

Azure Databricks

An Introduction

Why Databricks?

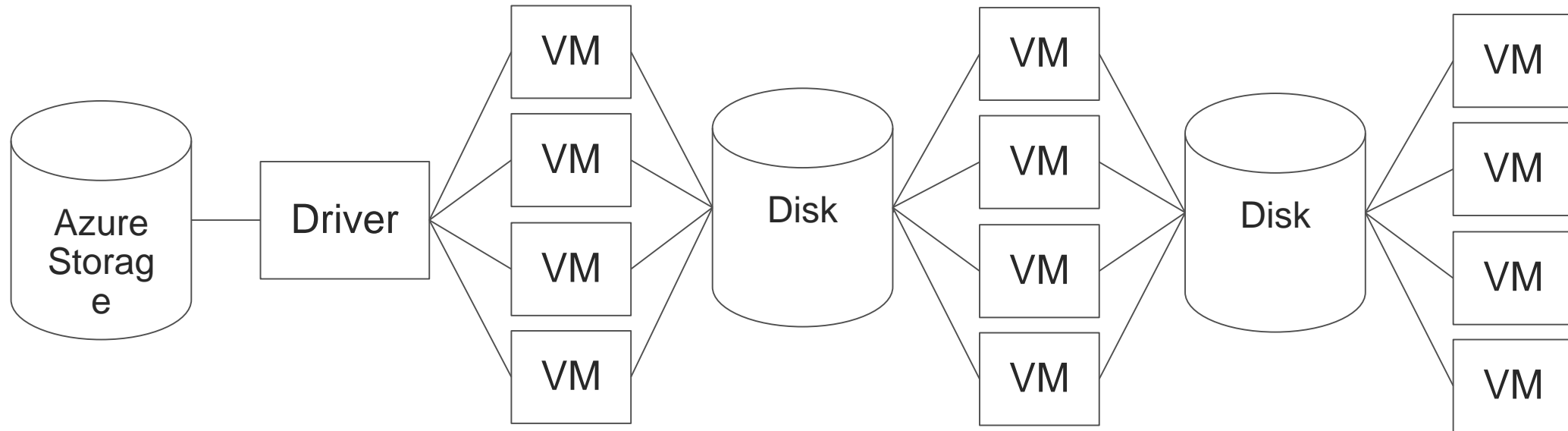


- Unified Analytics – Data Engineering and Data Science together
- In memory engine that is up to 100 times faster than Hadoop
- Separate Compute & Storage
- Built for the Cloud – not re-engineered from 'On Premise'
- Spark is the dominant workload in Hadoop – moving to Databricks is a logical evolution
- Security built in to the integrated and managed Azure Databricks platform
- Collaboration and concurrency amongst Data Science team is native with ADB

Commercials – Consumption economics means it is the natural Opex scale up and down options

