



线上研讨会

AI为擎 引领数智未来

——加速企业“数智化”创新，实现数据红利

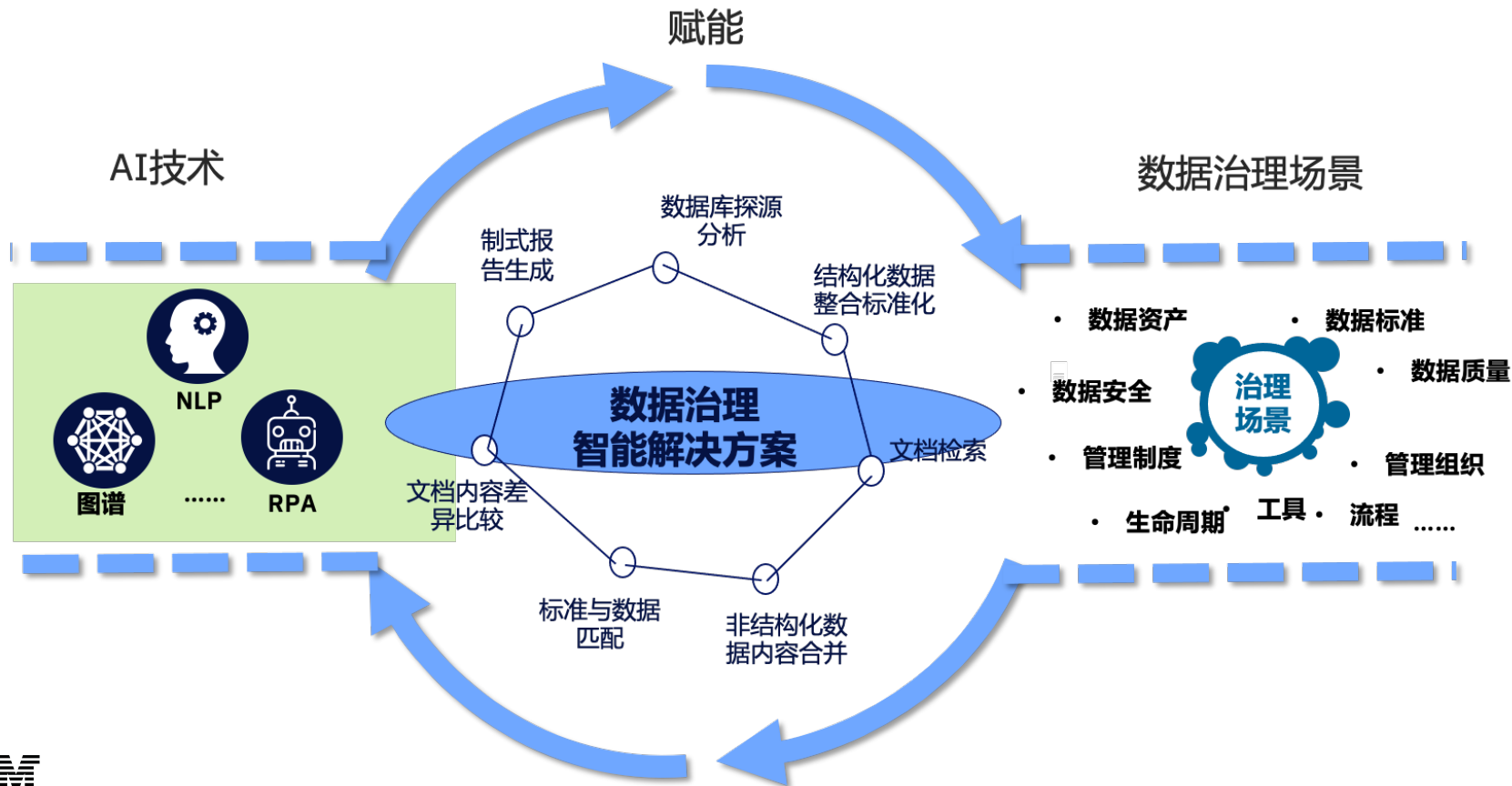
数据治理领域痛点分析

「**领域知识门槛高**」：传统的数据治理方式需要配合丰富的行业知识和深刻的业务理解，治理团队业务掌控能力弱会导致治理成果发生偏离。

「**需求市场价格敏感度高**」：客户对治理工作的价值普遍认可，市场潜在需求大，但计划投入的预算非常有限；传统方案的报价往往是在客户预算承受范围之外的，能够率先降低预算的厂商会占有极大的竞争优势。

