

Put your SAP Data to Databricks Lakehouse

SAP 데이터 마이그레이션 및
AI/ML 분석 활용 사례

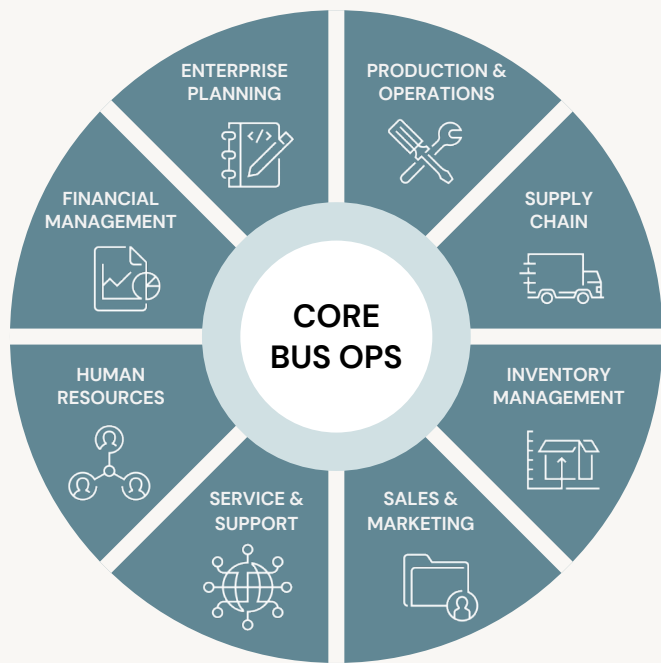
Databricks, Partner Solutions Architect

신효정

©2023 Databricks Inc. — All rights reserved



SAP 데이터로 핵심 비즈니스 운영 프로세스 강화

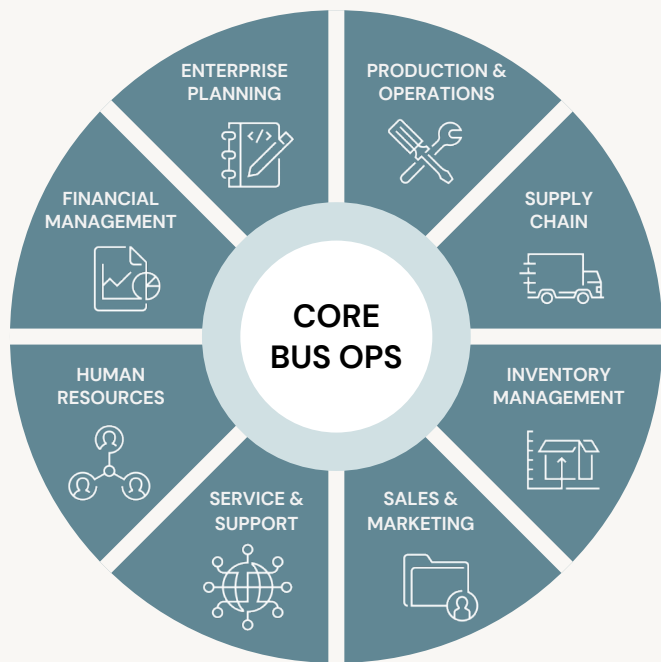


핵심적인 비즈니스 시스템의 데이터는 미션 크리티컬합니다. SAP는 조직 전체에 걸쳐 있으며 비즈니스 운영을 위한 데이터를 효율적으로 처리하고 정보 흐름을 원활하게 합니다

SAP 데이터는 공급망, 제조, 조달, 금융 및 HR 전반에 걸쳐 고급 분석 사용 사례를 구현할 때 매우 중요하게 사용 됩니다.

SAP 데이터를 쉽고 빠르게 통합할 수 있는 기능이 없다면 비즈니스 프로세스를 처음부터 끝까지 최적화하거나 의사 결정을 개선하는데 어려움이 있습니다.

