



# 快速部署企业级 AI 平台 - IBM Cloud Pak for Data

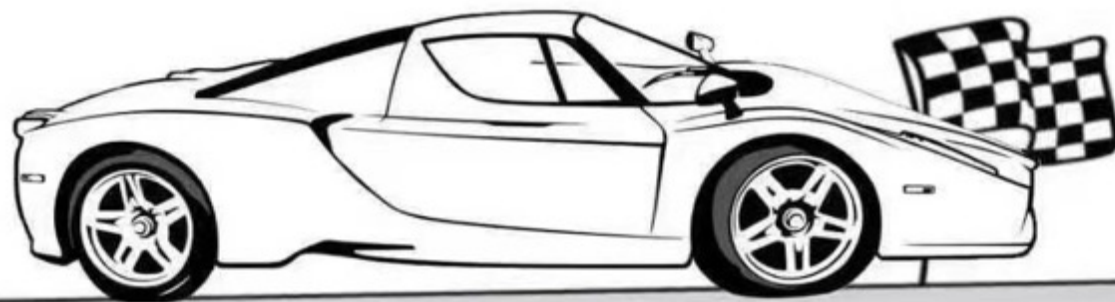


# 你眼中的大数据分析



数据提取

模型建立



深度学习, 人工智能

# 真实的大数据分析



需求讨论

提取数据

数据清洗

数据整合

缺失值处理

特征工程

模型评估



# 数据及人工智能团队可能的角色分工



CIO & CDO  
Establish goals



Executive  
Deliver on goals



Offering manager  
Leads business team



Data scientist  
Leads data team



Design manager  
Leads design team



Dev manager  
Leads dev team

---

## Stakeholders

---

## Leads



Data engineer



Data analyst



Data architect



Data trainer



Researcher



Designer



Developer

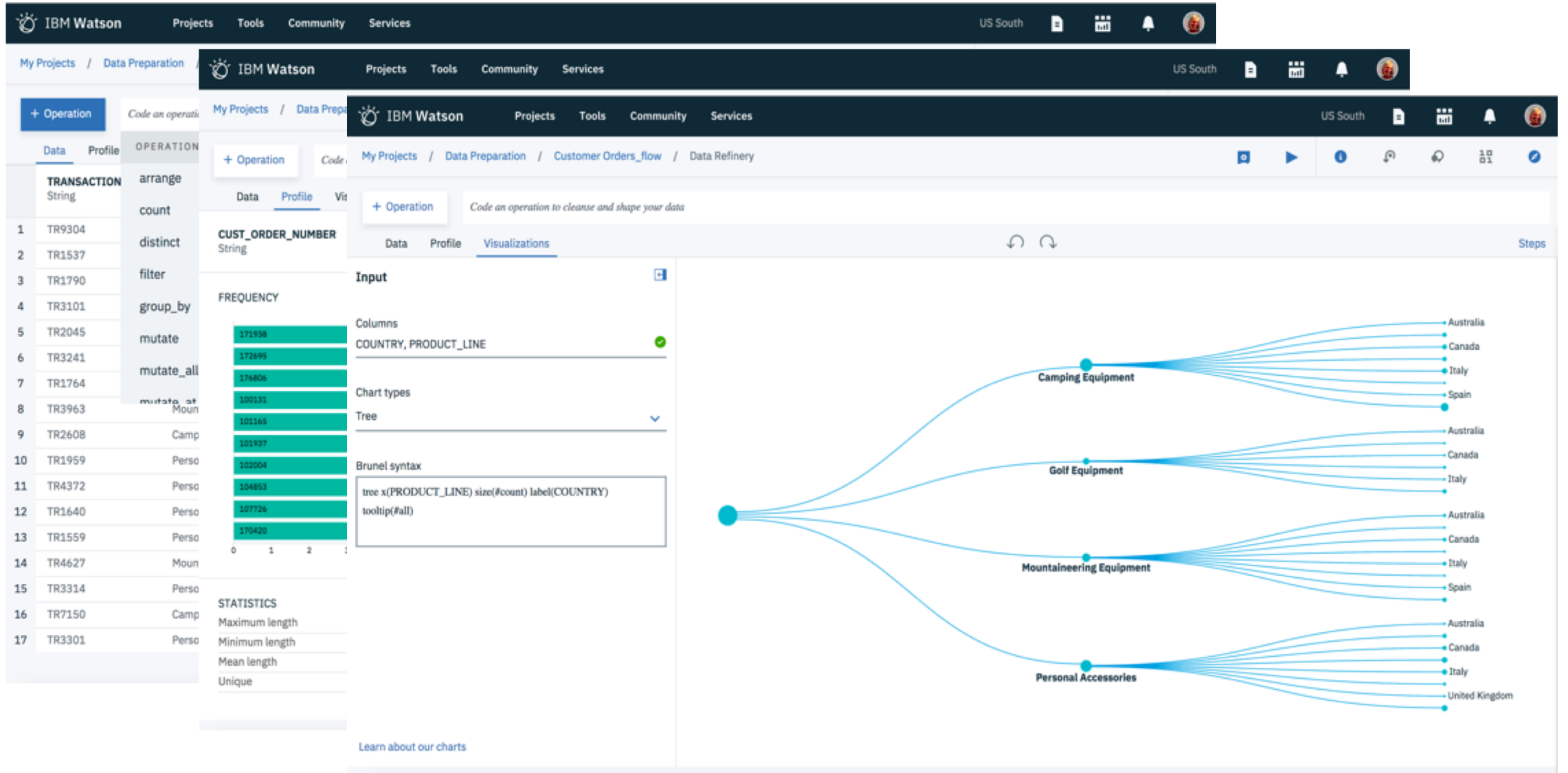


QA

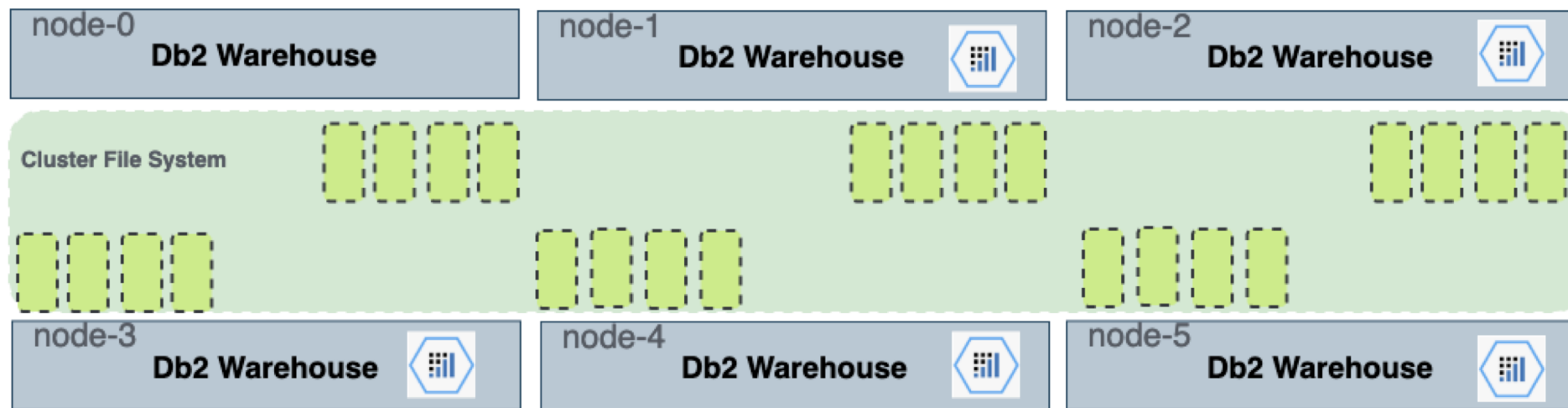
---

