

# ALTAIR RAPIDMINER

---

Data Analytics & AI Platform

## Altair-at-a-Glance

**\$666M**

FY24 Revenue

**81 Offices**

In 29 Countries

**3,000+**

Engineers, Scientists,  
and Creative Thinkers

**180+**

Altair and Partner  
Software Products

**16,000+**

Customers Globally

# Data Science + Rocket Science™ =



Altair One

一站式云端门户

**ALTAIR**  
HYPERWORKS

设计和仿真平台

**ALTAIR**  
HPCWORKS

高性能计算和云平台

**ALTAIR**  
RAPIDMINER

数据分析与 AI 平台

# RapidMiner——为所有人准备的机器学习平台

© Altair Engineering, Inc. Proprietary and Confidential. All rights reserved.

所有人都可以在RapidMiner平台上发挥自己的才能，以您想要的方式进行模型的构建



## 普通用户：可视化拖拽 workflow

简单，易用的界面建模方式。自动数据清洗和建模功能，新手也能快速构建机器学习模型



## 数据分析师：增强数据科学

超过200个功能模块，167个模型算子，支持数据分析师构建复杂的模型流程



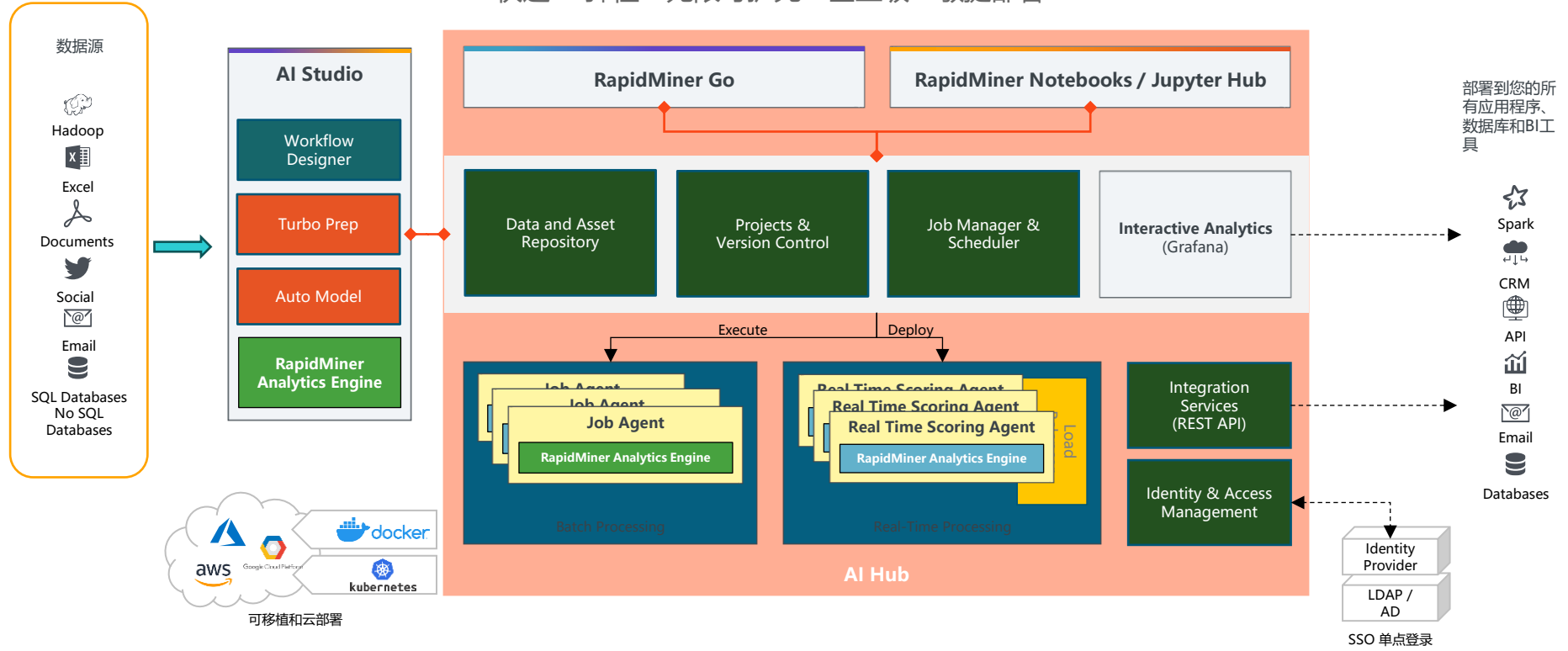
## 算法专家：Notebooks

算法专家可以在平台上编写代码，创建全定制化的模型

Survived	Passenger...	Name	Sex	Age	No. of Siblings...	No. of Par...	Ticket No...	Passenger...	Cabin
Yes	First	Anderson, ...	Male	41.000	0.000	0.000	19512	26.500	B12
No	First	Andrews, ...	Female	63.000	1.000	0.000	13052	77.950	D7
No	First	Andrews, ...	Male	39.000	0.000	0.000	110500	0.000	A36
Yes	First	Appleton, ...	Female	53.000	2.000	0.000	117369	51.479	C101
No	First	Arigunawati, ...	Male	71.000	0.000	0.000	PC 17609	49.504	7
No	First	Astor, Carl	Male	47.000	1.000	0.000	PC 17357	27.525	524 C84

