

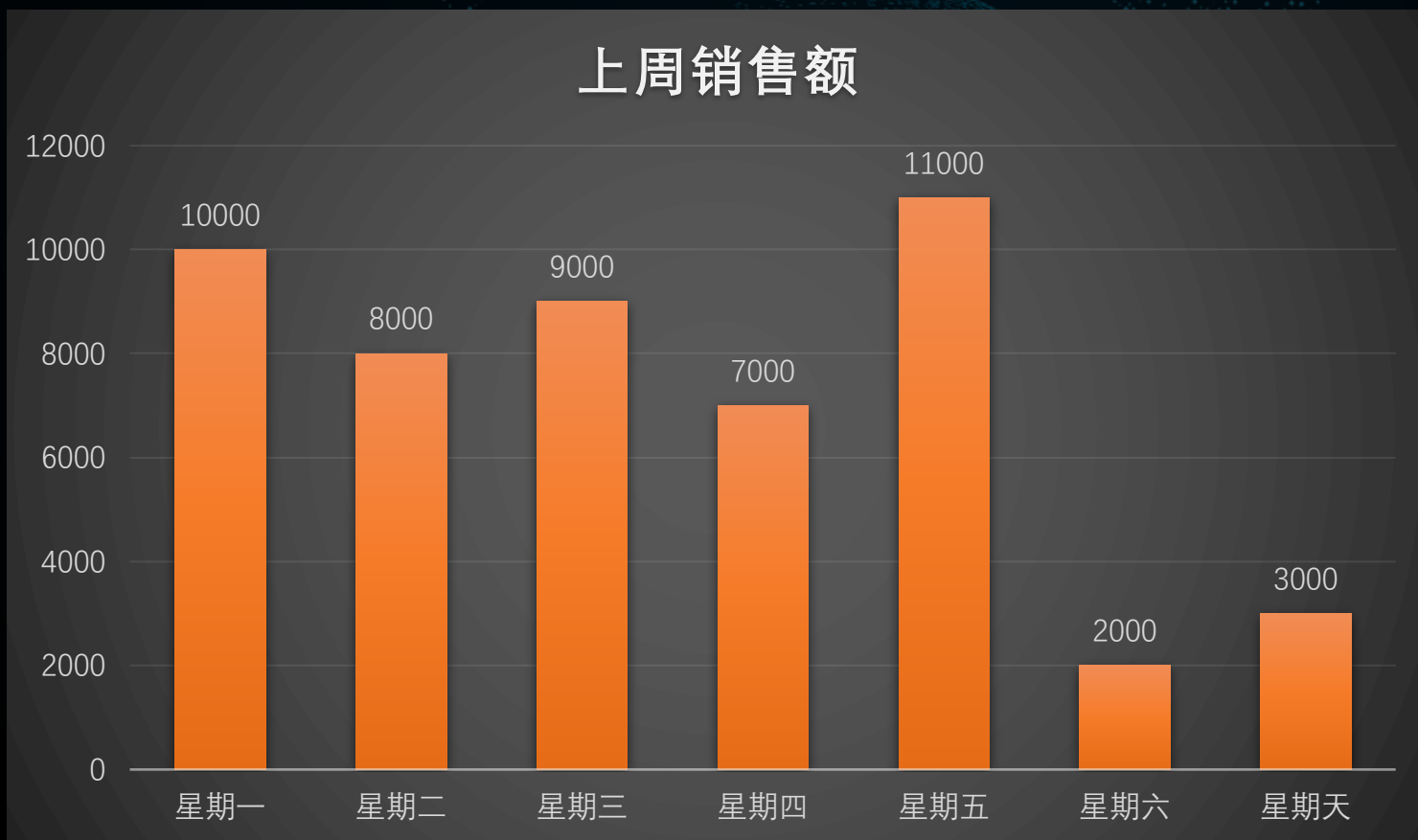
AI如何实现从个体赋能到全局智能

杨威

明略数据 技术合伙人&产品负责人

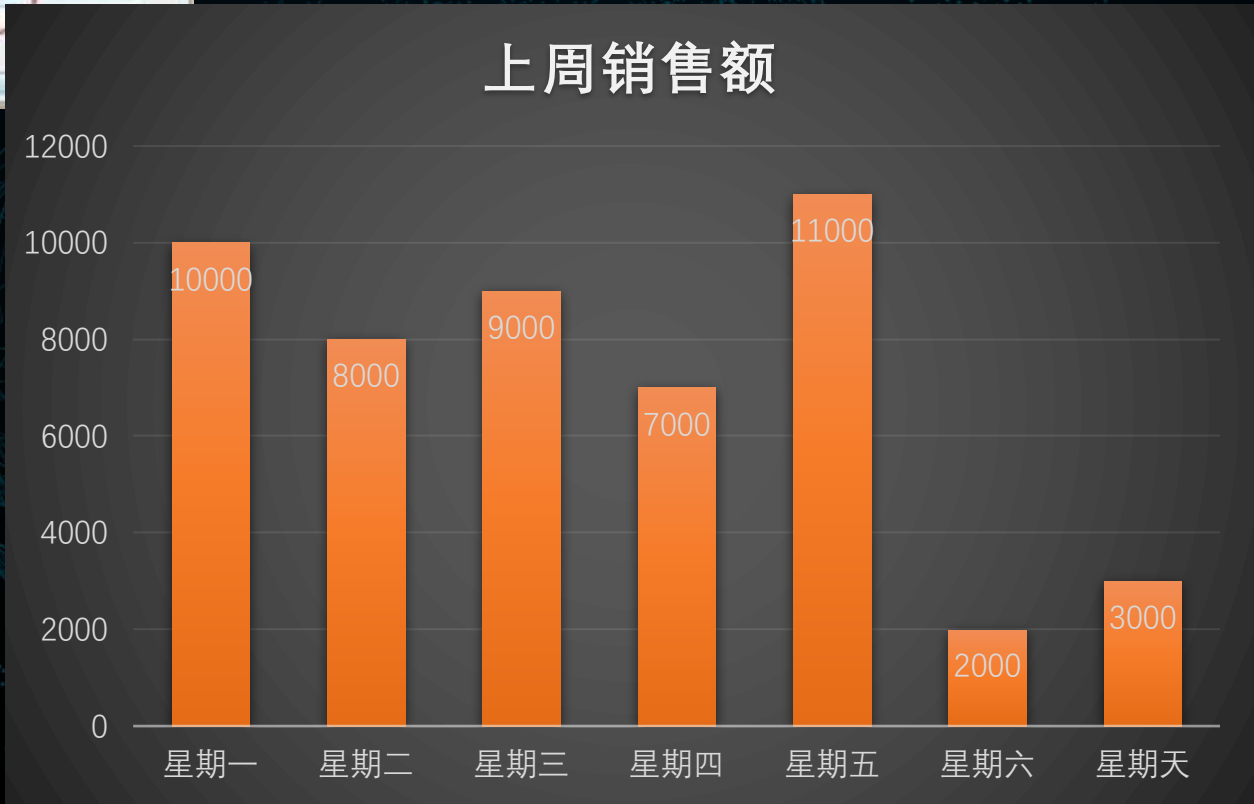
假如你是一个餐饮店的老板，请了一个店长帮你运营，
你想要了解上个周的营收报表时…