## Team

# 设计良好的数据处理工具

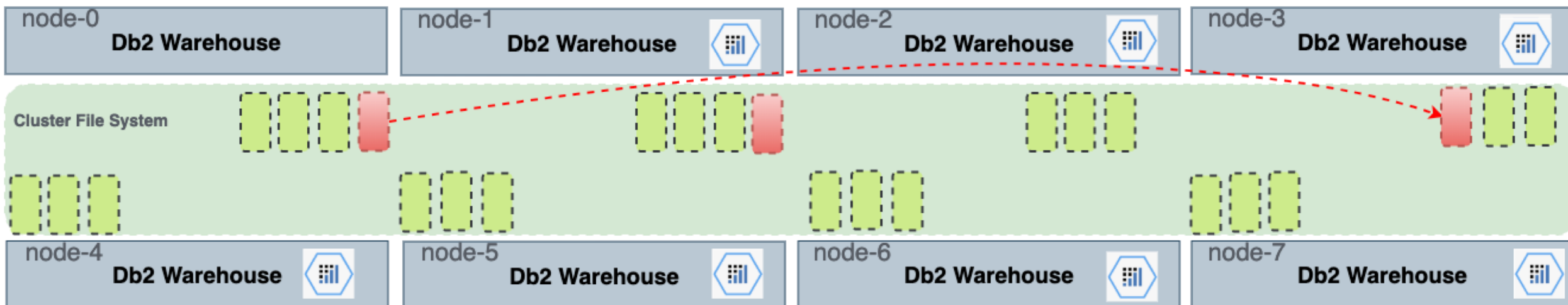


# 数据存储到高性能的分布式MPP数据库中



Scale in  
↑

↓  
Scale out



# 丰富的建模手段和工具



Notebooks RStudio

## 编码工具

将Python、R、Scala和最受欢迎的开源框架触手可及



Decision Optimization

## 优化

使用模型向导运行变量和约束优化模型



Watson Discovery

## 文本分析

解析并提取文本含义



SPSS Modeler



Data Refinery



AutoAI

## 可视化建模和自动化

鼠标点击进行数据探索，数据准备和建模，提高生产力和自动化



Cognos Analytics

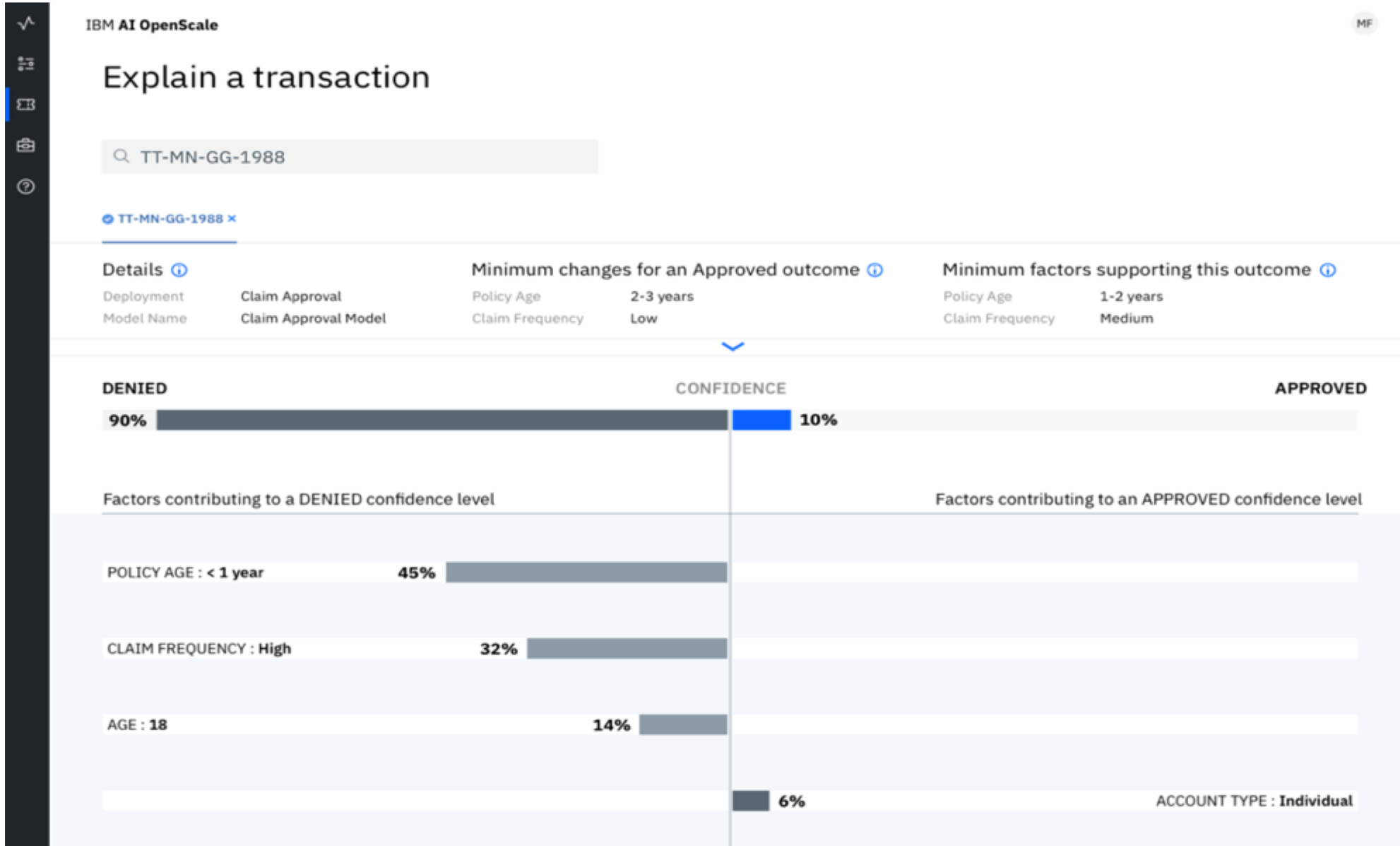
可视化和仪表盘  
推动和沟通见解

数据科学技能

高

低

# 模型运行的可解释性



# 数据、模型被当作资产管理起来

IBM Watson Knowledge Catalog Enterprise Catalog

Catalogs / Enterprise Catalog

Browse Assets Access Control Settings

Enterprise Catalog

Filter Any source Any type Any business term Any tag Clear all