# RapidMiner 数据分析平台

快速 + 弹性 · 无限可扩充 · 企业级 · 敏捷部署



部署到您的所有应用程序、数据库和BI工具

- Spark
- CRM
- API
- BI
- Email
- Databases

Identity Provider  
LDAP / AD  
SSO 单点登录



# 零代码,端到端的机器学习平台

**数据工程师**  
加速企业数据分析



**数据科学家**  
提高工作效率和协作



**行业专家**  
确保和商业目的一致



**业务分析师**  
简单易用的机器学习



**CXO**  
发挥竞争优势



自动化/流程化/编程化的 高度可视可控 整合数据分析项目平台

连接所有数据,不论来源  
种类存储方式



**数据整理**



**机器学习**



**模型部署**

整合与所有  
应用程序、  
数据库和  
任何BI 工具

## Studio

灵活、可视化的拖放设计器,具有自动化功能

## Go

全自动原型设计及数据探索

## Notebooks

以代码为中心的方法,实现完全自定义的特性化

## AI Hub

协作、自动化决策和控制

# 企业知识图谱和数据结构

通过将孤立的系统数据集集成到共享知识图谱中来激活您的企业数据，在相关但断联的来源之间创建新的关系，并提供通用语义以使每个人都在同一页面上。

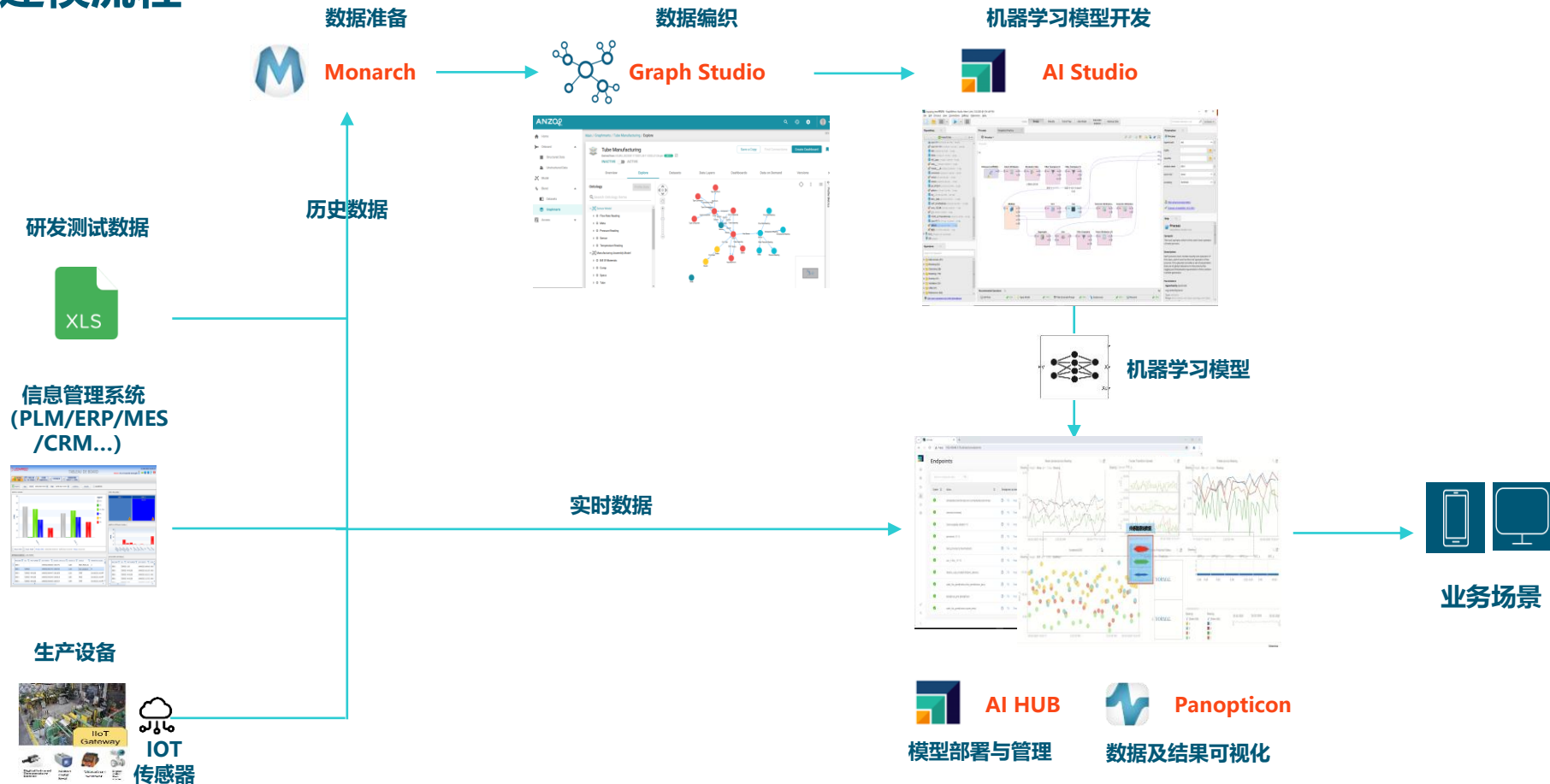


## 关键性能

- 1 适用于复杂、多样化的数据生态系统的叠加
- 2 简化访问并消除孤岛
- 3 针对新数据来源进行灵活、“动态”数据集成
- 4 语义层，为每个人提供对数据的共同理解
- 5 大规模并行图形引擎提供极大的规模和性能
- 6 强大的数据治理

# 建模流程

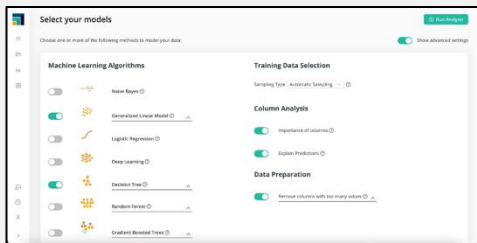
© Altair Engineering, Inc. Proprietary and Confidential. All rights reserved.