店长用Excel画了一张图：



店长用餐饮管理系统里的BI功能做了一张图

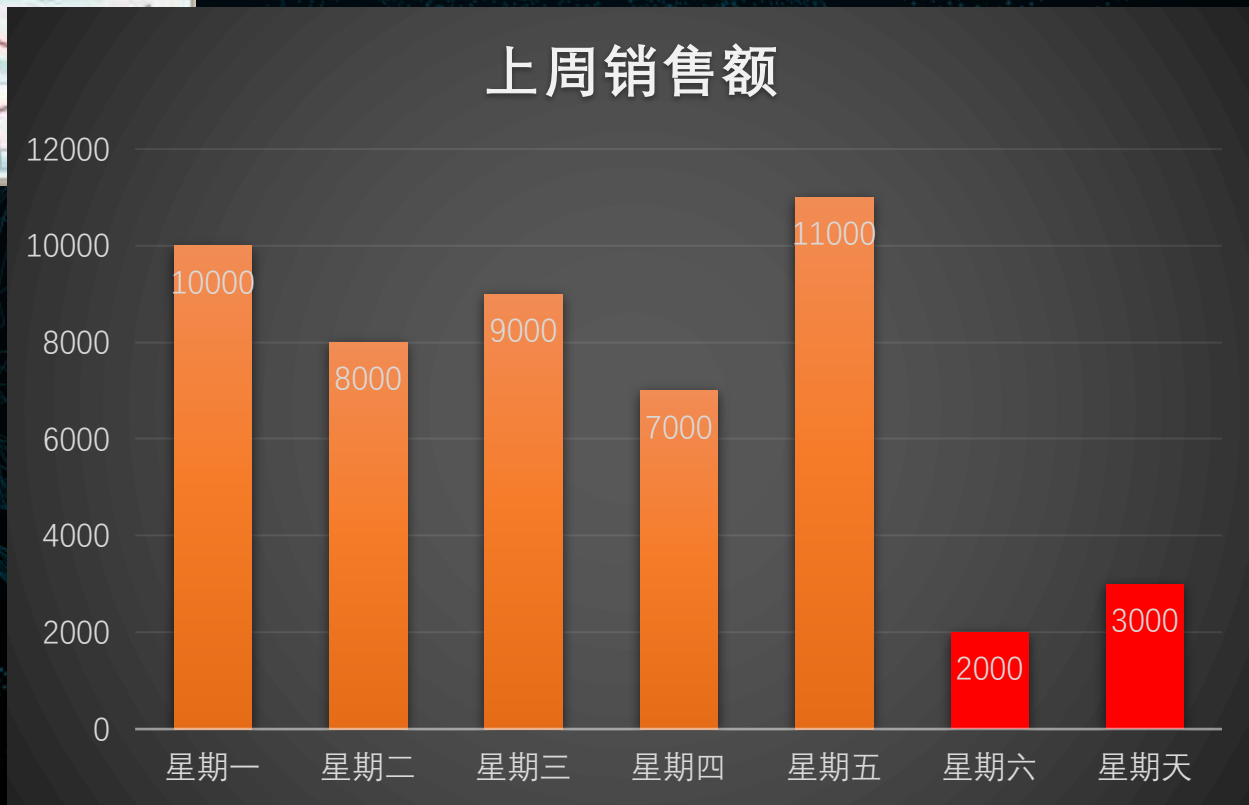
Date	Dept	Item	Price	Qty	Total
2017-10-27	01	01	79	12	948
2017-10-27	01	02	89	10	890
2017-10-27	01	03	99	10	990
2017-10-27	01	04	109	10	1090
2017-10-27	01	05	119	10	1190
2017-10-27	01	06	129	10	1290
2017-10-27	01	07	139	10	1390
2017-10-27	01	08	149	10	1490
2017-10-27	01	09	159	10	1590
2017-10-27	01	10	169	10	1690
2017-10-27	01	11	179	10	1790
2017-10-27	01	12	189	10	1890