SAP 데이터로 핵심 비즈니스 운영 프로세스 강화

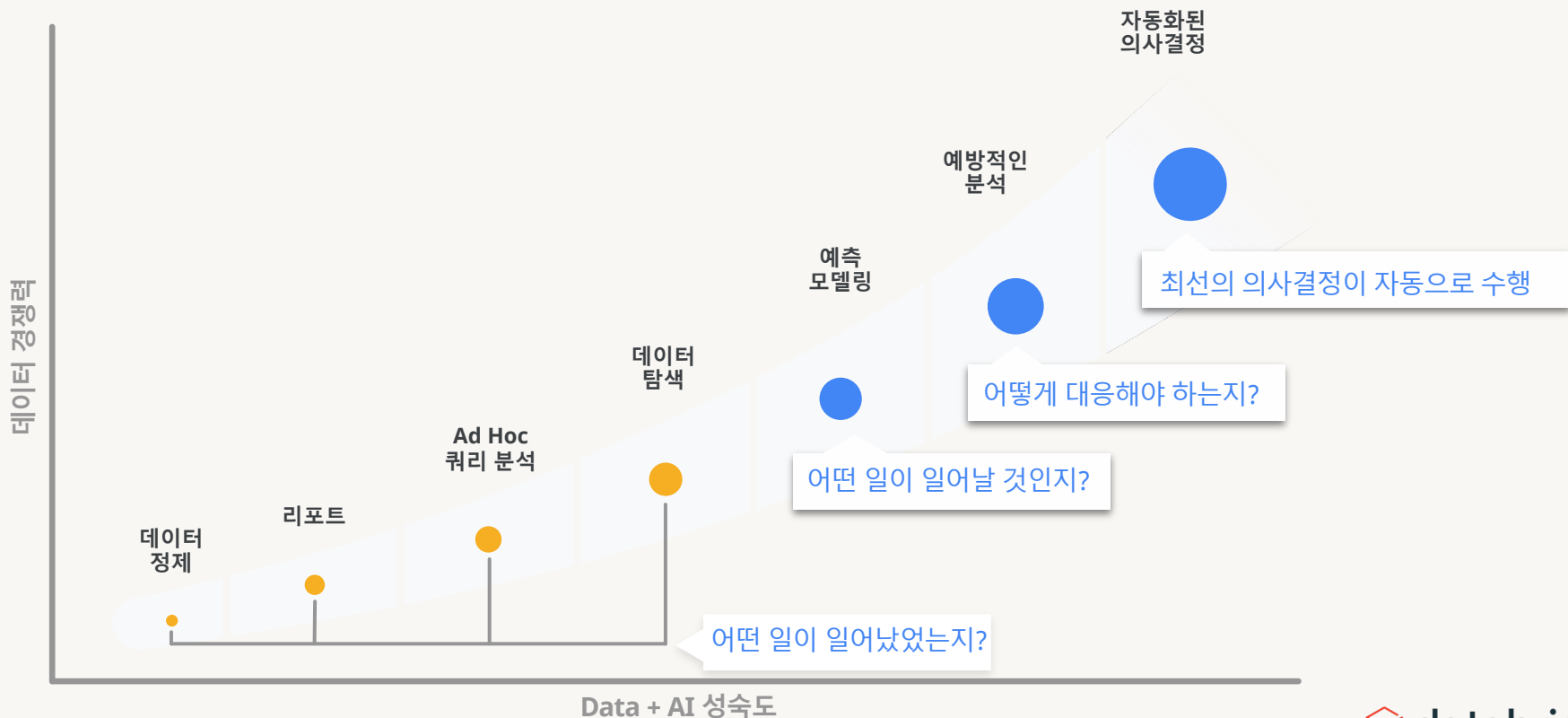


핵심적인 비즈니스 시스템의 데이터는 미션 크리티컬합니다. SAP는 조직 전체에 걸쳐 있으며 비즈니스 운영을 위한 데이터를 효율적으로 처리하고 정보 흐름을 원활하게 합니다

SAP 데이터는 공급망, 제조, 조달, 금융 및 HR 전반에 걸쳐 **고급 분석 사용 사례**를 구현할 때 매우 중요하게 사용 됩니다.

SAP 데이터를 쉽고 빠르게 통합할 수 있는 기능이 없다면 비즈니스 프로세스를 처음부터 끝까지 최적화하거나 **의사 결정을 개선**하는데 어려움이 있습니다.

데이터 성숙도 커브



오늘날, 너무 많은 플랫폼을 병합해 사용하고 있습니다.

모두 불필요한 비용이 발생하고 복잡성이 증가합니다.