# 一个面向所有人的平台

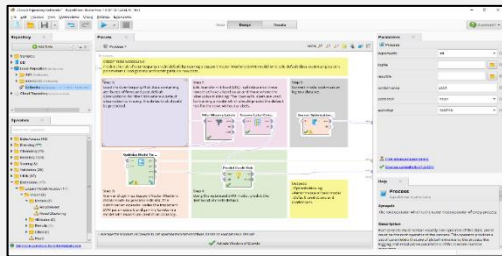
## 自动化数据科学

引导式交互体验 自动执行数据科学与机器学习与最佳实践



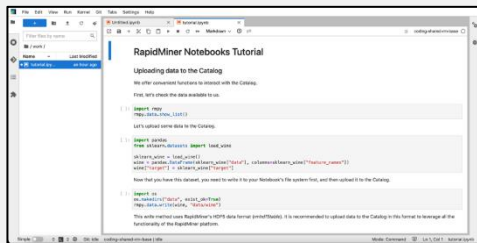
## 可视化 workflow

可视化设计数据流和DS/ML功能流



## 代码

开发代码执行DS/ML任务的代码



业务专家

易用性  
广度/深度  
可解释性



公民数据科学家

易用性  
广度/深度  
可解释性



专业数据科学家

易用性  
广度/深度  
可解释性



面向所有人的AI平台

# Altair Monarch

## 自助式数据准备工具



### 连接数据

- 常用关系型数据库连接: Oracle、MySQL等
- 连接大数据, 如: Hive、Spark SQL、AWS Redshift等
- 半结构化PDF数据抽取



### 简单易用

- 全可视化界面数据准备工具
- 80+ 预置数据准备功能
- 无需编写脚本或宏



### 数据导出

- 可以导出到任何BI和数据分析平台
- 过程可保存, 结果可审计



### 自动化

- 周期性、重复性的数据抽取工作
- 可视化界面设置工作流程
- 创建管理任务, 例如邮件分发、数据归档、通知发送

# Altair Monarch——激活和自动化“捕获”数据

使用专利工具提取报告（包括文本报告、企业业务报告、PDF 和高度格式化的 Excel 电子表格）中的数据，消除笨拙的手动数据输入并自动化流程。

## 关键性能

### 从企业报告中提取数据

将 ERP、MES、EMR 数据转换为行和列的专利技术

### 从 PDF、Excel 中提取数据

将 PDF 和高度格式化的 Excel 电子表格中的数据转换为行和列的专利技术

### 强大的后处理

超过 80 种单独的数据准备、转换和清理功能

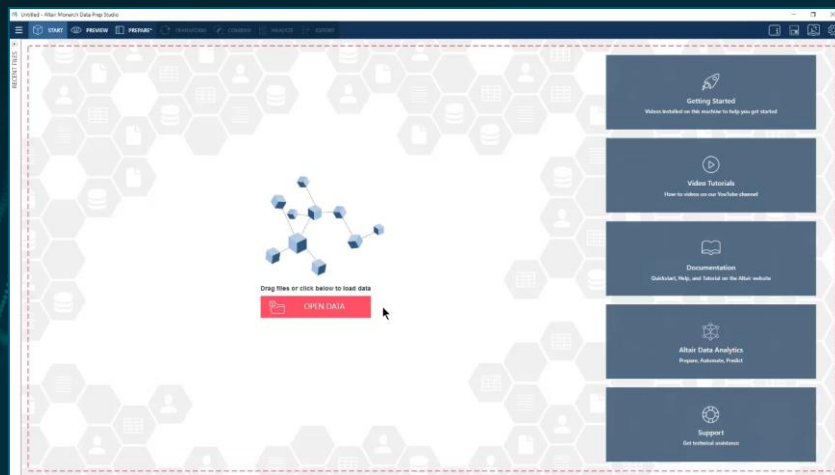
### 可重复、可靠、准确

经过企业测试和30年的开发，我们提供 100% 准确的提取

## 多种多样的格式



## 用户友好的界面



# 面向企业的数据抓取，数据预处理平台

© Altair Engineering, Inc. Proprietary and Confidential. All rights reserved.

- 支持从多源异构中提取数据，并转化为结构化数据，再进行相关的数据处理。

## 自助式数据提取， 全面的数据接口



## 自助式数据预处理



## 原文件预览与类正则表达式的捕获设置

The screenshot displays the Monarch PDF software interface for configuring a report design. The main window is titled "Report Design - Expense Report, 2017-07.pdf". The interface is divided into several panels:

- Templates Panel:** Lists templates with columns for Name, Role, and a delete icon. The "Append0" template is selected.
- Edit Field Properties Panel:** Shows configuration for the selected template, including Name (Client), Template Width (23), and Type (Character).
- Document Preview Panel:** Shows a preview of the PDF content with highlighted text. The text includes "Client: Centene", "Datawatch Corporation", "4 Crosby Dr.", "Bedford MA 01730", "Report Date 01-Aug-17", "Report Number 384", and "Client: Centene", "7711 Carondelet Ave.", "Clayton, MO 63105".
- Table Panel:** Shows a table with columns "Ab Client", "Ab Name", and "Ab Department". The table contains 6 rows of data.

Navigation buttons at the top include ACCEPT, CANCEL, TRAPPING, REPORT VERIFY, MULTI-COLUMN REG..., and DOCUMENT OPTIONS. A search bar and page navigation controls are also visible.

针对不同格式，不同内容的捕获细节设定

输出结果文件预览

# Monarch 数据处理功能

© Altair Engineering, Inc. Proprietary and Confidential. All rights reserved.

提供了变量级和文件级的数据预处理和文件导出

记录每一次数据处理步骤

The screenshot displays the Altair Monarch Data Prep Studio interface. The top navigation bar includes buttons for START, PREVIEW, PREPARE\*, TRANSFORM, COMBINE, ANALYZE, and EXPORT. The main workspace shows a data table with columns for various categories and values. A context menu is open over the table, listing actions such as Copy, Filter, Sort, and Delete Column. On the left, a 'TABLES' panel shows 'Expense Report, 2017-07' and 'Employee' sources. On the right, a 'CHANGE HISTORY' panel lists seven steps of data processing, including column moves, deletions, and renames. A red arrow points from the 'CHANGE HISTORY' panel to the 'ANALYZE' button in the top bar.