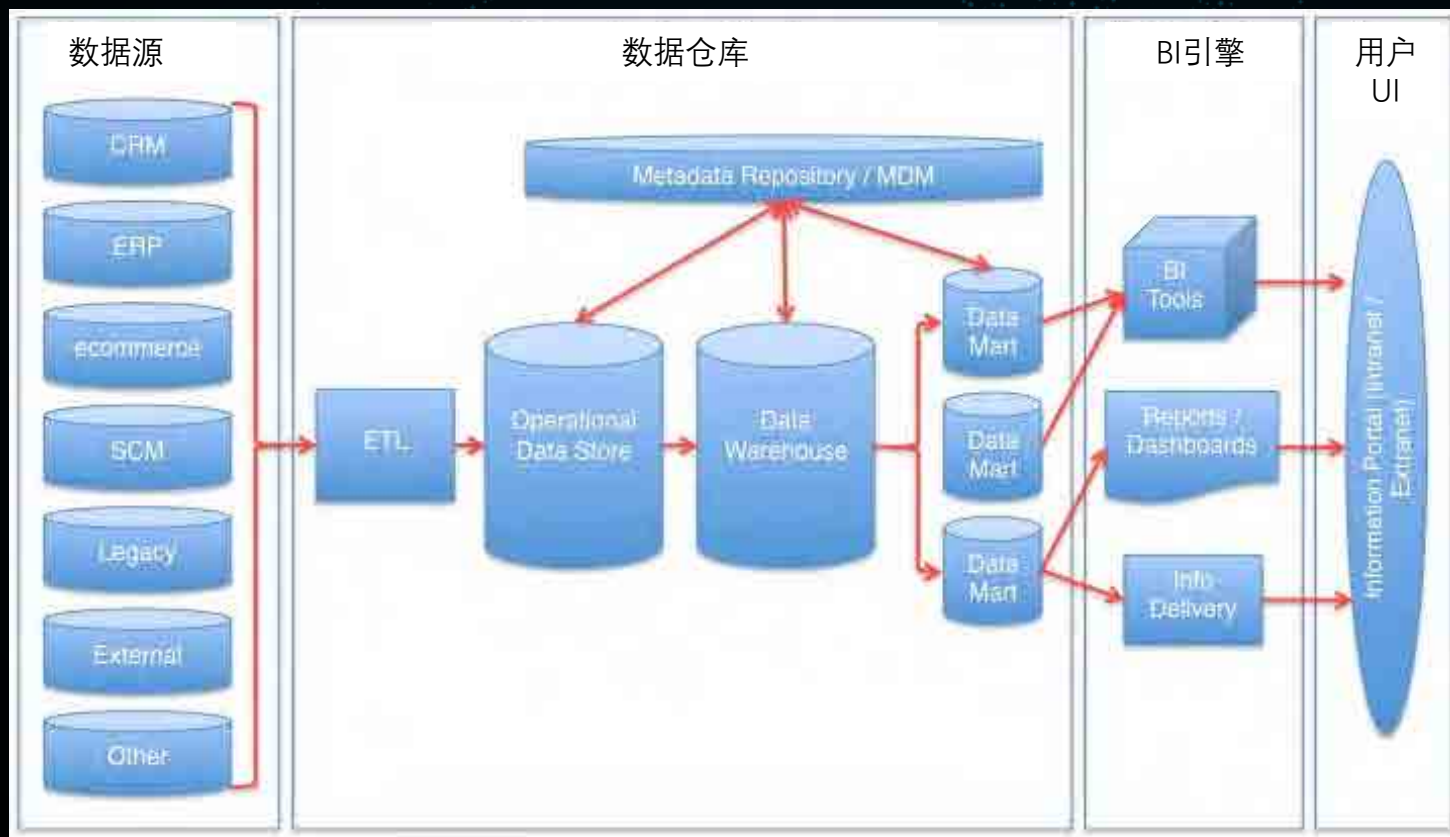


店长用餐饮管理系统里的BI功能做了一张图，且BI工具给出了一些提示

The screenshot shows a BI tool interface with a data table and a filter panel. The table has columns for 'Salesman', 'Sales', 'Jan', 'Feb', 'Mar', 'Apr', 'May', and 'Jun'. The filter panel below the table has three buttons: '指定行字段' (Specify Row Fields), '选中字段' (Select Fields), and '关联字段' (Associate Fields). Below these buttons are input fields for '指定行字段' (containing '2'), '选中字段' (containing 'Sales'), and '关联字段' (containing 'Order').

Salesman	Sales	Jan	Feb	Mar	Apr	May	Jun
王	10000	8000	9000	7000	10000	7000	8000
李	8000	7000	8000	9000	8000	9000	8000
张	9000	8000	9000	8000	7000	8000	9000
赵	7000	8000	9000	8000	7000	8000	9000
陈	8000	9000	8000	7000	8000	9000	8000
周	9000	8000	9000	8000	7000	8000	9000
吴	7000	8000	9000	8000	7000	8000	9000
郑	8000	9000	8000	7000	8000	9000	8000
冯	9000	8000	9000	8000	7000	8000	9000
孙	7000	8000	9000	8000	7000	8000	9000