데이터
레이크

데이터
웨어하우스

오케스트레이션

비즈니스
인텔리전스

데이터 사이언스
및 ML

스트리밍

거버넌스

데이터 사일로로 인한
높은 운영 비용

일관성 없는 정책으로
데이터의 신뢰도 저하

서로 다른 도구 사용으로
팀 간 생산성 저하



전혀 다른 접근 방식의 데이터 레이크하우스

단일 플랫폼으로 다양한 사용자 유형 지원



BI 및 데이터
웨어하우징



데이터
엔지니어링



데이터
스트리밍



데이터
사이언스 및 ML

조직 전체의 모든 데이터에 액세스할 수 있는
단 하나의 보안 및 거버넌스 모델

모든 정형, 반정형 및 비정형 데이터를 저장하고
관리하는 단일 플랫폼



클라우드 데이터 레이크

모든 원시 데이터
(로그, 텍스트, 오디오, 비디오, 이미지)

유일한 통합 레이크하우스를 제공하는 Databricks

단일 통합 플랫폼으로 다양한 사용자 유형 지원



Databricks
SQL



Databricks
워크플로우



Delta Live
테이블



Databricks
ML

Unity Catalog

Delta Lake



클라우드 데이터 레이크

모든 원시 데이터
(로그, 텍스트, 오디오, 비디오, 이미지)

엔터프라이즈 분석팀의 제한적인 환경



Difficulty connecting data sources

실시간 분석이나 예측을 위해 SAP 데이터를 다른 소스와 통합할 필요성이 커지면서 고객은 SAP 시스템 밖에서 SAP 데이터를 분석하기 위한 효율적인 방법을 찾고 있습니다.



Limited availability of advanced analytics

SAP에서 수집 가능한 것보다 더 광범위한 데이터를 대규모로 사용하여 고급 분석을 구현하려는 기업이 점점 더 많아지고 있습니다.



Need for lower cost total cost of ownership

SAP는 조직에 일관성과 효율성을 제공하지만 라이선싱/컴퓨팅/기술지원 비용은 데이터 볼륨과 비례합니다. 정확한 예측을 위한 모델링에 필요한 엄청난 양의 데이터에 대해서는 비현실적인 비용이 발생합니다.

Lakehouse를 통해 SAP 데이터 가치 극대화

SAP 데이터의 잠재력 활용 최대화, 실시간 운영 통찰력 강화, 분석 비용 절감



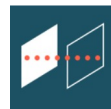
Flexibility in extracting & integrating SAP and third-party data

SAP 데이터를 쉽게 추출하고 non-SAP 데이터와 통합하여 보고 및 예측 분석을 통해 가치 사슬의 전체적인 그림을 이해합니다.



More accurate and robust analytics

처음부터 고급 분석을 위해 설계된 플랫폼으로 모든 종류의 데이터를 사용하여 보다 정확한 예측을 가능하게 합니다.



Improved outcomes and lower TCO

비즈니스 데이터 기능을 그대로 유지하면서 SAP 투자에 대한 결과를 극대화하고 TCO를 줄입니다.



SAP 데이터분석을 위한 현대화 전략

조직은 다양한 구현 방식을 선택할 수 있습니다.