将不同的数据源导入到同一平台

记录每一次数据处理步骤

Row	eg_110_05	Ab	eg_USD	Ab	eg_Telephone - Mobile	Ab	eg_NA Sales Engineering	Ab
1	110.05	USD	Telephone - Mobile	NA Sales Engineering	DW			
2	127.49	USD	Internet Services	NA Sales Engineering	DW			
3	106.34	USD	Telephone - Mobile	NA Sales Engineering	DW			
4	133.51	USD	Taxi/Bus/Train	NA Sales Engineering	DW			
5	141.09	USD	Taxi/Bus	Engineering	DW			
6	29.23	USD	Airline	Engineering	DW			
7	263.46	USD	Taxi/Bus	Engineering	DW			
8	49.96	USD	Lunch	Engineering	DW			
9	235.68	USD	Internet	Engineering	DW			
10	249.30	USD	Hotel	Engineering	DW			
11	81.26	USD	Dinner	Engineering	DW			
12	36.18	USD	Breakfast	Engineering	DW			
13	66.82	GBP	Taxi/Bus	ulting	DW			
14	67.63	GBP	Hotel	ulting	DW			
15	66.78	GBP	Taxi/Bus	ulting	DW			
16	68.60	GBP	Taxi/Bus	ulting	DW			
17	27.07	GBP	Dinner	ulting	DW			
18	21.94	GBP	Parking	ulting	DW			
19	13.83	GBP	Lunch	ulting	DW			
20	28.97	GBP	Dinner	ulting	DW			
21	64.88	GBP	Taxi/Bus	ulting	DW			
22	7.18	GBP	Breakfast	ulting	DW			
23	60.97	GBP	Hotel	ulting	DW			
24	61.96	GBP	Taxi/Bus	ulting	DW			
25	27.93	GBP	Dinner	ulting	DW			
26	60.37	GBP	Taxi/Bus	ulting	DW			
27	20.52	GBP	Dinner	ulting	DW			
28	61.93	GBP	Taxi/Bus/Train	UK Consulting	DW			
29	177.06	USD	Dinner - Employee	NA Support	DW			
30	68.62	USD	Telephone - Mobile	NA Consulting	DW			

The screenshot displays the Altair Monarch Data Prep Studio interface. At the top, a dark blue navigation bar contains icons for 'START', 'PREVIEW', 'PREPARE\*', 'TRANSFORM', 'COMBINE', 'ANALYZE', and 'EXPORT'. Below this, a light gray toolbar offers 'Join Tables' and 'Append Tables' options. The main workspace is divided into three sections: a left sidebar for table selection, a central visualization area, and a bottom data preview table.

**Join Configuration:** The central area shows a 'Join' node connected to two source tables: 'Expense Report, 2017-07.pdf' and 'Employee'. A dashed box labeled 'DRAG A TABLE' is positioned to the right of the join node.

**Data Preview:** The bottom table displays the following data:

	Date	Ab	eg_Sales	#	eg_110_05	Ab	eg_USD	Ab	eg_Telephone - Mobile
1	2002/7/17	Sales			110.05	USD			Telephone - Mobile
2	2002/7/17	Sales			127.49	USD			Internet Service
3	2002/7/17	Sales			106.34	USD			Telephone - Mobile
4	2002/7/17	Sales			133.51	USD			Taxi/Bus/Train
5	2002/7/17	Sales			141.09	USD			Taxi/Bus/Train
6	2002/7/17	Sales			29.23	USD			Airline Fees
7	2002/7/17	Sales			263.46	USD			Taxi/Bus/Train
8	2002/7/17	Sales			49.96	USD			Lunch - Employee
9	2002/7/17	Sales			235.68	USD			Internet Service
10	2002/7/17	Sales			249.30	USD			Hotel
11	2002/7/17	Sales			81.26	USD			Dinner - Employee
12	2002/7/17	Sales			36.18	USD			Breakfast - Employee
13	2009/7/17	Operations			66.82	GBP			Taxi/Bus/Train

## 可视化的数据表连接配置

## 结果数据预览

# Altair® AI Studio

## 预测性分析和数据科学平台



### 数据连接

支持连接到任何数据源，如  
TXT、XLS、Hadoop；  
内置数据准备操作，方便用户  
对数据进行预处理；



### 可视化建模

自动数据准备  
自动机器学习  
全流程可视化建模



### 大数据支持

连接Hadoop集群，进行大规模  
模型运算



### 开放式部署

支持PMML模型导出部署  
RapidMiner AI Hub端部署，  
提供Web API服务





# RapidMiner AI Hub

## 模型共享和协作

团队成员可以轻松共享、重用和协作开发机器学习模型。可以跟踪和管理不同版本的模型，轻松切换和回滚到不同的模型版本

## 安全和权限管理

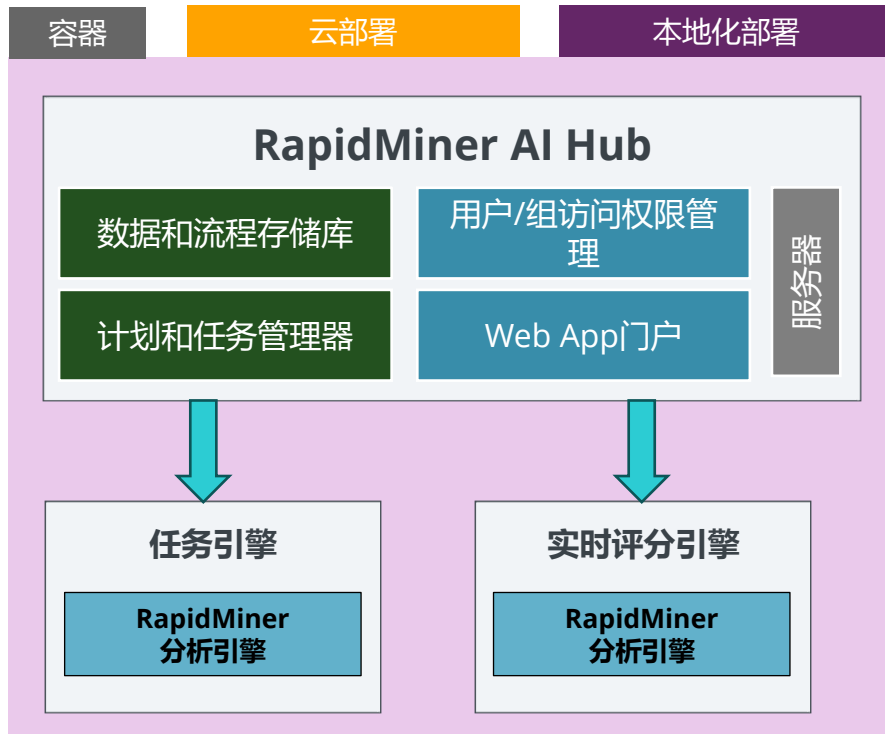
具有严格的安全和权限管理功能，确保只有授权用户能够访问和使用模型。用户可以定义和管理团队成员的访问权限，并保护模型和数据的安全性