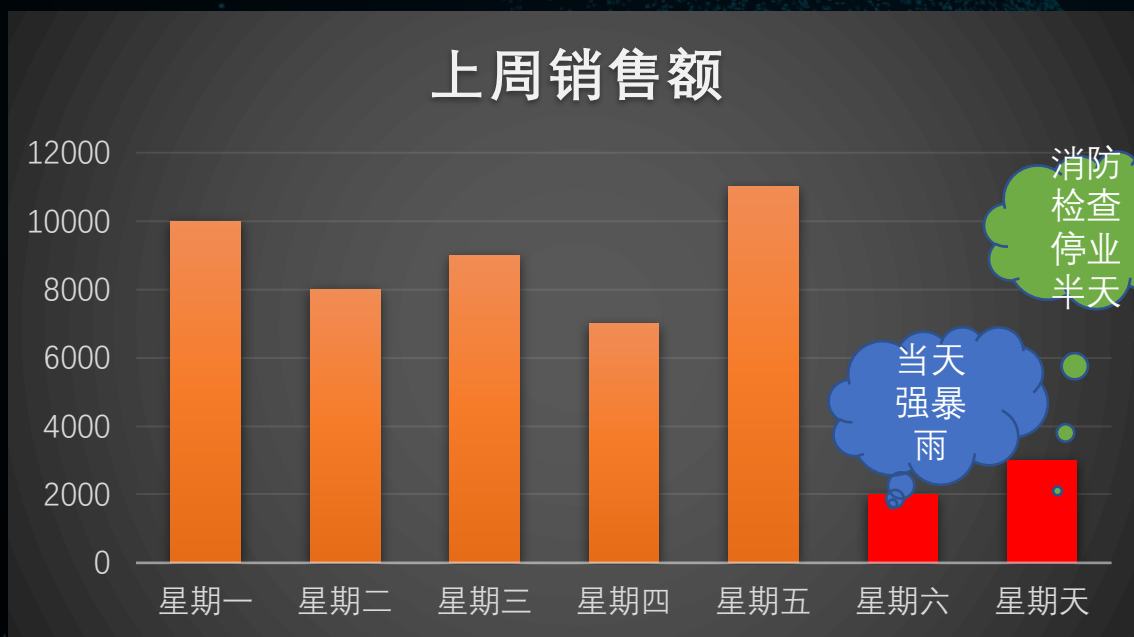




- 问：上周我们店每天的销售额是多少？
- 答：

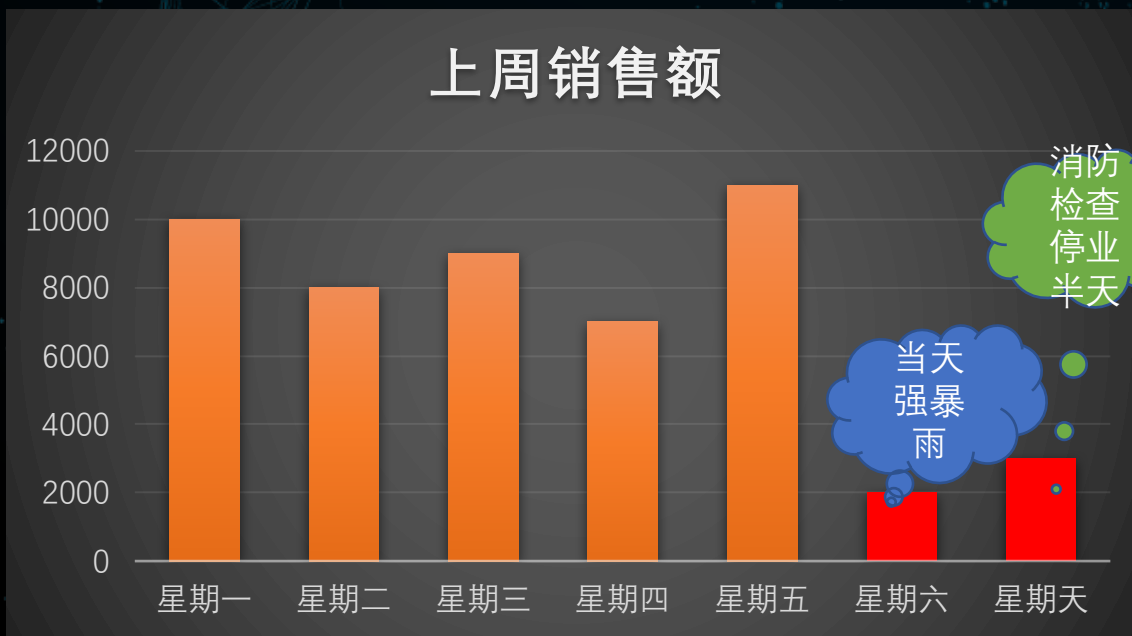


- 问：上周我们店每天的销售额是多少？
- 答：



• 我们要实现这样的一套人工智能系统，需要哪些能力？

- 能够理解自然语言的人机对话能力
 - NLP+自然语言理解技术
- 能够关联企业内外部各种数据的数据架构
 - 知识图谱
- 能够基于数据架构进行计算和推理的AI引擎
 - 机器学习+符号推理+大数据



与传统数据架构对比

数据仓库 vs 知识图谱

- 企业上一代数据组织架构是数据仓库，而新一代的企业数据组织架构是知识图谱

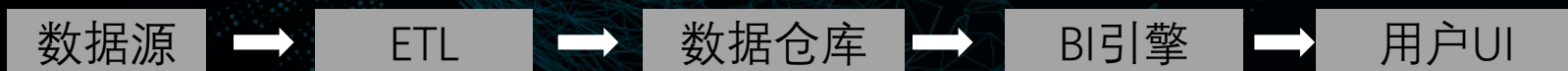
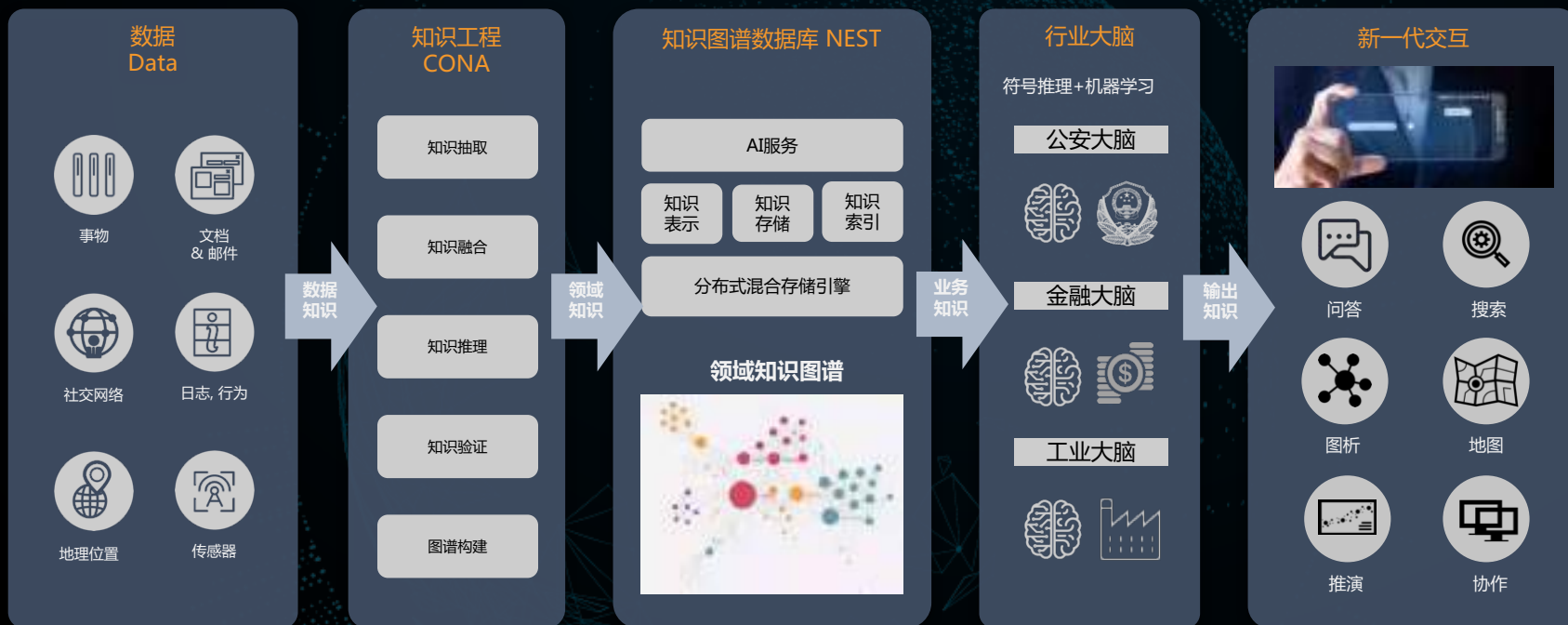
商业智能 BI vs 人工智能 AI

