

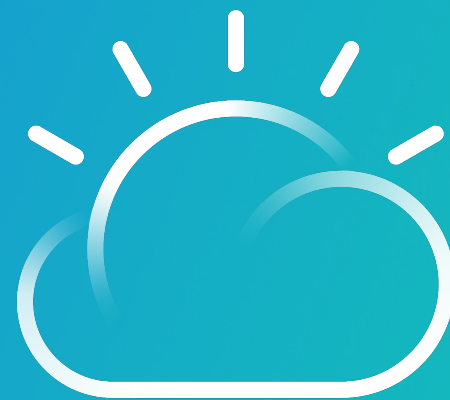
# 赋能员工专注于更有价值的工作

—

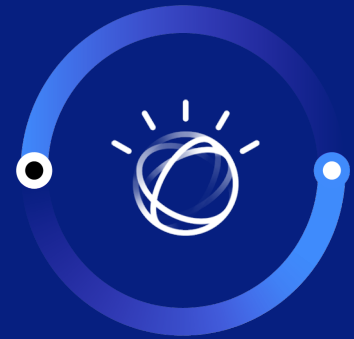
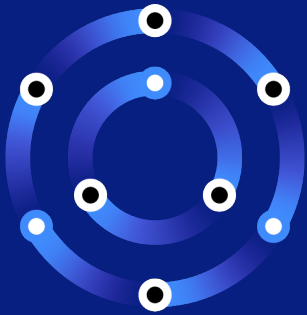
朱辉

IBM 全球机器学习研发中心总经理

IBM 中国认知计算与数据分析研发中心



# 企业正在加速人工智能之旅。然而在落地的过程中碰到诸多挑战，其中之一就是新人才和新技能的缺乏



## 收集组织数据

解决的难题：

- 数据孤岛，不可获取的数据源
- 缺乏质量或线性控制
- 无真正的自助式数据访问
- 查询效率低下

## 分析

解决的难题：

- 数据科学和人工智能方面的技能较低
- 数据科学项目实现价值的时间过慢
- 对 AI 模型风险和惩罚的恐惧
- 在 AI 上缺少投资回报

## 融合

解决的难题：

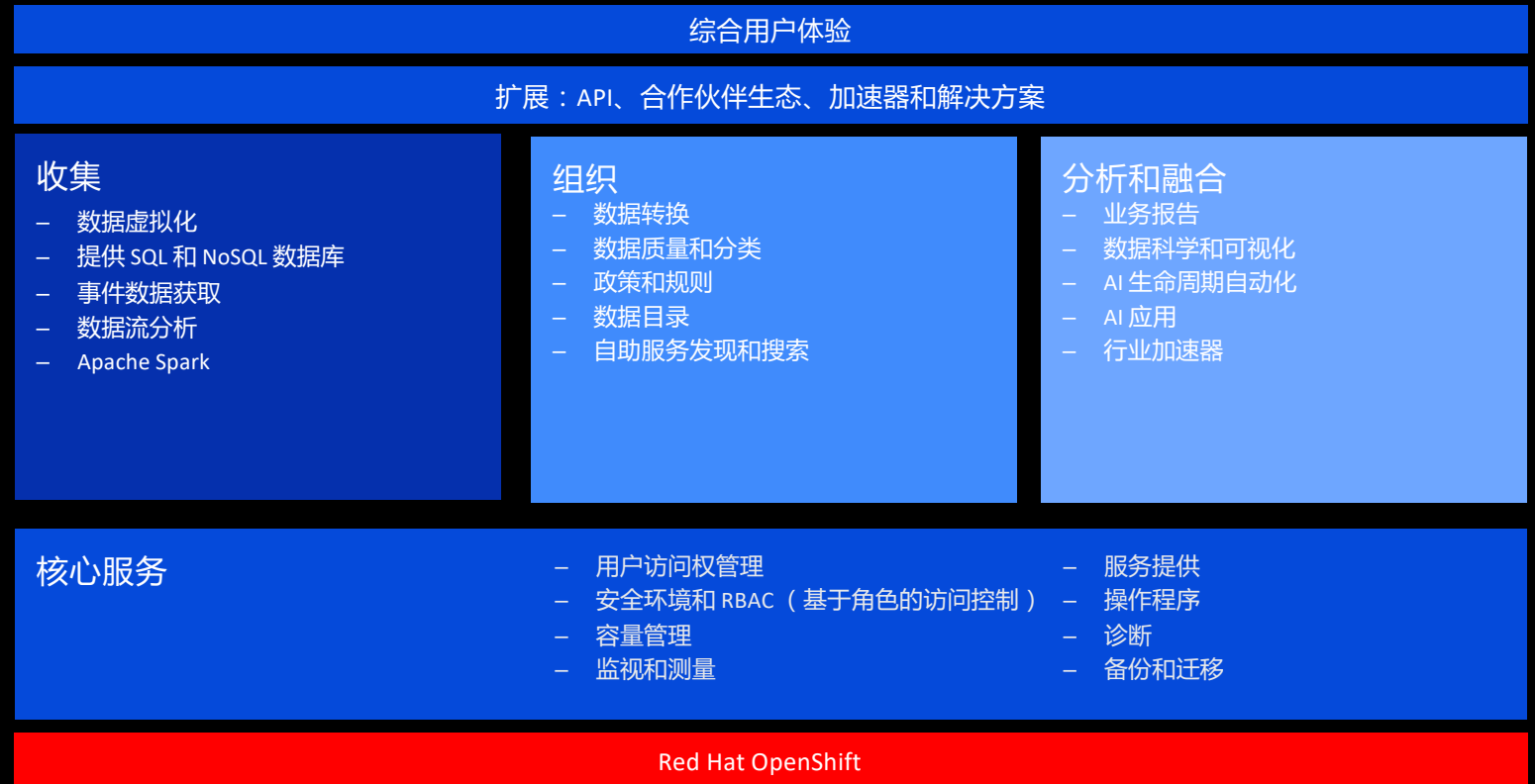
- 客户服务缺乏自动化
- 运营与财务脱节
- 预防金融犯罪
- 无法将 AI 引入核心 IT 领域

