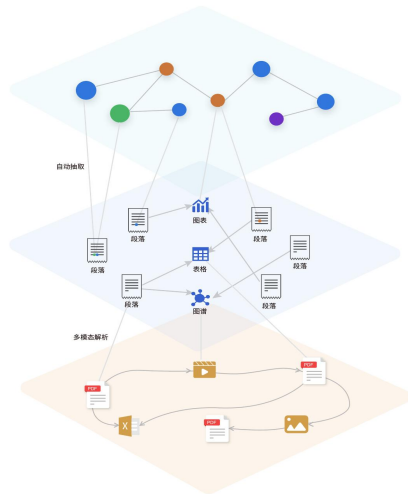
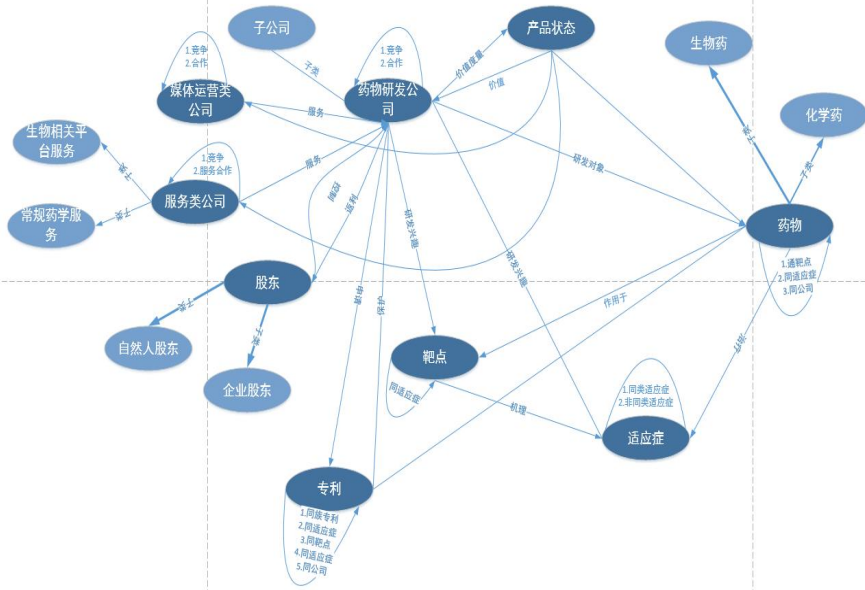




本体和语义图谱、高质量图谱

说明

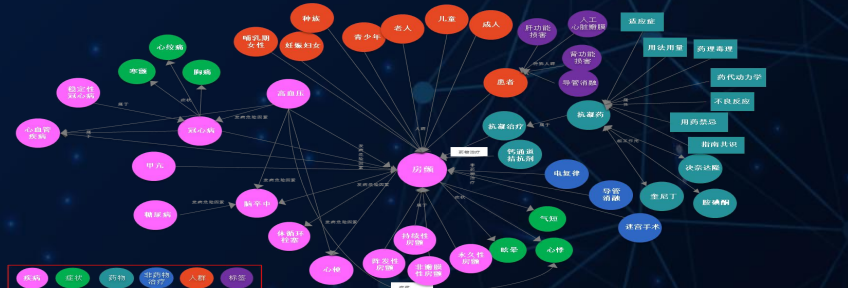
- 概念
- 关系
- 属性(单值)
- 属性(多值)
- 实体



知识点和标签级别的知识关联

段落、图表的结构化关联

文献/研究论文/学术会议



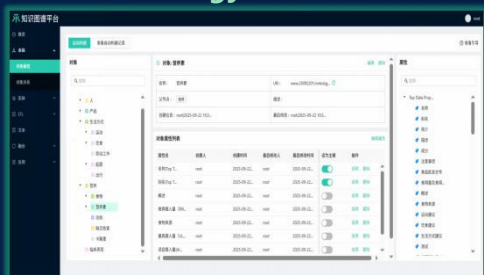


本体知识图谱平台多源异构数据集成，高质量知识图谱构建

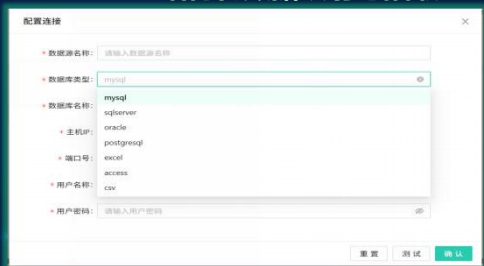
KG DATA

平台功能完整性：包含了从本体构建，结构化、非结构化数据抽取、融合，关联可视化分析推理，实体管理和运维等全流程构建和运维能力，支撑企业构建动态可运维知识图谱，开放API供上层应用查询和推理。