## 计划 & 任务管理器

长时间运行的作业可以在服务器硬件上更快更有效地分派。工作任务可以按计划时间运行，也可以设置为由外部事件触发

## 模型部署

将模型以API或服务化的形式部署。通过API，其他系统和应用程序可以通过请求和响应的方式与模型进行交互，实现实时的决策和预测



# Altair Graph Studio

## 数据编织平台



### 数据集成

支持结构化数据源、非结构化  
文档数据集成



### 大数据存储

分布式集群的图数据库  
数据自动转为知识图谱  
数据虚拟化技术



### 数据查询分析

高性能内存计算引擎  
通用OLAP查询接口  
SPARQL查询接口  
可视化分析报表

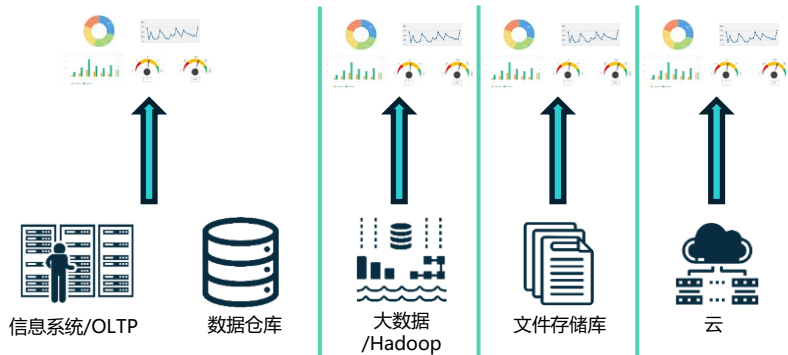


### 大模型

支持对话式数据查询  
支持嵌入行业知识

# 传统方法 VS 数据编织

## 传统方法



### 孤立的数据集

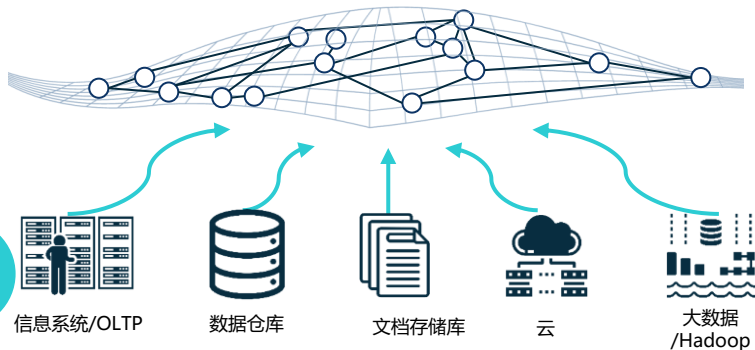
- 独立管理的数据“孤岛”，交互很少
- 数据源的增加增加了分析的复杂性和利用率
- 应用级数据管理

### 优势

回答重复的单一来源问题

VS

## 数据编织



### 一个有凝聚力的数据层，多个来源

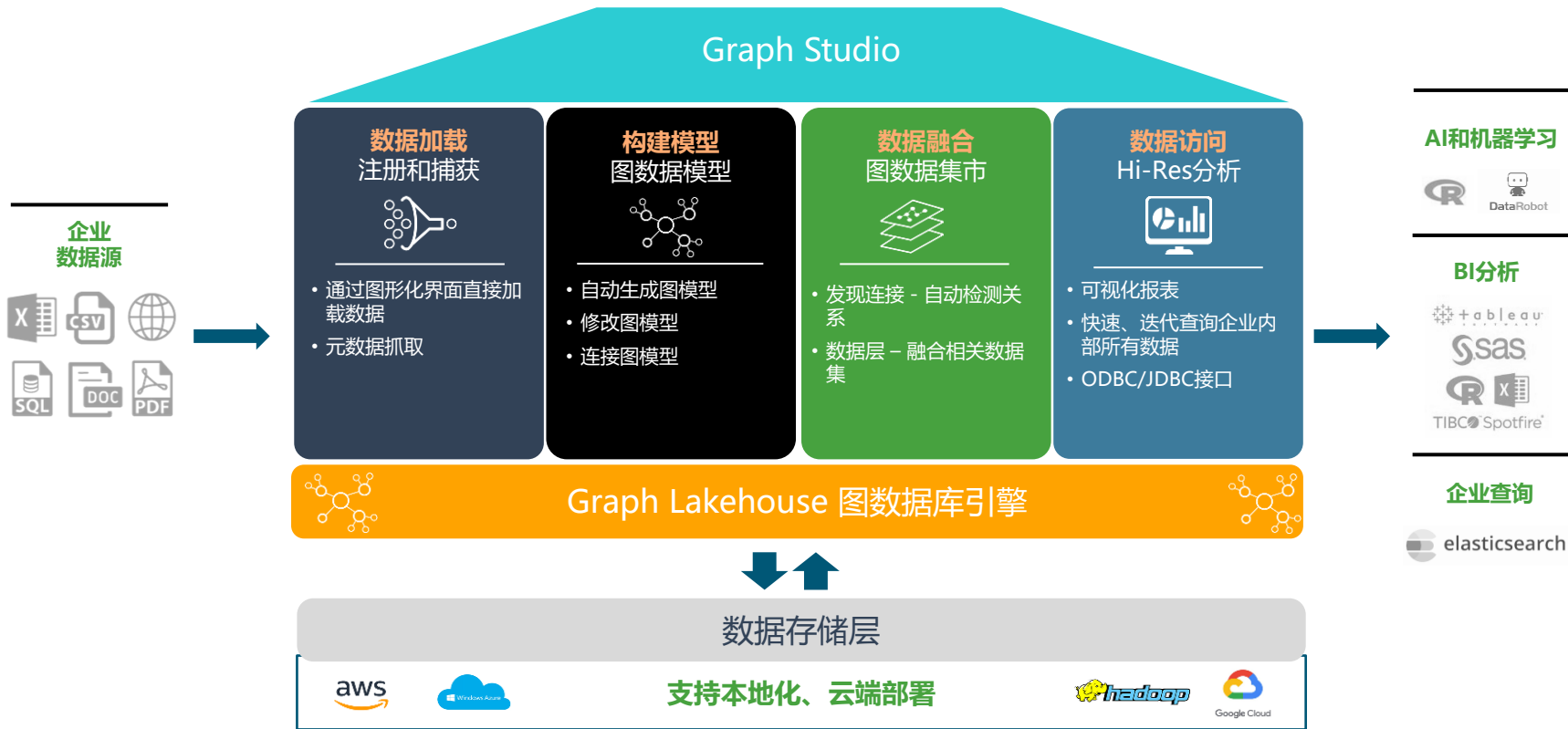
- 数据集成，将公司各种数据编织在一起
- 灵活的结构可实现“即时”数据集成
- 以知识图谱的方式呈现数据