Analytics Sidecar

개별 프로젝트/사용 사례를 지원하기 위해 필요에 따라 SAP에서 Databricks Lakehouse로 데이터 추출

PROS

- 특정 문제점 해결
- SAP 데이터 분석 플랫폼과 하이브리드 사용

CONS

- 각 요구 사항은 독립적인 프로젝트로 구현
- 개발 시간 추가 / TTM 증가

대상 사업부 범위

Enterprise Analytics

Lakehouse 중심 구현을 위한 엔터프라이즈 전략의 일부로 SAP 데이터를 Databricks Lakehouse로 추출

PROS

- 개방형 플랫폼으로 광범위하게 분석가 및 데이터 과학자 지원
- BW/HANA 플랫폼의 데이터 볼륨 최소화

CONS

- 많은 데이터 파이프라인 개발 및 유지보수

엔터프라이즈 범위

SAP BW / HANA EDW Replacement

SAP 데이터의 분석을 Databricks Lakehouse로 교체

PROS

- SAP 기능에 대한 의존도 최소화
- SAP 환경 비용 절감

CONS

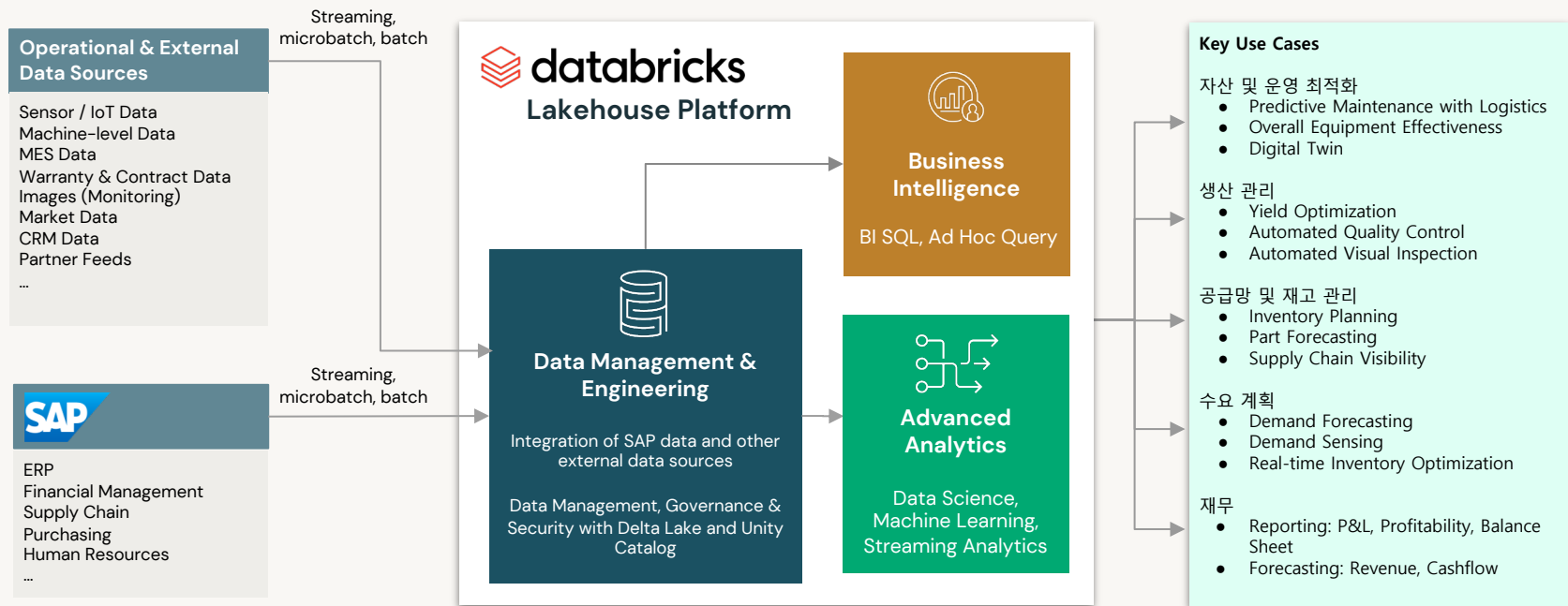
- 특정 요구 사항을 해결하기 위해 별도 솔루션을 추가해야 하는 경우 발생