Ontology本体构建



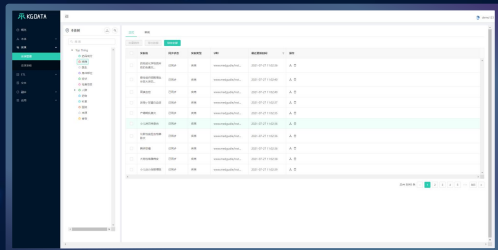
ETL结构化数据映射与抽取



大模型预标注配置



实体管理可运维，包括属性和关系等



冲突检测与图谱融合



关联分析可视化，提供API接口支持查询和推理





AI知识库建设: 内部业务系统数据采集对接-ETL工具

永 KG DATA

- 覆盖常用的关系数据库, 包括mysql、oracle、sqlserver、postgerSQL、csv、excel等作为结构化数据源;
- 数据源、数据集、映射、任务配置操作方便快捷, 达到业务人员可方便使用的标准。



数据源

提供数据来源



数据集

自定义数据集



映射配置

对象关系映射



任务抽取

永 知识图谱平台

概览

本体

实体

ETL

数据源管理

数据集管理

映射配置

任务抽取

一键导入

文本

融合

应用

< 返回 添加数据集

*数据集名称: 不为空、不允许特殊字符且长度不超过...

*数据集类型: ☒ 对象数据集 ☐ 关系数据集

数据集最多添加2个数据源里面的10张数据表

*选择数据源:

AZ产品一览表

将需要的表格拖动到右侧

请输入表名

excel_4587_TA

excel_4587_AZ产品一览表

excel_4587_成分

excel_4587_疾病

excel_4587_药品成分关系

excel_4587_疾病关系

每个数据集设置一个唯一主表

每张表至少设置一个主键

设置增量策略 帮助

excel_4587_AZ产品一览表

AZ产品一览表

auto_id

TA

商品名称

英文名称

通用名称

成分

注意事项

贮存

规格

包装

剂型

适应症

介绍

excel_4587_成分

AZ产品一览表

auto_id

成分

选择表(打星号的表已设置增量策略)

AZ产品一览表

excel_4587_AZ产品一览表

excel_4587_成分

excel_4587_药品成分关系

excel_4587_药品成分关...

AZ产品一览表

auto_id

商品名称

成分

设置增量策略

数据集配置



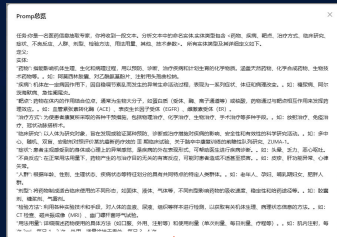
AI知识库建设：非结构化数据自动标签和语义图谱

永KGDATA

知识体系本体搭建 (内置)

实体抽取自动生成 (Prompt)

知识图谱抽取与评估多智能体 (AutoQG)



文档批量上传/云盘导入

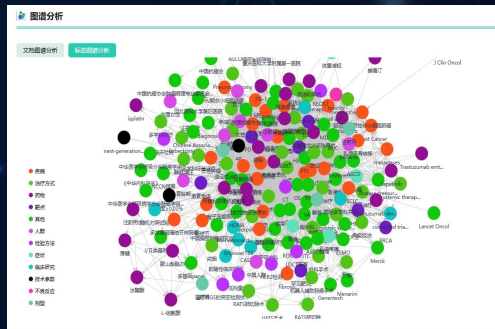
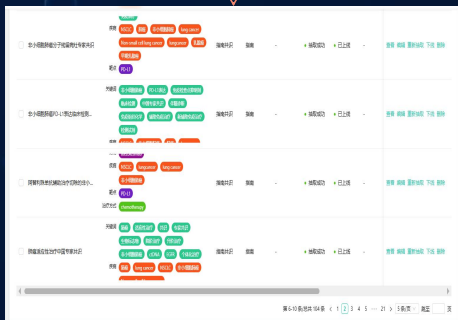
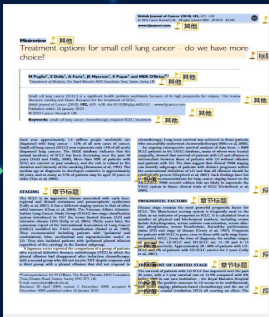
- 产品说明书
- 中英文文献
- 中英指南
- 其他各类材料