### 优势

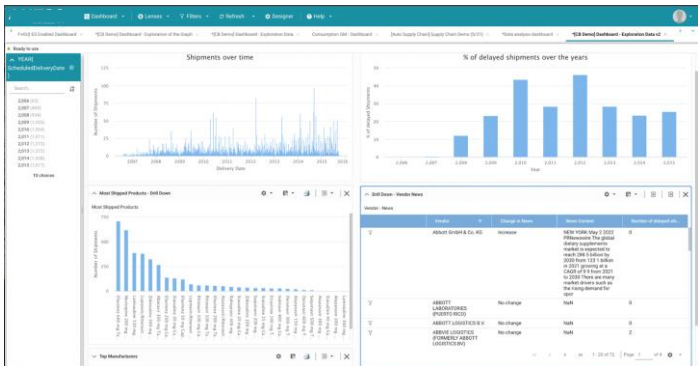
快速、大规模地回答临时的多源问题

# Graph Studio架构图

© Altair Engineering, Inc. Proprietary and Confidential. All rights reserved.



# 数据查询&可视化



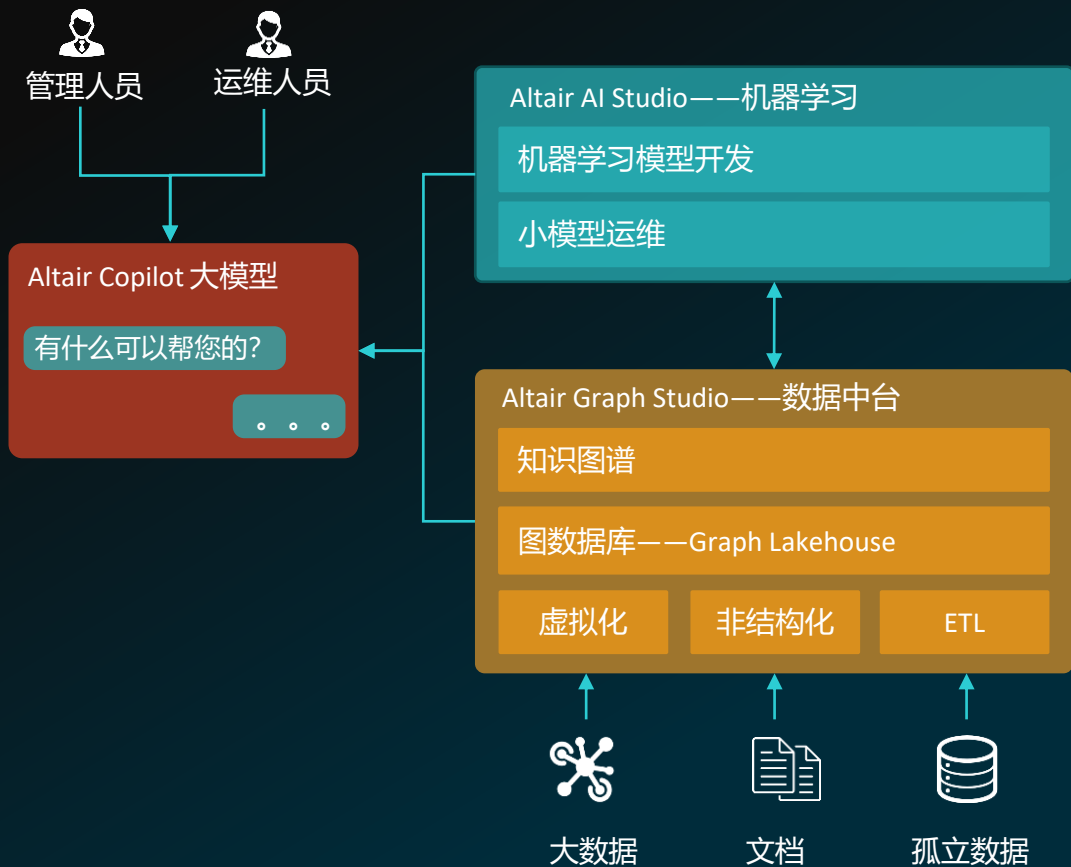
- **Hi-Res Analytics Dashboards** 用户能够轻松的基于知识图谱分析数据和构建可视化看板

- **Query Builder:** 使熟悉SPARQL的技术用户能够直接查询数据

AI、BI工具，如Altair RapidMiner、Power BI等  
OData访问协议

- ODBC
- JDBC

# Altair Copilot 在企业的应用



## Altair RapidMiner 带来的益处

1. 每位业务人员都有自己的个人AI Copilot
2. 灵活的工作流程，可以快速迭代，不受特定大型语言模型（LLM）的限制
3. 知识图谱为会话分析提供了商业背景和确定性
4. 激活组织内未充分利用的结构化和非结构化数据
5. 虚拟化技术使数据保留在原系统，降低集成成本

## 实时流数据可视化



### 数据访问

流式数据连接,  
如: Kafka、MQTT等  
时序数据库连接,  
如: KDB+、Influx DB  
关系型数据库



### 流数据处理

计算、  
聚合、  
过滤、  
异常检测、  
等等...