Watson Recommends Highly Rated Recently Added Collapse

- Data Asset RESOURCEUSAGE
- Data Asset JOBRUN
- Data Asset JOBEXEC
- Data Asset IGBUSINESSTERM
- Data Asset HOST
- Data Asset IGBUSINESSCA

Overview Access Review Profile Lineage

DATA ASSET CUSTOMERS

Overall Rating 4.0

Review Summary

My Review

Other Reviews

Showing 3 of 3 reviews

- Susanna Tai | May 19, 2019
- Marc Haber | Jun 18, 2019
- Elaine Hanley | Jun 11, 2019
- Administrator IIS | May 17, 2019

与其他用户共享知识

Showing 25 of 28 items

| Name               | Owner             |
|--------------------|-------------------|
| BRANCH             | Administrator IIS |
| CUSTOMERS          | Administrator IIS |
| Customers 2019.csv | WKC Admin         |
| Db2 IADB           | Administrator IIS |
| EMPLOYEE           | Susanna Tai       |

IBM Watson Knowledge Catalog Enterprise Catalog

Catalogs / Enterprise Catalog / CUSTOMERS

Overview Access Review Profile Lineage

Current profile 165 classifiers 24 Columns 3,000 Rows

EMAIL\_ADDRESS PHONE\_NUMBER GENDER CUSTOMER\_ID CREDIT\_CARD\_NO EDUCATION

STATISTICS

| Statistic      | Value |
|----------------|-------|
| Unique         | 3,000 |
| Minimum Length | 11    |
| Maximum Length | 35    |
| Mean Length    | 21.6% |

Profile unavailable  
The profile is unavailable because the data in this column is anonymized.