文档自动解构

- 文献标题, 摘要, 关键词
- 段落, 表格, 图片, 参考文献及其他

基于多智能体能力
实体、三元组自动抽取与
评估

多模态、多领域、多文档
知识图谱融合与关联



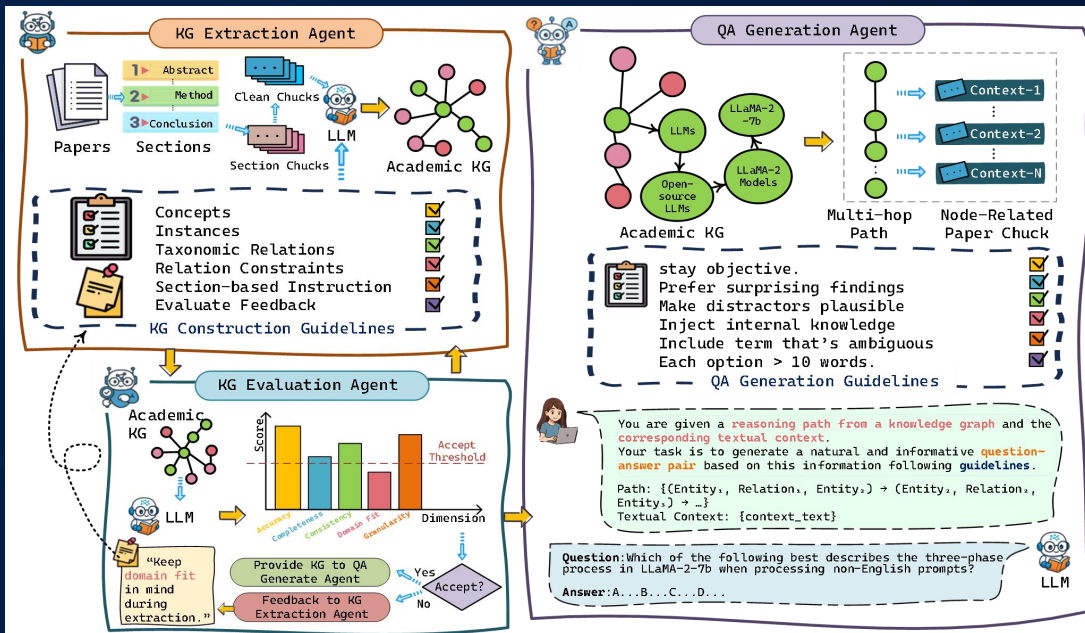


AI知识库建设：高质量知识图谱自动构建

KGDATA

AutoQG: 一个全自动、多智能体框架，实现基于证据的知识图谱抽取、评估与问答生成。

该框架实现了“构建—评估—生成”的闭环，多智能体分工明确，既保证了知识图谱的科学性和完整性，又能够高效地产出符合学术场景需求的问答对。——来自东南大学人工智能算法研究团队



AutoQG整体框架分为三部分：知识图谱抽取、知识图谱评估与问答对生成。

1、知识图谱抽取 (KG Extraction Agent)

从论文中自动抽取核心实体、概念和关系，构建结构化知识图谱，为评估和问答生成提供统一的知识基础。

2、知识图谱评估 (KG Evaluation Agent)

检查图谱的完整性、一致性与准确性，并将结果反馈给抽取环节，实现双向优化，持续提升图谱质量。

3、问答对生成 (QA Generation Agent)