### 可视化界面

内置多种图表选择,  
支持快速构建可视化分析;  
使用浏览器构建可视化仪表盘  
无需任何编码;

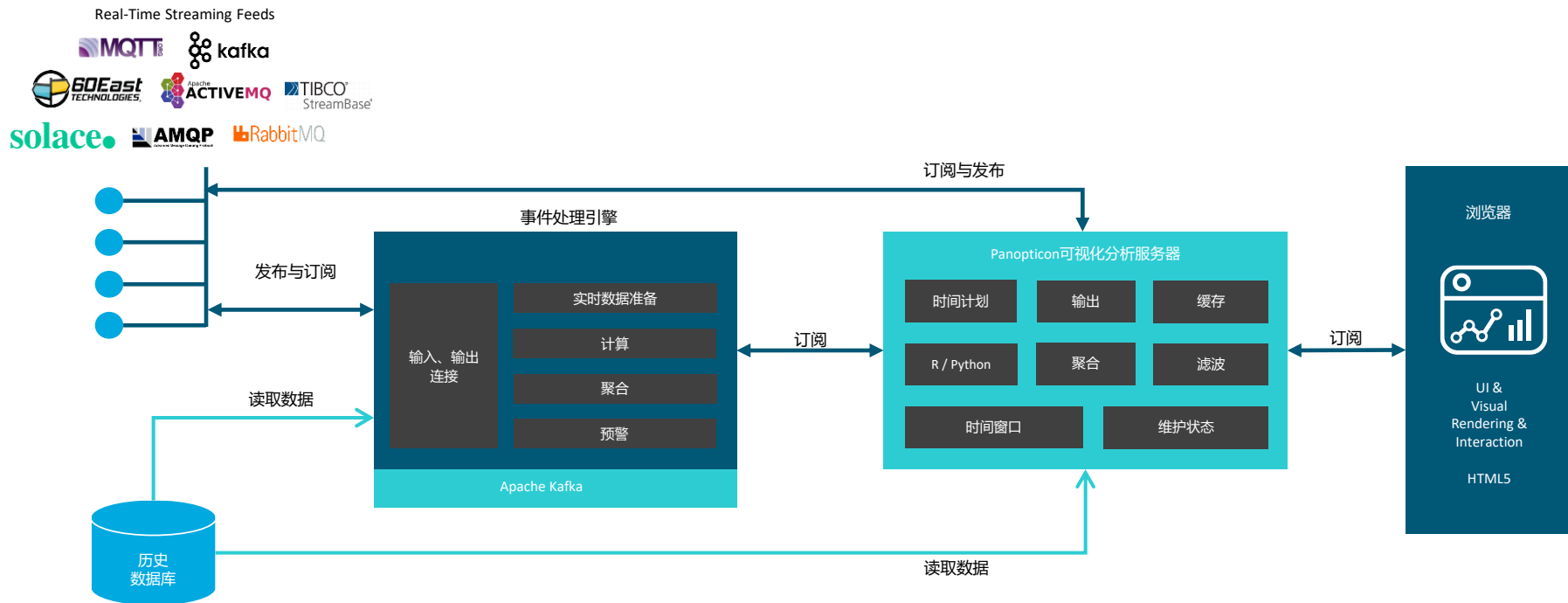


### 可视化部署

无需本地安装软件;  
简单点击, 即可发布仪表盘;  
支持浏览器访问可视化仪表盘;

# Panopticon 流数据分析平台: 底层架构

© Altair Engineering, Inc. Proprietary and Confidential. All rights reserved.