*장기 투자,
다년간의 로드맵*



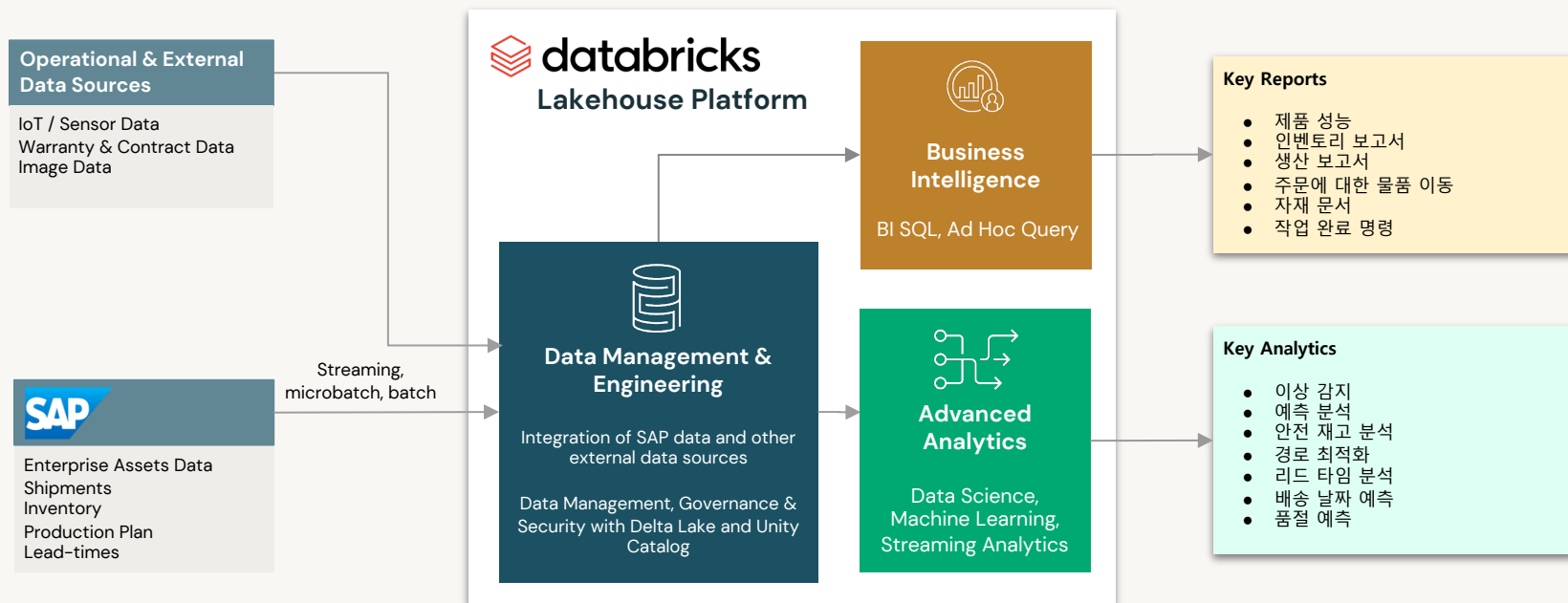
Lakehouse를 통해 SAP 데이터와 Non-SAP 데이터 통합

Databricks Lakehouse 플랫폼에서 SAP 데이터를 통해 구현할 수 있는 사례



SAP 데이터를 사용한 예지정비(Predictive Maintenance)

Lakehouse를 통한 물류 계획으로 유지보수 최적화



고객사례



Developed a solution on Databricks to seamlessly aggregate inventory data from SAP to run large scale simulations for managing spare parts inventory.

With Databricks they were able to **run models 50x faster and saved many millions against \$1bn+ in spare parts inventory, realizing value in weeks**



Built a data Lakehouse on Databricks with purpose-built analytics services for simplified and real-time access to SAP data.

With Databricks they achieved end-to-end data pipeline management, visibility, transparency, **accelerated time to value by 70% and reduced cost by 35%**



Created a virtual control tower on Databricks for end-to-end supply chain visibility across Procurement, Logistics, and Port Operations

With Databricks, they realized **\$2M+ annual savings in logistics costs, gained 10X improvement in supply chain visibility.**



Shell saves millions on spare parts inventory management with Databricks

To maintain production, Shell stocks over 3,000 different spare parts across their global facilities. It's crucial the right parts are available at the right time to avoid outages, but equally important is not overstocking which can be cost-prohibitive. Their current processes and technology stack for maintaining inventory faced challenges.

- **Disjointed Inventory Distribution**
- **Limited DSS (Decision Support System) Data Availability**
- **Lost Business Agility**

Shell achieved an ROI of the project in matter of weeks after the launch.