基于高质量图谱生成高难度单选/多选题，利用多跳关系，确保考察深度理解与推理能力。



AI知识库建设：多模态知识融合能力

永KGDATA



指南、文献等专业文档
自动化构建图谱

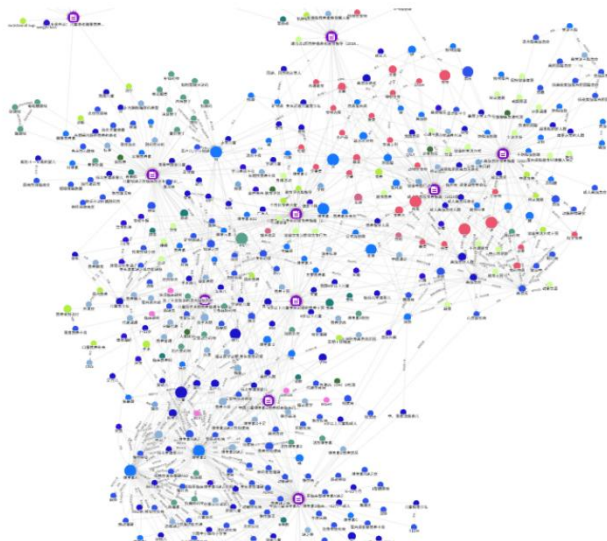


视频、音频等多模态数
据自动化构建图谱



知识图谱融合关联，并索引
到原段落内容

图谱分析





AI知识库建设：多模态（图片）处理能力

永KGDATA

兼容多模态多格式的自动化打标签能力，包括PDF、word、PPT、视频、图片、音频等。基于文档中的图片或者图片文件可自动化处理，标签，关联形成图谱。

自动缩放 25 / 362 > 1 页 / 图

联结成的长链称为肽。相应氨基酸的数目为 由于羧基和氨基都在 α -碳原子上，故

98

第五卷 蛋白质和氨基酸

图 5-1-2 氨基酸的肽键结构

称为 α -氨基酸。除缬氨酸基团只有 1 个氢原子外，其余 α -氨基酸的 α -碳原子都是不对称碳原子。因此氨基酸存在 L 型和 D 型两种异构体。人体蛋白质中氨基酸都是

量不足，须额外补充的称为条件性必需氨基酸。如精氨酸在创伤等情况合成不足，额外补充以增强免疫功能。

整个肽链从 N-末端到 C-末端依次排列，显示了氨基酸的排列和顺序。

标签信息 基本信息 图谱信息 AI 总结

缩略图标题 肽键

缩略图名称 图 5-1-2 氨基酸的肽键结构

缩略图正文

蛋白质是由一个多肽链或多肽链组成。包含多个氨基酸单元通过肽键连接。每个氨基酸单元由一个 α -碳原子 (C) 组成。该碳原子连接着氨基 (NH₂)、羧基 (COOH)、氢原子 (H) 和侧链 (R)。侧链 (R) 决定了氨基酸的性质。氨基酸的侧链 (R) 决定了氨基酸的性质。氨基酸的侧链 (R) 决定了氨基酸的性质。

1. "H" 末端: 氨基酸的侧链 (R) 决定了氨基酸的性质。氨基酸的侧链 (R) 决定了氨基酸的性质。
2. "肽键": 氨基酸的侧链 (R) 决定了氨基酸的性质。氨基酸的侧链 (R) 决定了氨基酸的性质。
3. "H" 末端: 氨基酸的侧链 (R) 决定了氨基酸的性质。氨基酸的侧链 (R) 决定了氨基酸的性质。
4. "H" 末端: 氨基酸的侧链 (R) 决定了氨基酸的性质。氨基酸的侧链 (R) 决定了氨基酸的性质。