「**非结构化数据管理难度大**」：由于非结构化数据类型和表现形式多样，对于传统的数据治理而言，非结构化的元数据提取和管理还是一个新领域。

数据治理智能解决方案技术框架



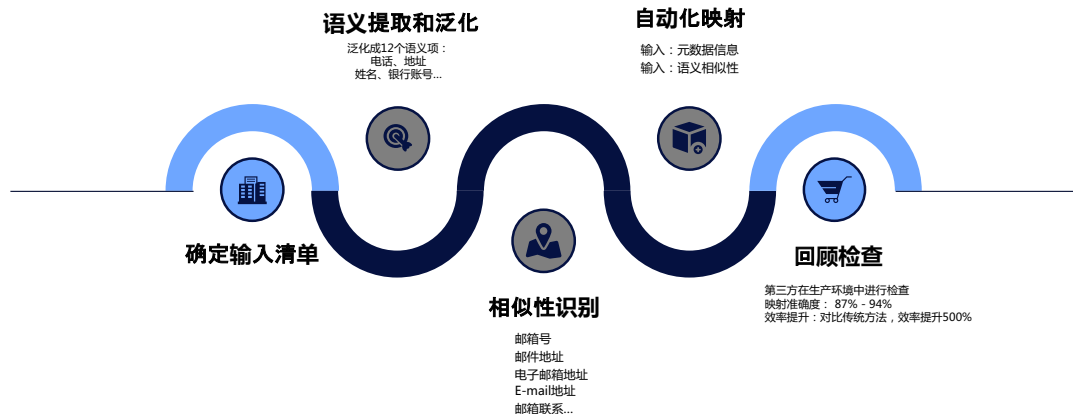
案例分享：某证券智能化数据治理

落实外部数据分类分级监管要求

项目背景：2018年9月，证监会发布了《证券期货业数据分类分级指引》，对证券公司加强数据安全治理，建立健全数据管理制度提出了更高的要求。

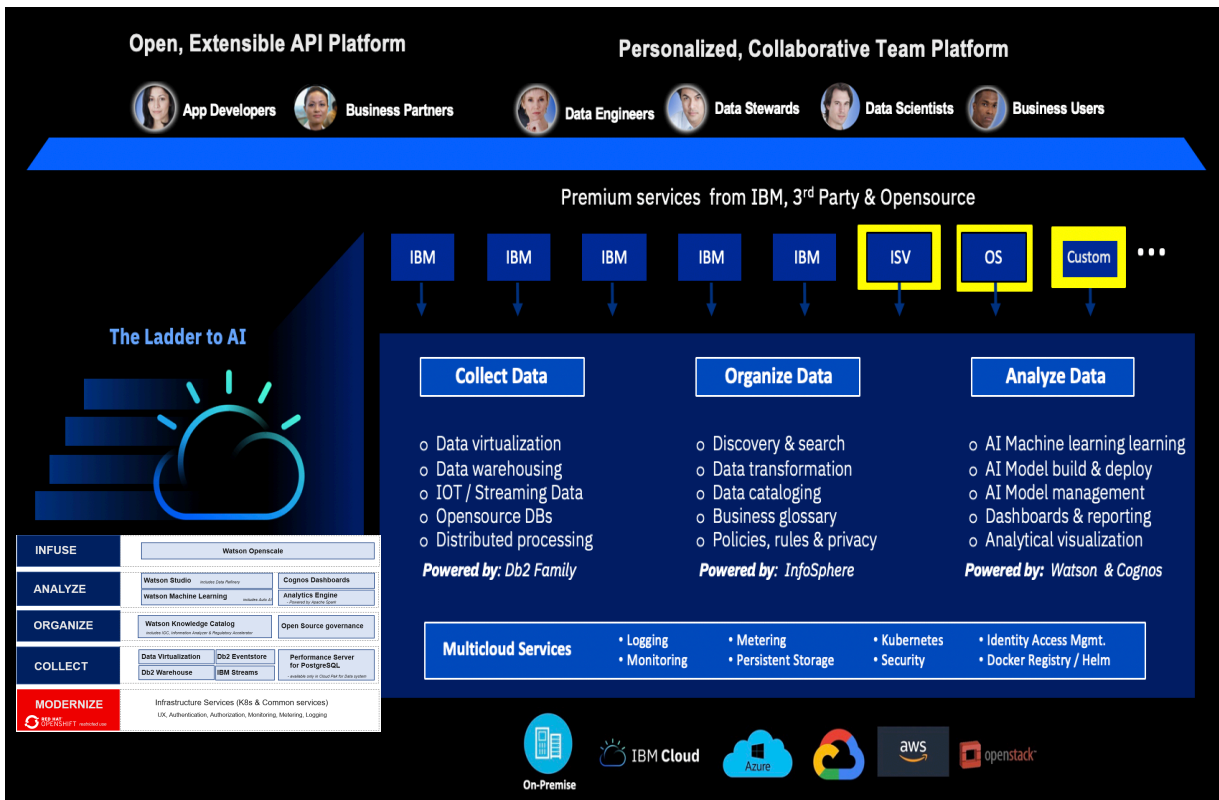
- 客户重要信息梳理包含地址、电话、口令、密码、姓名、银行账号、用户名、邮箱、账号、账户余额、证件号码、证件类型等**12**个类别。
- 重要信息映射范围**10**个系统，包含**11344**张表，**203763**个字段，其中涉及客户重要信息的表**4681**张，占比**41%**，字段**13839**个，占比**7%**，重要信息在系统中较为分散。

用智能化的方式：在**3**个工作日内实现**10**个系统**4000+**表的映射扫描，而传统方式至少需要一名有经验的数据分析师**15**个工作日以上时间



Cloud Pak for Data - 数据中台快速落地的平台

集成的一站式数据及人工智能平台，助力企业AI之旅，数据驱动业务创新



- ✓ **云原生架构**，包括容器化工作负载、微服务和多云供应，灵活、高效、成本低