- 企业上一代数据应用主要为BI，而AI将成为为新一代企业主流应用

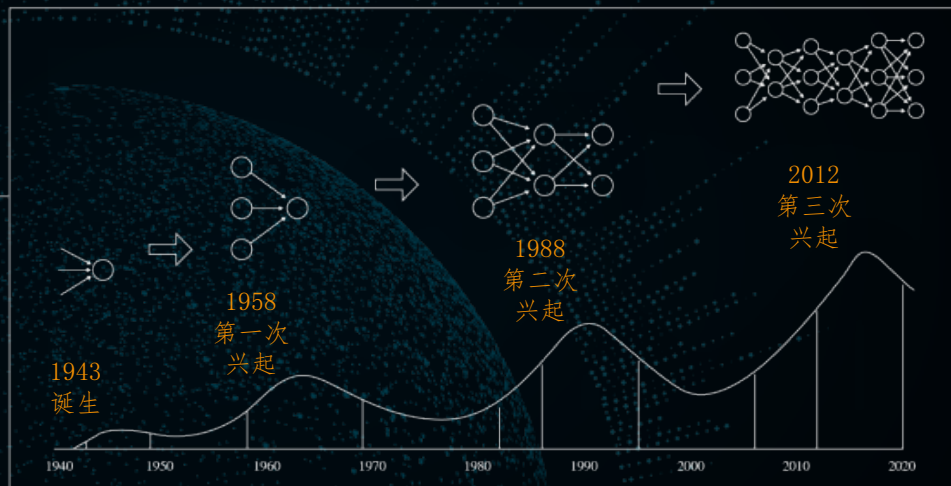
传统UI vs 新一代交互

- 企业里复杂的定制化软件界面将会被以对话形式为主的问答系统代替

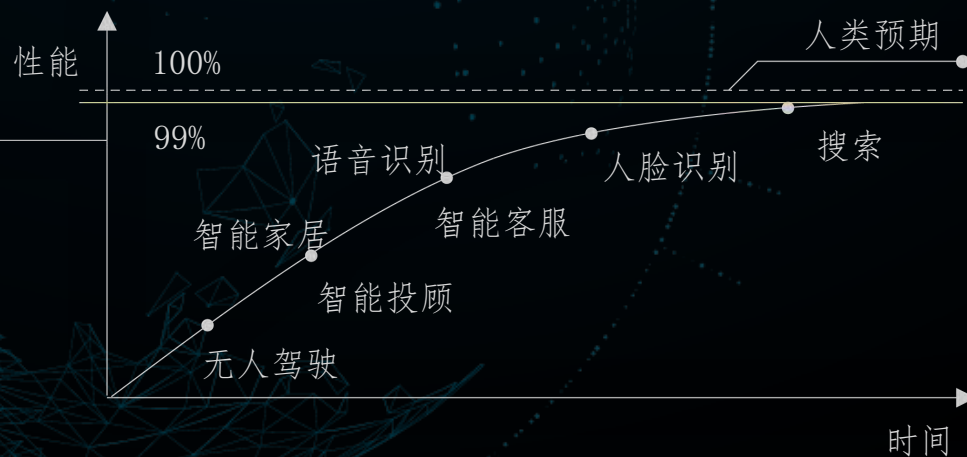
明智系统产品体系



人工智能历史曲线



人工智能技术与人类预期



Google Bai**du**百度 Bing 搜**狗**搜索

YAHOO! 淘**宝**网
雅虎中国 Taobao.com

搜索：最成功的人工智能商业应用

价值

AI算法无法做到100%准确，但随着数据反馈越来越多，算法正在变得越来越准确

效率

以效率思维选择工具，“人类+工具”让人类实现自身能力的延伸和飞跃

易用

易用的人工智能才能被广泛使用通过为个体赋能形成群体智慧

人工智能的商业化核心是让用户合理的接受机器的错误

企业级人工智能市场缺乏好的产品，好的体验

- 使用价值低：大量业务处理过程仍需人工干预
- 使用体验差：操作不符合人类思维习惯
- 使用效率低：功能复杂，培训成本高
- 不可解释性：黑盒操作，没有证据支撑结果

小明：企业级人工智能统一入口

新一代人机交互范式，以对话的形式提供业务支持

23th Wednesday
2017-04

822案件

盛世天骄打人案件

长安公园拾胎案

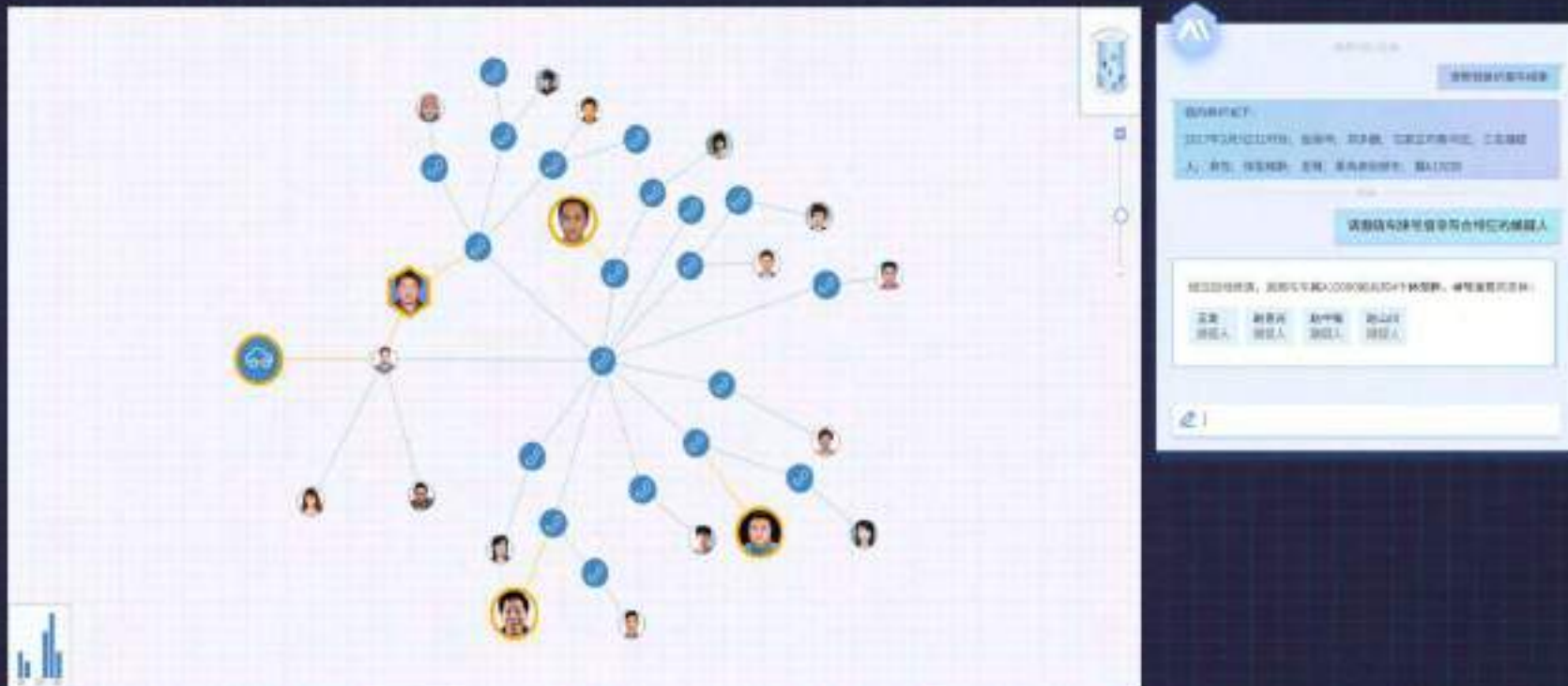
729案件

案件分析界面截图，显示了案件的基本信息和嫌疑人的详细资料。

2017年3月5日21时	赵南华	郑多健	石家庄市新华区
案发时间	嫌疑人	受害人	案发地点
三名嫌疑人	男性	体型魁梧	年龄
2000+	170+	170+	20-30
冀A10080	车牌		