# IBM Cloud Pak for Data

统一、模块化、可在任意云平台部署

应用开发人员 | 合作伙伴 | 数据工程师 | 数据管理员 | 数据科学家 | 业务用户

## AI 阶梯



IBM Cloud | Amazon Web Services | Microsoft Azure | Google Cloud | 超融合系统

# IBM Cloud Pak for Data 3.0 关键亮点

## 降低数据操作成本

- BigSQL 的新服务可优化数据库操作
- 支持 Red Hat OpenShift Container Storage 软件，改善存储的同时降低成本。
- InfoSphere Virtual Data Pipeline 现已成为标配；额外的安全层可降低个人数据落入不当人士手中的风险。

## 借助 AI 应用加速实现价值

- IBM Planning Analytics 可作为扩展；快速自动化规划、预算和预测。
- 为满足客户服务数字化需求，Watson Assistant 现已增加自然语言处理功能
- 内置 AutoAI 功能，可使大部分 AI 流程自动化，确保数据科学家能够高效工作。

## 借助不断成长的生态系统增强创新能力

- 支持 IBM Power Systems
- 第三方数据即服务，包括 Weather Company 数据、人口统计信息、股票市场等。
- 集成 Red Hat Marketplace
- 行业加速器，包括媒体、交易产品服务组合管理、公用事业的预测性客户分析。

## 降低安全风险

- InstaScan 的新功能可以识别数据中的热点，同时减少合规所需的时间
- 集成 IBM Guardium，提高平台数据安全性
- 更新至 Red Hat OpenShift v4.3，增加平台底层的核心容器安全。
- 全新的 Watson OpenScale Model Risk Management (MRM) 能力满足监管要求

AutoAI 为全球的数据科学家“减负” ——  
通过 AutoAI，构建 AI 模型的速度更快，质量更优，迅速弥补 AI 人才缺口，加速企业的 AI 转型

## AutoAI ( IBM )



- ✓ 迁移学习
- ✓ 数据准备
- ✓ 先进的数据提炼器
- ✓ ML 和神经网络选择
- ✓ 特征工程
- ✓ 超参数优化
- ✓ 一键部署
- ✓ 可解释性和去偏见化
- ✓ AI 模型精度的完善
- ✓ 集成在 AI 平台中
- ✓ 反馈训练回路
- ✓ 分布式深度学习