主动策略实施的动态数据屏蔽



# IBM 数据及人工智能集成平台

## 通往人工智能阶梯 AI Ladder

**融合** - 与业务流程集成，大规模自动化部署

Watson API's - Watson Solutions - Watson OpenScale  
**Watson**

**分析** - 利用机器学习和数据可视化提供洞察

Watson Studio - Cognos Analytics - SPSS  
**Business Analytics**

**组织** - 创建可信赖数据分析基础

InfoSphere Datastage - Watson Knowledge Catalog - MDM -  
**Unified Governance & Integration**

**收集** - 数据访问更简单

Db2 - Hadoop - MongoDB - PostgreSQL  
**Hybrid Data Management**



实现平台简约和灵活

各种类型的数据，无论它在哪里

# Cloud Pak for Data v2.5 新的组合包

融合

Watson Openscale

分析

Watson Studio *includes Data Refinery*

Cognos Dashboards

Watson Machine Learning *includes Auto AI*

Analytics Engine

*- Powered by Apache Spark*

组织

Watson Knowledge Catalog

*includes IGC, Information Analyzer & Regulatory Accelerator*

Open Source Governance

收集

Data Virtualization

Db2 Eventstore

Performance Server

for PostgreSQL

Db2 Warehouse

IBM Streams

*- available only in Cloud Pak for Data system*

平台现代化

基础架构服务 (K8s & Common services)

UX, Authentication, Authorization, Monitoring, Metering, Logging

restricted use

# IBM 更多的扩展功能

## 应用及解决方案

通过开箱即用的解决方案和应用程序模块化加速AI的部署

- Watson Assistant
- Watson Discovery
- Watson Natural Language Understanding
- Watson Knowledge Studio
- Watson Explorer
- Watson Services
- Watson Compare and Comply

## 工具及模型部署

在流程和应用程序中构建、部署自定义模型

- Watson Studio
- Watson Machine Learning
- Watson OpenScale
- SPSS Modeler
- SPSS Statistics
- Decision Optimization