案件描述

2017年3月5日21时许，赵南华（男，郑多健）石家庄市新华区盛世天骄小区内下车时，三名嫌疑人上前阻拦赵南华下车并对其进行“逼停”，双方发生口角进而产生肢体冲突并造成多人受伤等恶性，赵某鼻骨骨折并，经接受家人提供的信息，嫌疑人中有一男性体型魁梧，且被其同伴称作“发哥”。此外，在案发现场中，发现嫌疑人员上了一辆黑色奔驰轿车，车牌号为冀A10080。





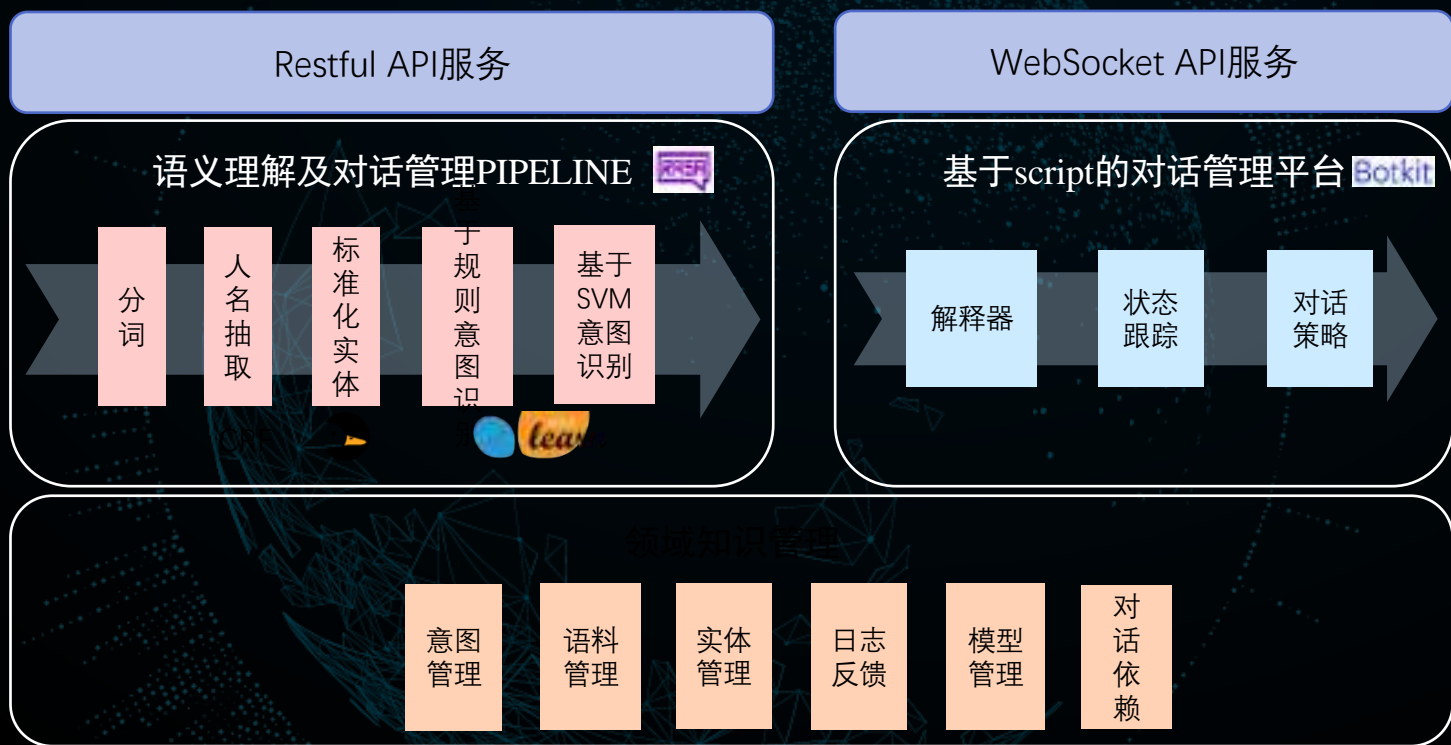
移动轨迹、移动轨迹的关联分析

关联分析、移动轨迹、移动轨迹、移动轨迹

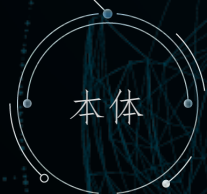
移动轨迹分析及其关联分析在移动轨迹的移动轨迹

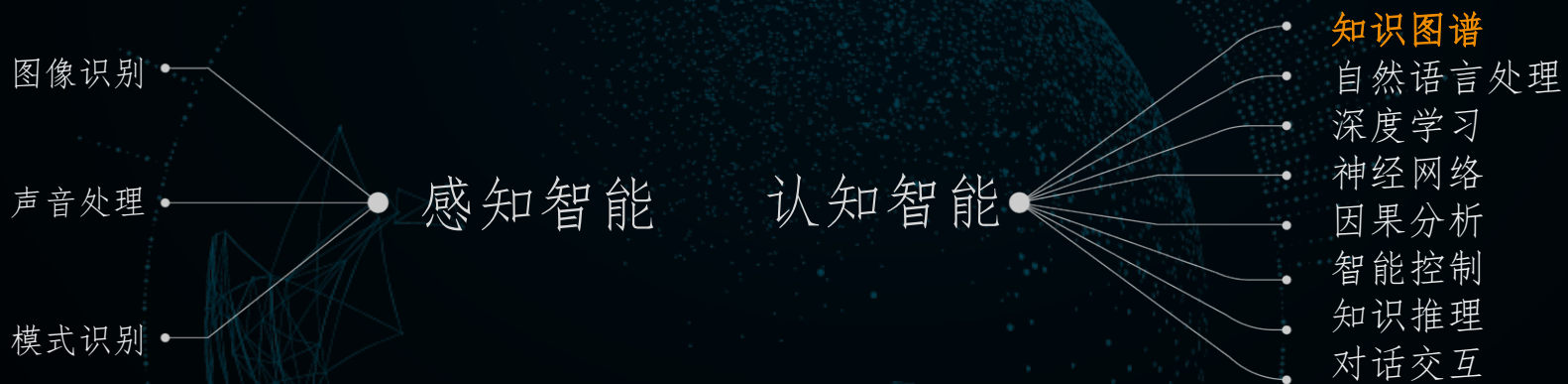
轨迹已播放完毕

小明技术架构

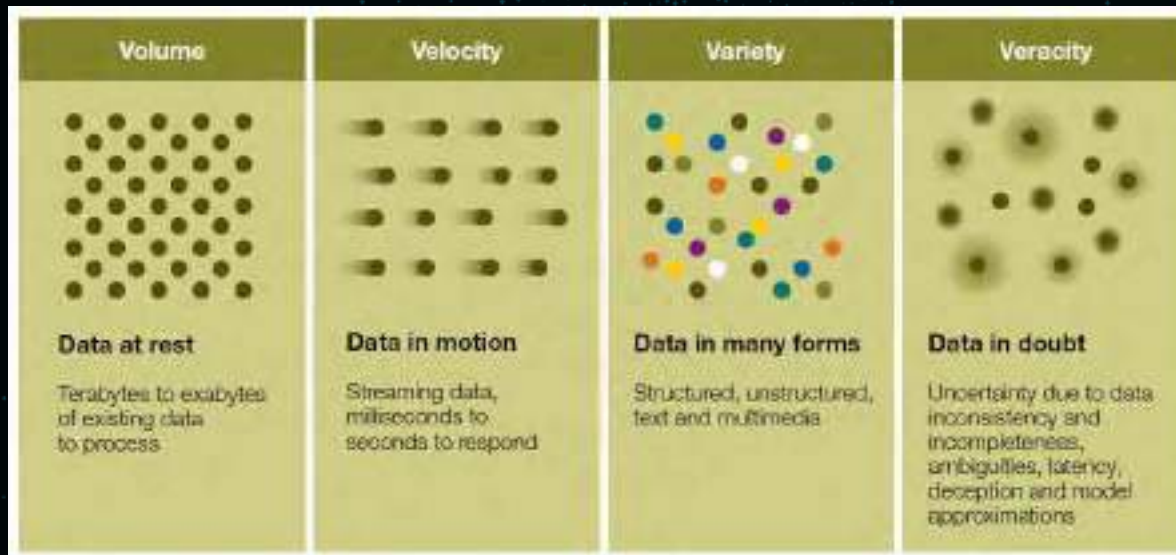


小明的核心支撑：知识图谱









企业级人工智能是复杂性业务的综合体现



知识图谱是机器大脑中的知识库、下一代人工智能应用的基础设施
让计算机像人类一样理解世界，用人的思维模式更精准地量化真实世界



2012年
Google推出知识图谱



如今百度知识图谱可直接反馈答案
并推荐相关信息

- 更高质量
- 表示更加规范
- 语义自描述
- 更加可计算
- 更加 Machine-Understandable

明略数据“知识图谱数据库”

蜂巢
NEST

知识图谱数据库是使用基于图数据库的混合存储技术实现大规模知识图谱数据存储，是机器大脑中的知识库、人工智能应用的基础设施。



基于知识图谱数据库的落地实践



格物致知 ← 明智系统 → 人机同行

- 将先贤经验/知识转化为规则，混合深度学习，探究事物本质，量化世界
- 通过知识图谱来建构新一代企业数据架构，让以前模糊的商业决策变得更加清晰高效，构建未来真正的竞争力。

NEST
知识图谱数据库

中国企业发展新模式
新技术能力、全面在线化、
智能大脑

通过简单好用的新一代交互入口，将行业知识赋能个人，释放人类的创造力，帮助人们在时代大拐弯中顺利掌握未来的驱动力。

LiteMind (小明)
简单好用的企业级人工智能入口

明智系统产品理念

明智系统核心技术

多规融合知识图谱构建技术；知识图谱数据库：NEST；行业人工智能技术；人机交互：LiteMind（小明）

1. 多规融合知识图谱构建技术

在企业的数据库、文档、传感器、访问日志、邮件、图像、音视频等数据中包含了大量反应企业业务的数据事实和领域、业务知识，明智系统通过多规融合的知识图谱构建技术，从多源异构的数据中抽取实体信息，链接和融合实体，推理补全属性，识别语义并建立关系，并最终将知识存储于知识图谱数据库中。这其中包含的实体-关系抽取技术，是知识图谱构建的核心技术之一。