Products and Services

Microsoft Azure
Azure Databricks
Alteryx
SAP Hana

Organization Size

86,000 employees

Industry

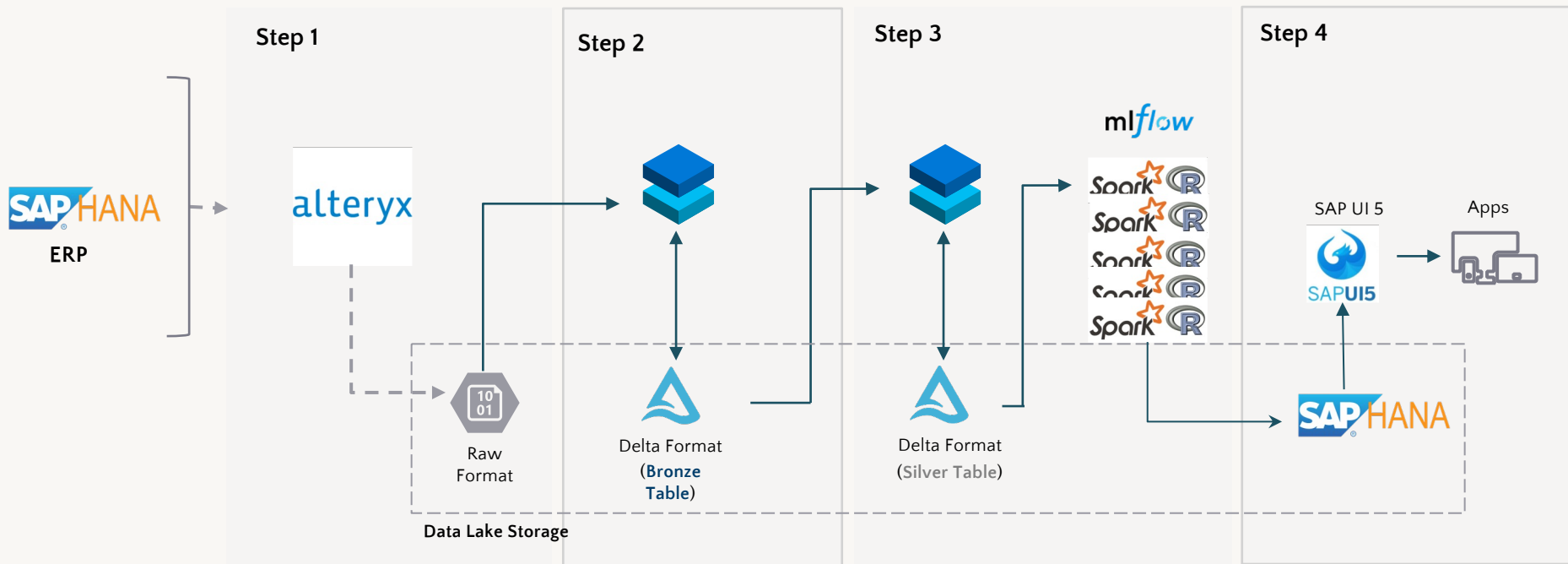
Mining, Oil and Gas

Location

The Netherlands



Spare Parts Inventory with SAP & Databricks



<https://www.itnews.com.au/news/inside-shells-spares-parts-data-science-project-476863>




Result:

- **Predictive Modeling:** Scalable predictive across more than 3,000 types of materials at 50+ locations.
- **Historical Analyses:** 10,000 Markov Chain Monte Carlo iterations
- **Massive Performance Gains:** Reduced the inventory analysis and prediction time to 45 min from 48 h
- **Reduced Expenditures:** Cost savings equivalent to millions of dollars per year.

Appendix


Databricks Solutions


databricks


- Platform
- Solutions
- Learn
- Customers
- Partners
- Company

[Try Databricks](#)





- [Watch Demos](#)
- [Contact Us](#)
- [Login](#)


**DATA+AI
SUMMIT**
BY  databricks

Discover the Lakehouse for your industry





 **Communications, Media & Entertainment**
Earn more attention, capture more imaginations


[Learn how](#)





 **Financial Services**
Build more trust, ensure peace of mind


[Learn how](#)





 **Healthcare and Life Sciences**
Discover and deliver better care

[Learn how](#)