## 业务分析

可视化、理解并对基于企业数据采取对应措施

- Cognos Analytics
- Planning Analytics

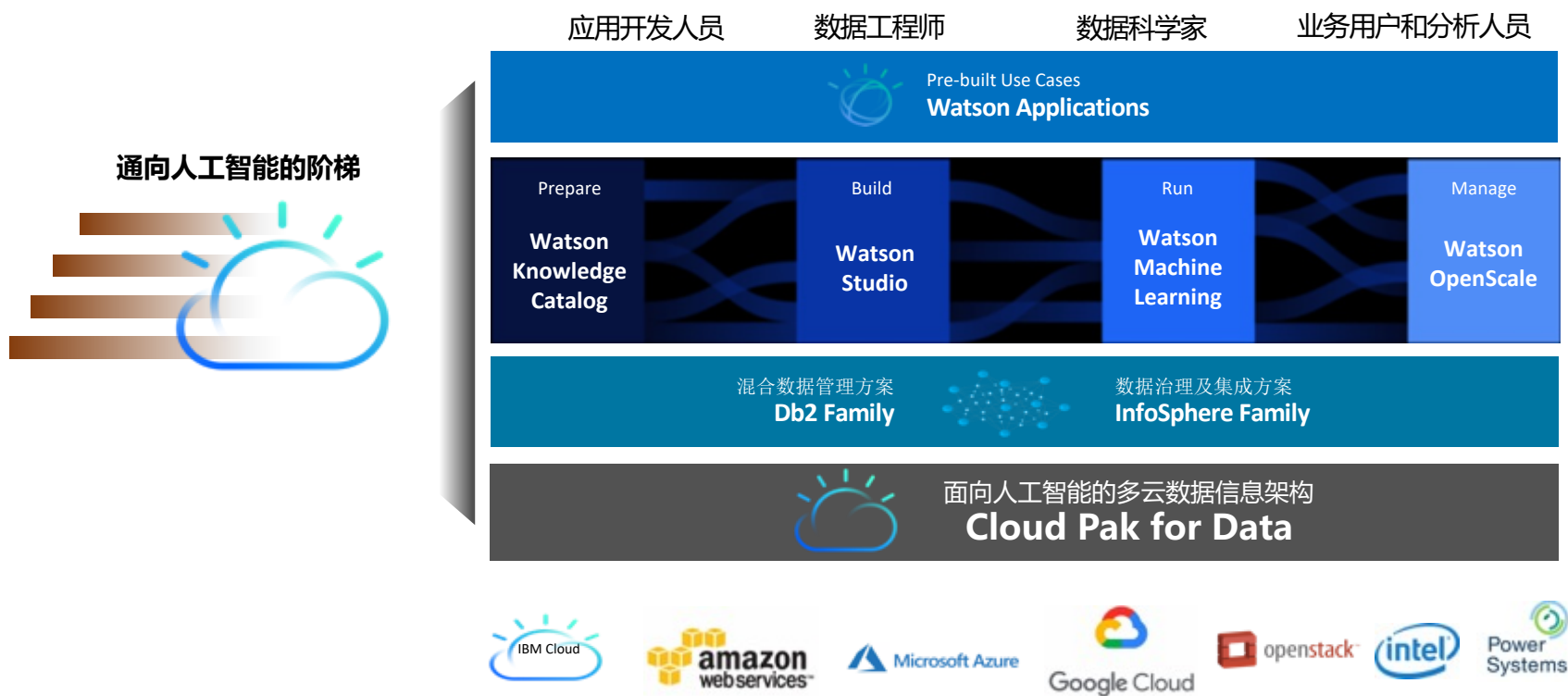
## IBM Cloud Pak for Data + Add-ons 插件

公有云，私有云，传统集群和单机



# IBM 人工智能场景服务组合

为您提供构建企业级人工智能所需的一切，并支持任何云环境

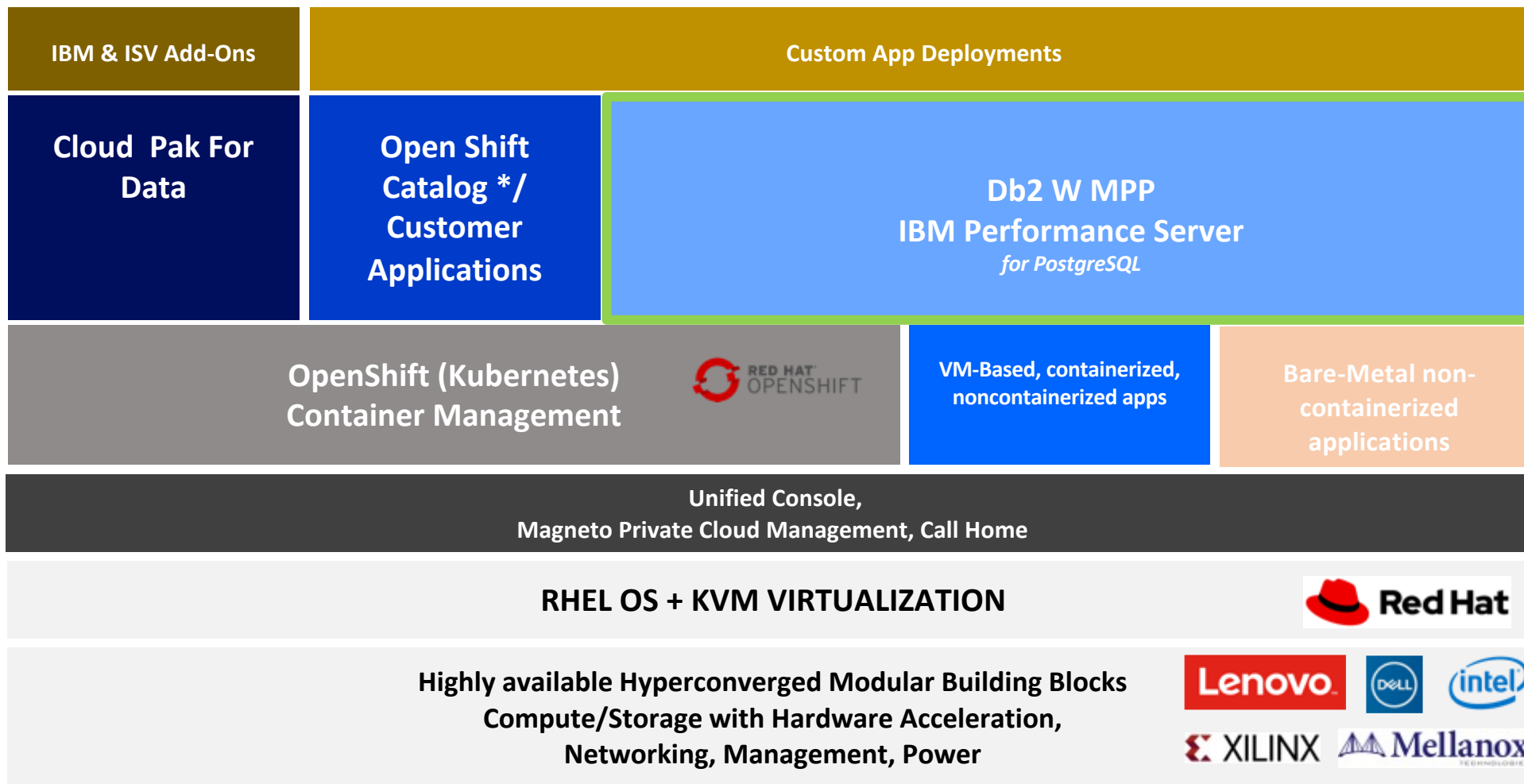


# Cloud Pak for Data System 数据及人工智能一体机



- 4小时内实现真正的即插即用企业数据和人工智能平台
- 云原生的设计
- 运行于Red Hat OpenShift上的Kubernetes集群
- 提供数据虚拟化服务
- 收集、治理、分析和融合数据
- 提供一系列的组件模块，实现快速搭建和扩展一个高性能的数据和AI的私有云环境

# Cloud Pak for Data System 数据及人工智能一体机



## 统一管理客户端

|                                      |  |  |  |
|--------------------------------------|--|--|--|
| Multi-Tenancy                        | Machine Learning for Performance and Capacity Optimization | Application-Specific Hardware Accelerators           | Service Catalog + CI/CD Upgrades                   |