- 结构化知识抽取：针对结构化数据，支持关系型数据库、流式数据、CSV/Excel等多种数据格式，大规模自动化的采集、清洗、归类、标准化、关联所有数据，形成统一数据视图；通过基于规则的和基于机器学习的知识抽取技术，识别、链接、融合实体，并抽取、补全属性，建立实体间的关系关联；
- 非结构化知识抽取：针对海量的文本、网页数据，通过自然语言处理（NLP）技术和机器学习技术将非结构化的文本数据，转化成知识图谱的实体-关系。支持业务人员可基于规则自由定义概念、实体和关系，基于规则的语法模式抽取固定模式，再通过机器学习扩展和清洗关系，智能化制定和理解行业领域的语义。同时，支持概念及概念间关系自动获取，包括分类层次关系、结构关系、同义关系、实例关系等。

2. 知识图谱数据库技术--NEST

知识图谱是机器大脑中的知识库、人工智能应用的基础设施。明智系统通过使用基于图数据库的混合存储技术实现大规模知识图谱数据存储。不同于以往关系型数据库的存储，明智将所有数据以实体-关系-事件-属性的形式存储，这种形式最直观的展示了数据以及数据背后的关联，真正做到了“所见即所得”；明智系统充分利用知识图谱数据存储，优化搜索引擎，并实现基于语义的检索和多种复杂推理

- 混合存储技术：包含图数据库、列式存储、全文索引和文档库
- 复杂图谱技术：边爆炸、超级节点、多点并发、全文索引等问题解决与性能优化
- 语义检索：基于NLP技术和知识图谱技术进行实体识别、全文检索和子图匹配实现基于自然语言的复杂语义搜索
- 推理引擎：结合领域知识，实现自动长链推理、意图识别、因果分析，同时通过学习让引擎更加理解领域

传统数据库 vs 知识图谱数据库

找出与尾号2134的红色轿车有关系的名叫“发哥”和“体型略胖”的嫌疑人

使用传统数据库

程序员或DBA



传统关系型数据库

数据存储以结构化为主，有数据量限制，不易扩展



存储数据对象或陈述性知识

存储治理后的数据，用于之后的各种查询和统计



搜索代码繁琐

需要写程序，并通常需要多个系统配合



人工推理

需要大量时间进行人工分析



使用知识图谱数据库

普通用户



分布式知识图谱

存储结构化、半结构化、非结构化数据，支持PB及以上数据量，使用可动态扩展的本体模型



存储陈述性和程序性知识

不仅存储多元异构的大规模数据，还存储基于数据的规则与模型，用于解决“做什么”和“如何做”等问题



自然语言语义搜索

智能入口，无专业要求



系统推理，辅助人工

系统自动推理辅助分析，大大节省时间



3. 行业人工智能技术

基于明略知识图谱数据库NEST提供的基础AI能力如实体识别、问题理解、推理执行、假设预测、因果分析等能力之上，我们结合公安、金融、工业等实际业务场景，将符号推理和机器学习、深度学习等AI技术相结合，最终实现解决实际业务问题的行业人工智能。