- ✓ 业界领先的异构数据源**数据虚拟化** -- 连接到远程系统，而不将任何数据移动到我们的平台
- ✓ **内置数据治理**，保证数据可信、合规、业务就绪
- ✓ AI开源编程语言、深度学习框架和开源接口（如Python, TensorFlow, Jupyter, and SparkML）
- ✓ AutoAI自动机器建模
- ✓ **数据开源框架**，如Apache Hadoop, MongoDB
- ✓ 通过**实时流数据**应用程序实现连续智能
- ✓ 混合多云部署
- ✓ IBM及第三方生态系统

Cloud Pak for Data最主要的4个企业使用场景



随时随地管理数据

1. 管理所有企业数据，而不管它位于何处；通过内置的数据治理和隐私，实现自助服务和数据消费（数据虚拟化）；为数据湖提供数据治理能力；

2. 获得控制并利用实时数据进行分析

(快速数据摄取和流分析)



实用性数据科学和AI

规模化构建、部署、管理和治理模型和数据，以改进业务成果例

如：A. 客户流失

B. 交叉销售/追加销售

C. 预测性维护



转移到云原生数据工作负载

转移到云原生

- a. 提供和扩展数据及人工智能服务
- b. 一次构建，随时随地部署-多云支持
- c. 内置自动化、扩展和协作以提高生产率



更明智的治理

支持自助服务分析的治理

自动发现元数据、管理治理规则和策略、实施隐私等，以降低风险并确保合规性

例如：加速GDPR合规性

借助Watson Knowledge Catalog 组件完成一站式数据治理