MEMSQL kx ONETICK

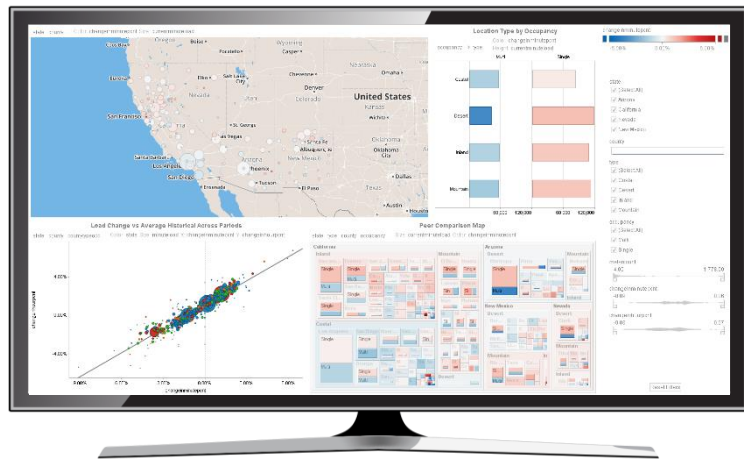
mongoDB influxdb

ORACLE IBM DB2 cassandra elastic



## 物流车辆实时运行状况监控

实时监控车辆运行情况，包括车辆信息，例如：GPS位置，车速，车辆运行时间等。同时还能监控驾驶员的驾驶行为，例如是否系安全带，急刹车数量等



## IoT 智能电表监控与分析

企业能够监视和分析来自工业和商业应用中的各种传感器（在本例中为电表表）的实时数据流。用户可以使用可视化分析显示来发现趋势，并找出异常值

# 案例分享

# AI X 仿真

## Problem

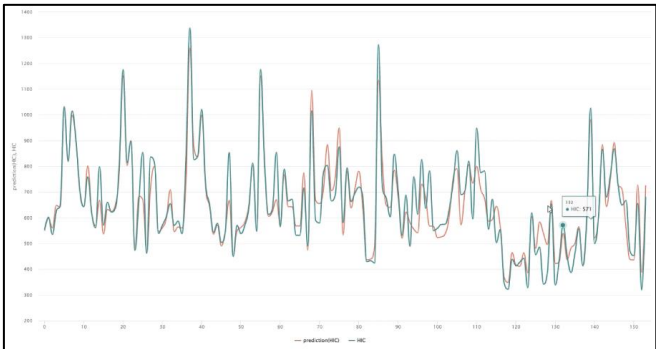
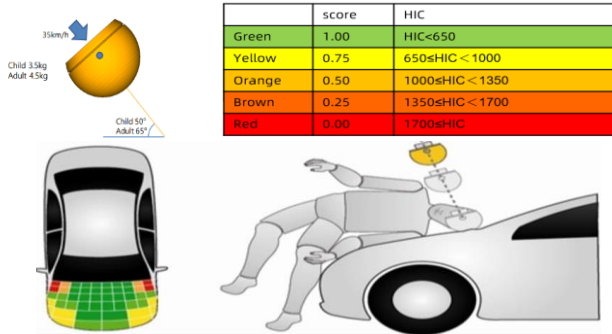
- 在汽车行业，行人保护至关重要，尤其是车辆与行人碰撞时。头部伤害指标（HIC）是衡量行人头部碰撞伤害程度的关键。然而，目前HIC值的预测主要依赖耗时耗力的物理试验和仿真方法，且难以全面考虑实际碰撞场景中的变量。所以汽车行业急需高效、准确的方法来预测行人头部碰撞时的HIC值，以优化车辆结构，提高行人保护性能。

## Solution

- 收集多辆车型的仿真数据并进行二次开发做数据的特征提取，提取表征碰撞情境的特征，如引擎盖材料和形状的参数，撞击器的参数等。
- 模型构建与训练：利用RapidMiner的机器学习算法库，基于提取的特征构建HIC值预测模型，并进行训练和验证，后续有了新的设计，直接提取参数后用模型进行预测。

## Benefit

- 利用RapidMiner开发的行人头部碰撞HIC值预测系统，能够显著提高预测准确性，缩短车辆研发周期并降低成本。该系统使车辆设计师能够更有针对性地优化车辆结构，提升行人保护性能，减少碰撞事故中的行人伤害。



# AI X 测试

## Problem

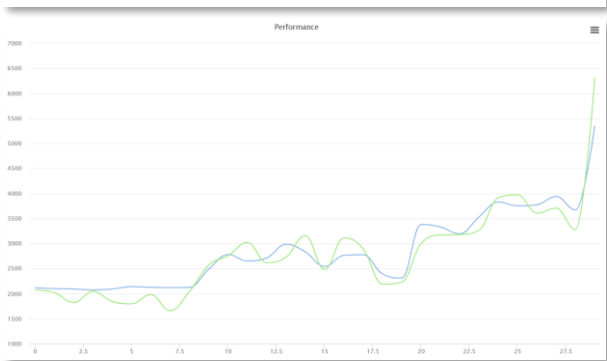
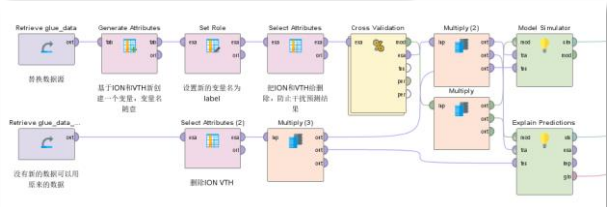
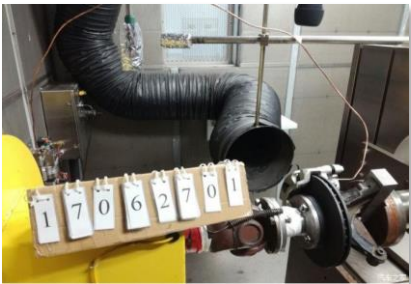
- 刹车片制造商在投入市场前需进行全面的性能测试，确保其符合安全和效能标准。磨损值是关键性能指标之一，测试涉及多种路况、刹车次数、温度和速度，实验耗时且成本高。缺乏刹车片配方数据也影响了磨损预测的准确性。

## Solution

- 基于用户历史上大量的刹车片产品和磨损测试数据，基于RapidMiner AI Studio构建了多个预测模型，包括直接预测某种路况下的磨损值，以及使用部分路况的磨损来预测整体磨损情况。随着配方数据的纳入，预期将进一步提升模型的预测精度。

## Benefit

- 我们的模型在预测刹车片整体磨损的结果上，**偏差在3%左右**，能够帮助用户**减少约66%的实验测试时间**。随着配方数据的完善，预期将大幅度减少非必要的实验次数和时间，降本增效，提升产品竞争力。



# AI X 材料研发

## Problem

- 粘合剂生产商面临配方开发的复杂性、实验的不确定性以及严格的性能测试要求等核心痛点。传统试错法效率低下，难以同时满足高性能粘合剂的多项严苛标准，且实验结果可能因操作或环境差异而无法重现，导致开发周期长、成本高。

## Solution

- 为解决这些痛点，企业转向利用AI技术，通过RapidMiner AI Studio的模拟功能进行高性能粘合剂的模拟实验。收集历史数据，建立预测模型，利用机器学习算法优化迭代，快速筛选和优化配方，同时处理多目标优化问题，找到满足所有性能要求的最优配比方案。

## Benefit

- 采用AI技术后，企业显著提高了研发效率，降低了成本，加速了产品上市。通过智能算法快速定位到最佳配比，解决了传统试错法的局限性，确保了高性能粘合剂的质量稳定性和市场竞争力，为不同行业的制造领域提供了可靠的粘合剂解决方案。