| Magneto Hardware Monitoring + Alerts | Visual Order Configurator                                  | Intuitive System Performance & Troubleshooting Tools | Automatic Hardware Discovery and modular expansion |
|                                      | Call Home  |  |  |



## 系统管理软件



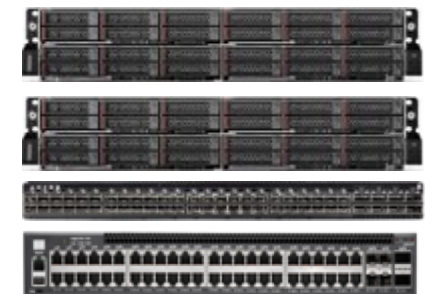
## 系统硬件



| 计算节点  | 网络                   | 存储                           | 机架 & 电源 (Optional) |
|---|----------------------|------------------------------|--------------------|
| Hyper Converged with Encrypted NVMe Storage + Hardware Acceleration | Fabric Leaf (10/25G) | Storage Gateway: SAN & Cloud | Pre-wired rack     |
| Storage Optimized   | Fabric Spine (100G)  | Object Store                 | Single Phase PDU   |
| Memory Optimized  | Management (1/10G)   | Tier 1 Flash                 | 3-Phase-WYE PDU    |
| GPU Optimized   |                      | Tier 2 Near Line             | 3-Phase-Delta PDU  |