氨基酸的侧链 (R) 决定了氨基酸的性质。氨基酸的侧链 (R) 决定了氨基酸的性质。

图片内容理解

轻松抗酶-疗效放心: 康新博® 浓度稳定疗效保障。单药治疗侵袭性曲霉病的一线优选药物 0922

其他 6

自动缩放 1 / 4 >



相关信息 图谱信息 AI 总结



图片标签提取

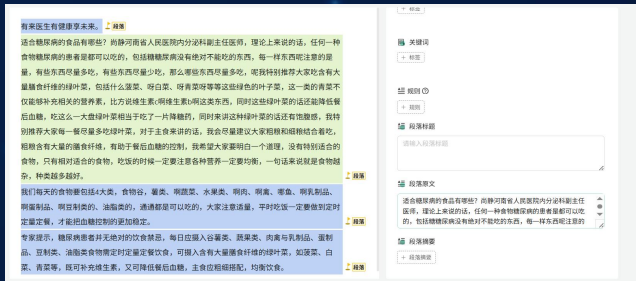
图片关联图谱



AI知识库建设：多模态（视频/音频文件）处理

永 KG DATA

针对上传的视频，将语音转成文字，并分段处理；视频转文字和标签抽取结果支持修改。



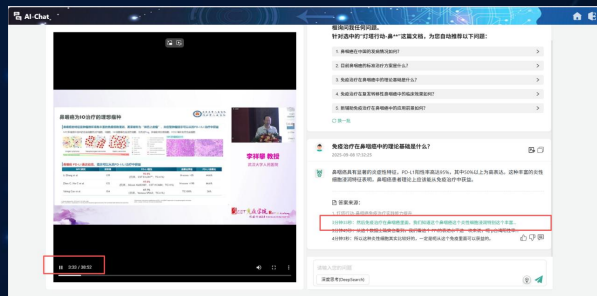
录音文件或者视频文件
输入

语音处理与分段（语音转文字API）

AI总结和标签提取



ID关联到Knowledge Base



问答、检索、个性化推荐



AI知识库建设：结构化、非结构化知识融合



所有ETL/文本抽取的结果可设置进入到待审核列表，在审核前，可以使用查重功能进行融合处理；支持融合ETL抽取、文本抽取、人工创建的数据来源根据融合相似度算法进行本体和实体的相似度计算，可以设置**自动融合或者人工干预融合**。

融合配置项

✕

本体

* 本体相似度②:

80

%

* 是否人工干预②:

是

实体

* 实体相似度②:

90

%

* 是否人工干预②:

是

取消

确定

永 KGDATA

demo123

首页

概述

人 审核

告 汇报

旧 ETL

工 文本

数据配置

数据治理

告 治理

数据管理 / 已部署任务

来源名称	来源设备	目标设备	任务类型	Y	成功率	Y	任务结果	Y
设备	传感器A	传感器B	人工融合	92.15%	成功			
设备	传感器C	服务器D	人工融合	95.52%	成功			
设备	空中通信基站设备	空中通信基站设备	人工融合	90.28%	成功			
设备	早期识别能力设备	早期识别能力设备升级	人工融合	90.37%	成功			
设备	天气数据源	天气数据源	人工融合	91.31%	成功			
设备	数据标准化	数据标准化	人工融合	91.67%	成功			
设备	血加模型	血加模型	人工融合	92.51%	成功			
设备	中量数据	中量数据	人工融合	92.55%	成功			
设备	网络层	网络层	人工融合	92.38%	成功			
设备	月报不报	月报不报	人工融合	90.07%	成功			



Agent应用：智能对话，精准召回与循证智能体

永KGDATA

RAG落地挑战

存在忽略上下文语义关联的问题

检索召回不够精准且存在幻觉挑战

缺乏意图识别和难以应对复杂查询

企业级多模态知识管理需求

突破多模态知识构建的瓶颈

构建成熟的企业级知识体系

升级更加智能化的交互应用

双轮驱动解决方案

用户提问

诊断出来蛋白质不足，有什么建议？

- KG (术语标准化、图谱推理)
- DeepSearch (语义理解与意图识别)

- 蛋白质缺乏的影响或者危害；
- 蛋白质缺乏有哪些生活方式建议，包括饮食、运动与各类行为？
- 蛋白质缺乏，有哪些产品可以推荐？
- 蛋白质包含动物蛋白和植物蛋白(如大豆分离蛋白等)，如何补充？

企业级知识库

500个营养素知识点、500个营养产品知识点

FAQ知识库

FAQ Search

KG Search

GraphRAG Search

高质量领域知识图谱

GraphRAG 多模态知识库



LLM

LLM

KG Search

Vector Search

共10个相关段落

蛋白质的定义和缺乏危害知识点

补充蛋白质的产品介绍与使用方法

蛋白质缺乏的危害

蛋白质缺乏的生活方式建议

蛋白质缺乏的营养补充和产品推荐

大豆分离蛋白的功效信息

蛋白质补充的饮食建议

蛋白质的定义

蛋白质缺乏的危害

大豆分离蛋白的最新研究

Rerank+LLM 排序Prompt

[从10个相关段落中，找到能够回答问题的段落并排序...]

段落原文、溯源

共召回6个相关段落

FAQ内容：

蛋白质的定义和缺乏危害知识点

图谱内容：

蛋白质缺乏的生活方式
蛋白质缺乏的营养补充和产品推荐

知识库内容：

蛋白质补充的饮食建议
大豆分离蛋白的功效信息
大豆分离蛋白的最新研究

段落召回完整

满足合规要求

LLM总结Prompt

[从6个相关段落中，进行总结回复，增加溯源验证...]