公安大脑：

- 基于人、事、地、物、组织构建的包含社会关系网络与事件的知识图谱，实现高效线索研判和深度挖掘预警。其能力包括案件线索解析、人案关系智能推演、犯罪团伙挖掘、时空轨迹比对、高危人员预测预警、情报数据挖掘分析等。

金融大脑：

- 基于企业、产品、证券、机构、个人、账户、交易等构建金融知识图谱，挖掘隐藏在复杂网络之下的关联关系风险和资金流动异常。通过机器学习、图挖掘等技术应对金融监管、风控、营销等领域的业务需求，构建担保风险模型、风险传导模型、反洗钱模型、企业派系模型、一致行动模型、实际控制人模型、图谱全息画像模型等

工业大脑：

- 将故障、传感器、时域、频域等数据关联起来，构建工业和物联网领域的知识图谱，基于知识图谱数据库实现对海量工业领域知识的存储、实时更新、检索、追溯，从而智能挖掘系统的异常特征设备状态监控、设备健康管理、点巡检、多维分析、视觉识别等

4. 人机对话技术—LiteMind (小明) : 简单好用的企业级人工智能入口

LiteMind (小明) 通过可感知、被信任的拟人交互过程实现人机同行, 降低系统使用难度, 提升企业级应用的用户体验。

例如在公安领域, 普通的警察无需学习图析等复杂操作, 对小明输入“给出与某车牌相关的胖子”这样的自然语言, 明智系统即可自动完成线索推导。

自然语言理解能力

以分类模型和序列标注为核心的语义理解模型能够精准的将自然语言转化为框架语义表示的结构化信息。同时, 接口的开放性支持不同领域知识的快速接入, 方便定制出领域自然语言理解接口。

对话能力

小明提供多种对话模型, 并允许以脚本配置的方式自定义具有复杂逻辑的对话模型, 为不同的业务系统提供统一的自然语言交互接口。

实体抽取及标准化

以“规则+模型”的方式提供了多个支持二次开发的实体识别模型。基于知识图谱构建实体识别能力, 能够快速定位和理解自然语言中的实体, 以及能够识别领域无关的实体, 并且可以通过简单配置快速识别未登录的领域实体。

明智系统构建行业大脑

基于知识图谱数据库，通过统一入口“小明”，以对话的形式提供业务支持

公安 大脑

基于人、事、地、物、组织构建的社会关系网，有助于实现公共安全领域的高效线索研判和深度挖掘预警。

实现人案关系可视化的智能推演分析，提供精准高效的可疑团伙和高危人员预测预警。业务上达到大案快破、小案多破、提前预警，实时管控的作战效果。

金融 风控大脑

基于“企业、个人、机构、账户、交易、以及行为数据”来构建金融行业知识图谱，挖掘隐藏在复杂网络之下的关联关系风险和资金流动异常；全面提升风控专家在海量数据下，精准甄别、有效防范和化解业务风险的效率，实现监管科技的目标。

工业 安全大脑

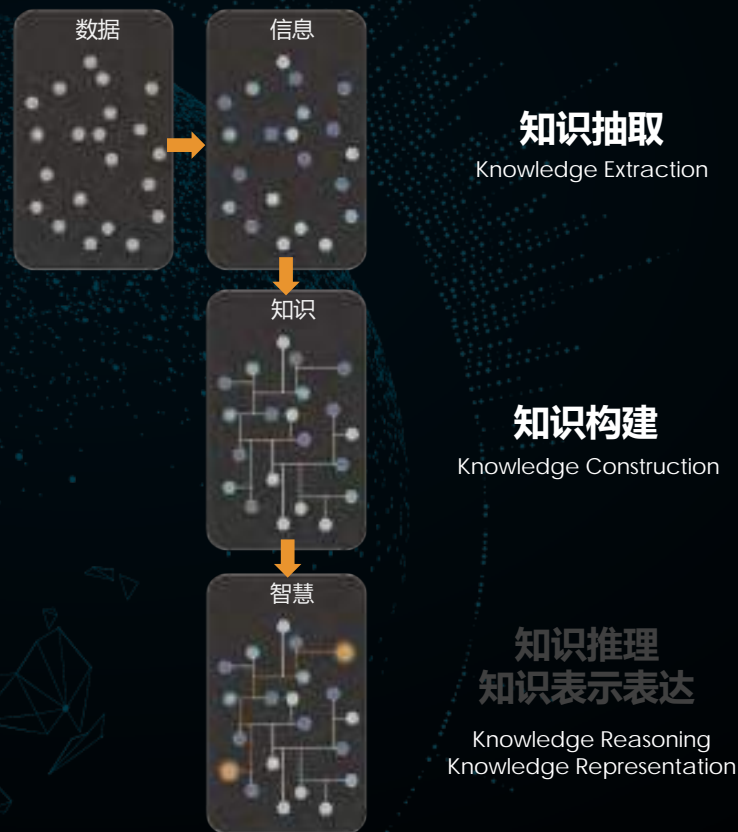
围绕设备，将故障、传感器、时域、频域等数据关联起来，智能挖掘系统的异常特征，做到设备状态实时监控、故障预警，高效避免系统安全事故，极大降低维护成本

2014年，明略凭借核心团队八年以上的大数据处理、分析的经验积累，打造高效易用的、跨平台的企业级数据整合管理系统（支持500节点以上的超大集群、数据之间秒级交互、3秒内快速灾备恢复），帮助客户将不同数据源的、结构化和非结构化数据融合在一起。

2015年，明略以客户为核心，立足行业为客户治理、连接、挖掘数据，发现新数据。率先推出公安、金融、工业、税务、政企、营销6大行业解决方案，通过数据挖掘找到业务规律，将经验转化为信息，释放数据价值。

2016年，明略经过36个业务领域的探索，与688位客户的深度接触，将图数据技术落地到行业应用中，构建适合多维多层关系挖掘的新一代企业数据架构，经过关系关联挖掘找到业务规律，为决策者提供用于解决问题的结构化信息——知识。

2017年，明略将业务范围锁定在公安、金融、工业与物联网三大行业，并围绕**公安大脑**、**金融风控大脑**、**工业安全大脑**，与客户一同构建**行业知识图谱**，开展决策智能自动化工程。行业人工智能“**明智系统**”诞生！



Thanks !

问见研习社