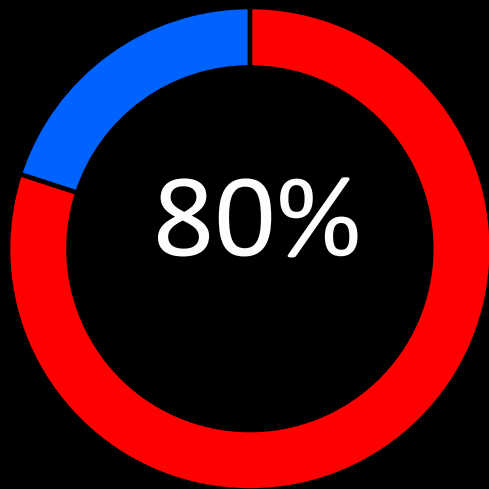
## 相较于

## AutoML ( 其他 )

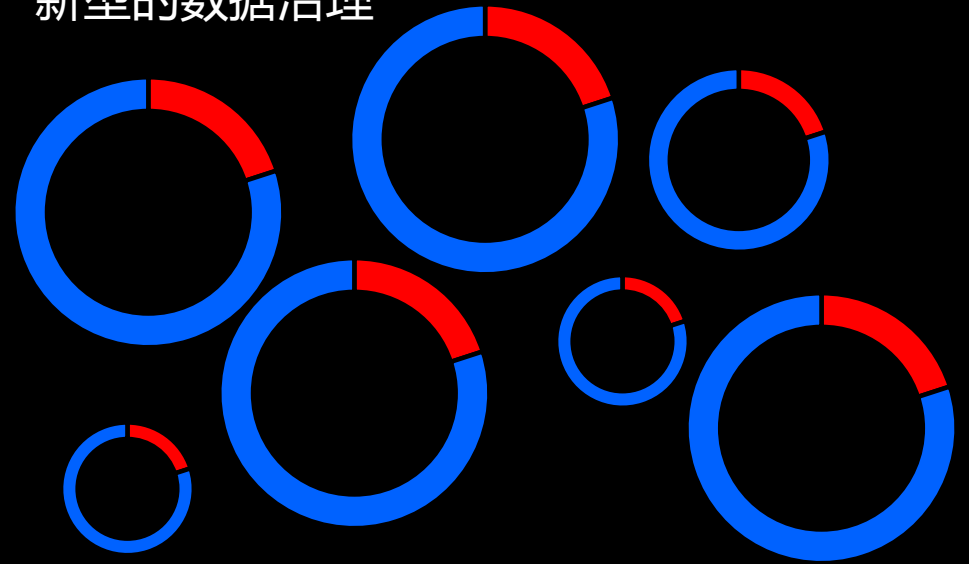
- ✓ 数据准备
- ✓ 特征工程
- ✓ 超参数优化
- ✓ ML 和神经网络选择

95%的组织看到数据质量差带来的负面影响，数据准备往往要花费几个月甚至几个季度，大大降低了数据所能发挥的作用

### 传统的数据治理



### 新型的数据治理



单次迭代  
耗时长，风险大

多次快速迭代  
发现理解、摄取整合，评估质量，清理数据  
周期短，见效快

# 使用数据虚拟化技术大大降低数据工程师的技能需求

- 消除数据孤岛，构建企业级的数据地图
- 无需移动数据，连接多种数据库 —— 甚至是全球分布的数据库
- 无需编写复杂的跨库脚本，轻松查询多个数据源，简化分析流程
- 降低了 25% to 64% 的 ETL 请求
- 数据虚拟化比起传统的联邦查询快 40 倍

数据无处不在，企业正在从越来越多样化的数据源收集数据，以分析和运行其业务，这些来源可能有数千，甚至数百万。收集、治理、存储、处理和分析这些数据中心的复杂性、成本、时间和出错风险正在成倍增加。因此，数据浪费是企业要面临的首要挑战。传统解决这个问题的方式大量精通数据库工程师进行手工抽取和转换和集成，而数据虚拟化技术给企业提供了新的选择。

IBM Cloud Pak for Data ,  
为企业发展持续赋能，让数据和人工智能  
更快地成为您的竞争优势