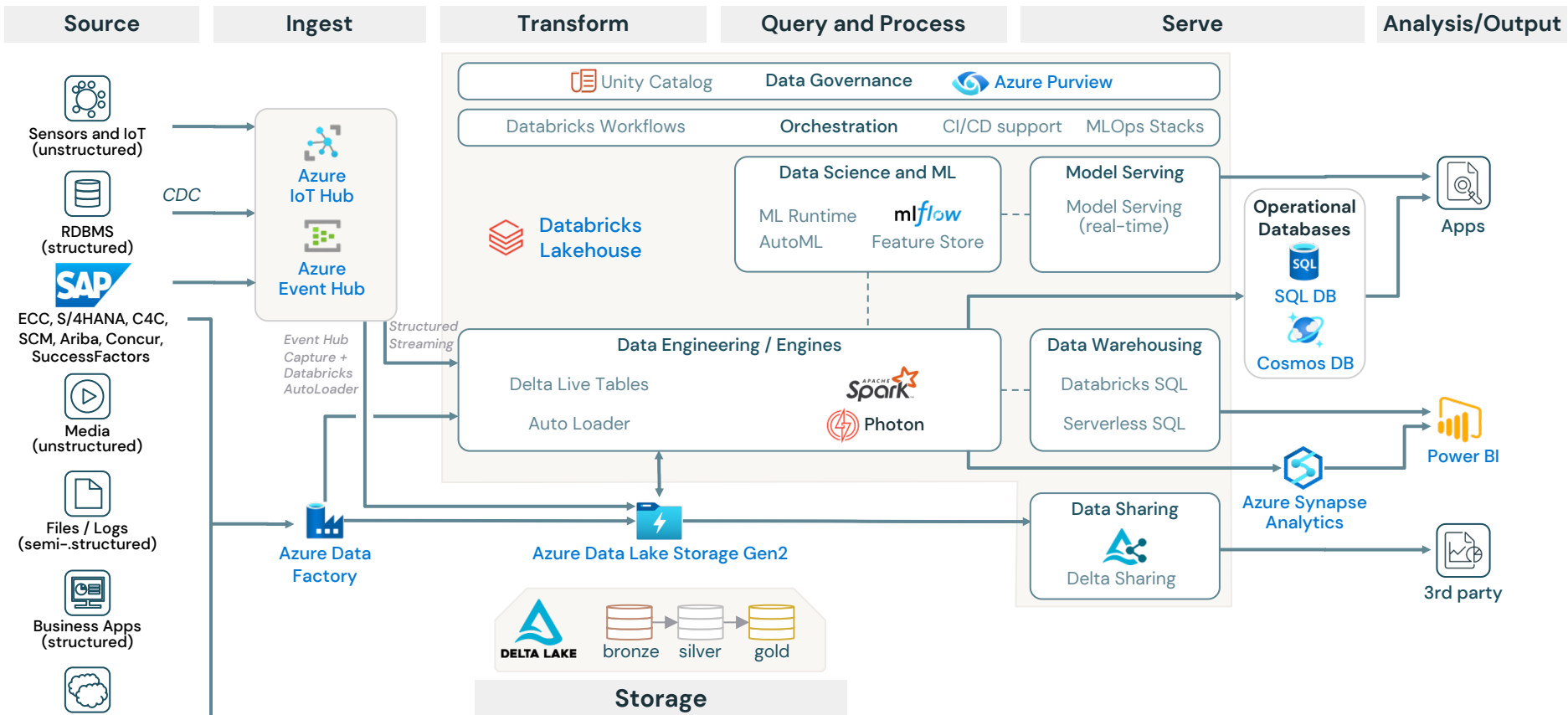
   

 **Retail and Consumer Goods**
Lead customers on their journey and champion your brand

[Learn how](#)

Reference Architecture



DATA+AI SUMMIT

샌프란시스코 현지 + 온라인 버추얼
2023년 6월 26(월)-29(목)

지금 행사 등록하기

레이크하우스로 모든 데이터를
통합하는 방법 알아보세요!



세계에서 가장 큰 규모의 데이터, 분석 및 AI 컨퍼런스가 샌프란시스코와 온라인에서 라이브로 진행됩니다. 나흘 동안 진행되는 이 이벤트에서 업계의 선구적인 리더가 진행하는 키노트, 기술 세션, 실습 교육 및 네트워킹 기회를 활용하실 수 있습니다.

- 데이터와 AI 분야의 사고 리더가 전하는 Apache Spark, Delta Lake, MLflow, TensorFlow, dbt 등 다양한 최신 트렌드와 혁신에 관한 소식을 들어보세요.
- 업계의 다양한 기업과 팀들이 어떻게 데이터 레이크하우스 패러다임을 적용하여 데이터, 분석, AI를 하나의 플랫폼에서 통합하는지 확인하실 수 있습니다.
- 다양한 실습 교육 워크샵에 참석하셔서 레이크하우스에 대해서 자세히 배울 수 있습니다.



감사합니다.