Agent智能应用

总结回复

- 蛋白质的定义和缺乏危害知识点；
- 蛋白质缺乏的生活方式建议，包括营养补充、饮食建议；
- 大豆分离蛋白的功效信息和最新研究

产品推荐

某蛋白产品的产品介绍与使用方法

个性化健康建议

素食主义者需要合理搭配不同种类的植物蛋白，提升吸收利用率。

报告生成

营养健康报告



Agent应用：智能对话，精准召回与循证智能体

永 KG DATA



DeepSearch

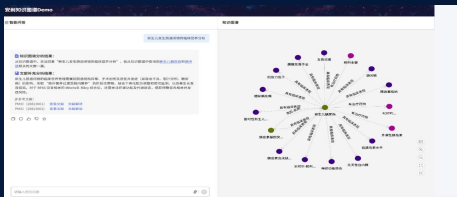
循证证据
(例如FAQ)

知识图谱
或者数据库

GraphRAG
原文增强检索

GraphRAG
AI总结

新问题
反馈机制





Agent应用：跨文档多源异构数据写作

永KGDATA



外部MCP链接：全球权威数据源接入



权威医学数据库：PubMed, ClinicalTrials.gov, Cochrane Library



药品监管机构：FDA, EMA, NMPA 官方数据库



专业知识库：各国药典、药品说明书、临床诊疗指南



市场信息源：医药市场销售数据、竞争格局分析报告



企业内部数据：历史立项报告、内部研发数据、知识库



多源信息处理：从数据到知识的转化



自动检索

基于任务关键词，全网自动检索相关信息



去重与清洗

智能识别并移除重复、错误或低质量数据



抽取与结构化

提取关键信息并统一格式，转化为结构化数据



证据整合

交叉验证多源信息，构建统一的知识图谱



Agent应用: 推理图谱&决策 - 慢病实时监测、主动管理

永KGDATA





Agenda

永KGDATA



**企业级AI知识平台
挑战及进展**



**Palantir本体论
及商业化**



**Ontology Agent
产学研实践**



**落地建议和
AI原生知识平台
未来展望**



中国市场落地挑战



数据土壤太碎

系统烟囱林立，主数据混乱，权限链路复杂。缺乏构建清晰 *Ontology* 的底层数据治理基础。



采购机制错位

大客户招标压价与重交付模式，难以支撑高 *ACV* 的工具化平台成长，导致项目陷入价格战。



权力设计缺失

本体论落地涉及跨部门流程重组。没有“一把手”的强力背书，再好的平台也难以真正动刀。

“复制产品不行，复制体系才有戏。”



企业决策指南：何时投资本体论？

✓ 采用信号

受监管行业

涉及医疗医药、法律、国防等对准确性有严格要求的领域。

互操作需求

需要与多个外部组织或内部异构系统进行深度集成。

知识复杂度高

领域术语众多、逻辑关系复杂且容易产生歧义。

⏸ 暂不采用信号

任务简单狭窄

基础的NLP、图像识别或已有成熟硬编码规则的场景。

优先快速上市

处于早期探索阶段，需要极高的迭代速度和灵活性。

黑箱接受度高

用户对决策过程的可解释性没有强制性要求。

📖 最优混合策略

GraphRAG 交互技术

利用图检索增强生成，让 LLM 与本体论进行动态交互。

多层本体和图谱，平衡性价比

通过语义本体图谱和高质量本体图谱和推理图谱

从小规模开始

先构建核心概念本体，根据业务反馈逐步扩展规模。

本体论在AI时代的新使命

本体论不是过时的技术遗产，而是连接系统、团队和标准的关键基础设施。未来的竞争优势属于那些能够最有效地组织、管理和利用领域知识的企业。

拥抱融合架构

利用 *GraphRAG* 技术让 LLM 与本体论交互，平衡灵活性与严谨性。

降低应用门槛

让非技术专家也能参与领域知识的数字化建模。

构建可信AI

将本体论作为“单一真实来源”，打造透明、可审计的智能体系统。

共同探索 AI 与知识管理的未来



打造世界领先认知智能引擎 加速人类知识工作自动化



吴刚

13901264290



✉ wg@kgtdata.com

🌐 www.kgtdata.com