Starter Bundle



# IBM Cloud Pak for Data 概览





# IBM Cloud Pak for Data 扩展组件

The screenshot displays the IBM Cloud Private for Data console interface. At the top, the header includes the text 'IBM Cloud Private for Data' and a search bar. The main content area is titled '附加组件' (Add-ons) and features a search bar with the text '查找附加组件'. A left-hand navigation menu lists various categories: 'All categories', 'AI', 'Industry accelerators', 'Storage', '仪表板' (Dashboards), '分析' (Analysis), '开发人员工具' (Developer tools), '数据源' (Data sources), and '数据监管' (Data governance). The 'AI' category is selected, and the main area displays six AI extension components in a grid:

- Watson OpenScale**: Includes an icon of a scale, 'IBM' and 'Premium' tags, and the description: 'Infuse your AI with trust and transparency. Understand how your AI models make decisions to detect and mitigate bias.'
- Watson Assistant - wa-cloudpak**: Includes a speech bubble icon with a question mark, 'IBM' and 'Enabled' tags, and a 'Premium' tag. Description: 'Build conversational interfaces into any app, device, or channel.'
- Watson Assistant**: Includes a speech bubble icon with a question mark, 'IBM' and 'Enabled' tags, and a 'Premium' tag. Description: 'Build conversational interfaces into any app, device, or channel.'
- Watson Discovery**: Includes an icon of a document with a magnifying glass, 'IBM' and 'Enabled' tags, and a 'Premium' tag. Description: 'Find answers and uncover insights in your complex business content.'
- Watson Speech to Text**: Includes an icon of a document with a speech bubble, 'IBM' and 'Enabled' tags, and a 'Premium' tag. Description: 'Convert audio and voice into written text.'
- Watson Text To Speech**: Includes an icon of a document with a speech bubble, 'IBM' and 'Enabled' tags, and a 'Premium' tag. Description: 'Convert written text into natural-sounding speech.'

# IBM Cloud Pak for Data 主要的网上资料

Cloud Pak for Data Knowledge Center 每个季度都有更新

[https://www.ibm.com/support/knowledgecenter/zh/SSQNUZ\\_2.1.0/com.ibm.icp.data.doc/zen/overview/overview.html](https://www.ibm.com/support/knowledgecenter/zh/SSQNUZ_2.1.0/com.ibm.icp.data.doc/zen/overview/overview.html)

Cloud Pak for Data 的网上视频及试用版本

<https://dsgapp.cn.edst.ibm.com/ibm-dte/cloud-private-for-data/index.html>

**IBM Cloud Pak for Data**  
IBM Cloud Pak for Data 是一个统一和简化数据收集、组织和分析的数据平台。

IBM Cloud Private™ for Data 是一种新型数据和分析平台，可简化和统一您收集、组织和分析数据的方式，从而加速数据科学和 AI 的价值实现。该平台提供广泛的核心数据微服务，可以选择从不增长的服务目录中添加其他服务。而且无需移动数据。通过这种受管理的多云平台，利用您所需的安全性和控制，提供所需的灵活性，同时保持云的优势，同时保持数据的位置。

[产品之旅](#) ↓  
[动手实验室](#) ↓

**IBM Cloud Pak for Data 概述**  
⌚ 5分钟的演示  
了解 IBM Cloud Pak for Data 的价值，了解它在产品生态系统中的适用位置以及最引人注目的用例。

**IBM Cloud Pak for Data - 数据虚拟化简介**  
⌚ 9分钟的演示  
介绍数据虚拟化的概念，并简要介绍如何设置虚拟表并加入 ICP for Data。

**快速产品演练**  
⌚ 4分钟的演示  
本期简报介绍 IBM Cloud Pak for Data 的端到端之旅。

# 进一步了解 IBM Cloud Pak for Data IBM 业务代表为您答疑解惑（工作日 9:00-17:00）



400-669-0260



[请填写联系信息表](#)



关注 IBM Cloud 官方微信号，及时获取 IBM Cloud 最新资讯