Hadoop MapReduce

MapReduce in Hadoop



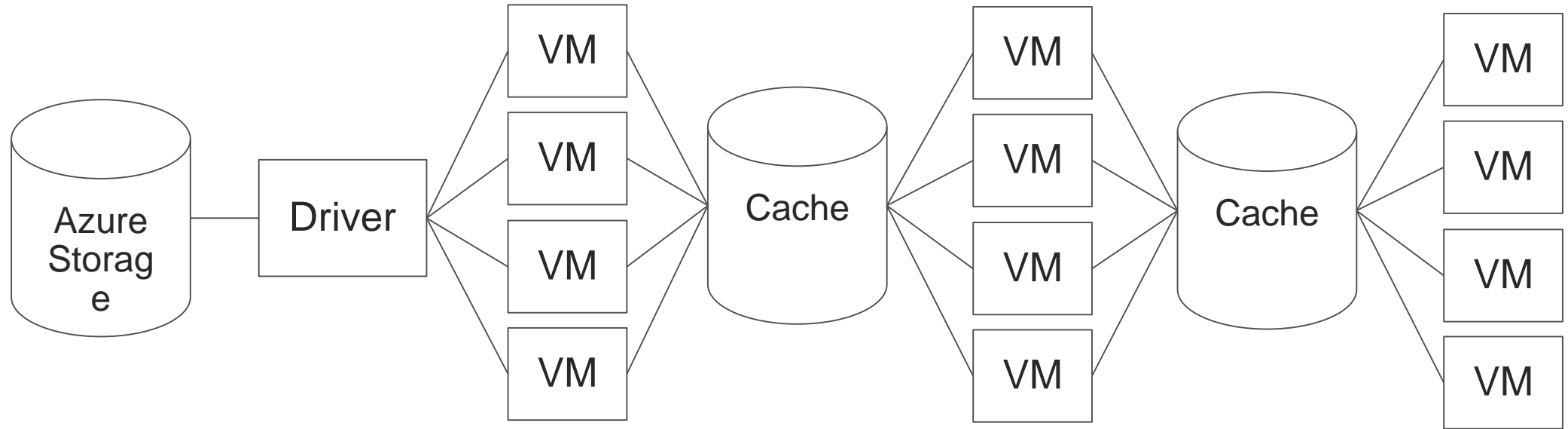
Azure Storage > Driver > VM/Parallelization > write to Disk > VM/Parallelization > write to disk > repeat...



Writing to disk takes time... every time you run this process in MapReduce

What is Azure Databricks?

Apache® Spark™ is FASTER and EASIER than MapReduce in Hadoop



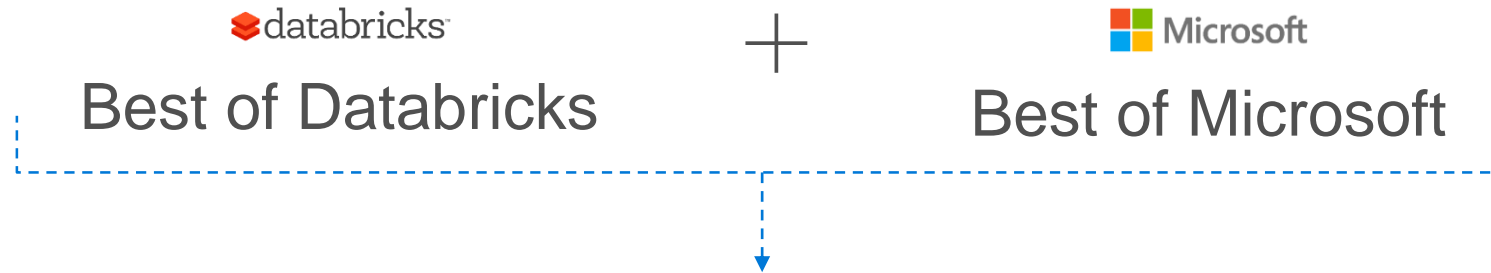
Faster – In Spark data stays in cache this give Spark the speed over MapReduce (writing to disk)



Easier – You can use the language you are most comfortable with in Spark (Python, Scala, R, SQL)

What is Azure Databricks?

A fast, easy and collaborative Apache® Spark™ based analytics platform optimized for Azure



Designed in collaboration with the founders of Apache Spark



One-click set up; streamlined workflows



Interactive workspace that enables collaboration between data scientists, data engineers, and business analysts.



Native integration with Azure services (Power BI, SQL DW, Cosmos DB, Blob Storage)



Enterprise-grade Azure security (Active Directory integration, compliance, enterprise-grade SLAs)

Azure Databricks key audiences & benefits



Data scientist

- Integrated workspace
- Easy data exploration
- Collaborative experience
- Interactive dashboards
- Faster insights
 - Best Spark & serverless
 - Databricks managed Spark



Data engineer

- Improved ETL performance
 - Zero management clusters, serverless
- Easy to schedule jobs
- Automated workflows
- Enhanced monitoring & troubleshooting
 - Automated alerts & easy access to logs
- Zero Management Spark
- Cluster democratization (High-concurrency)

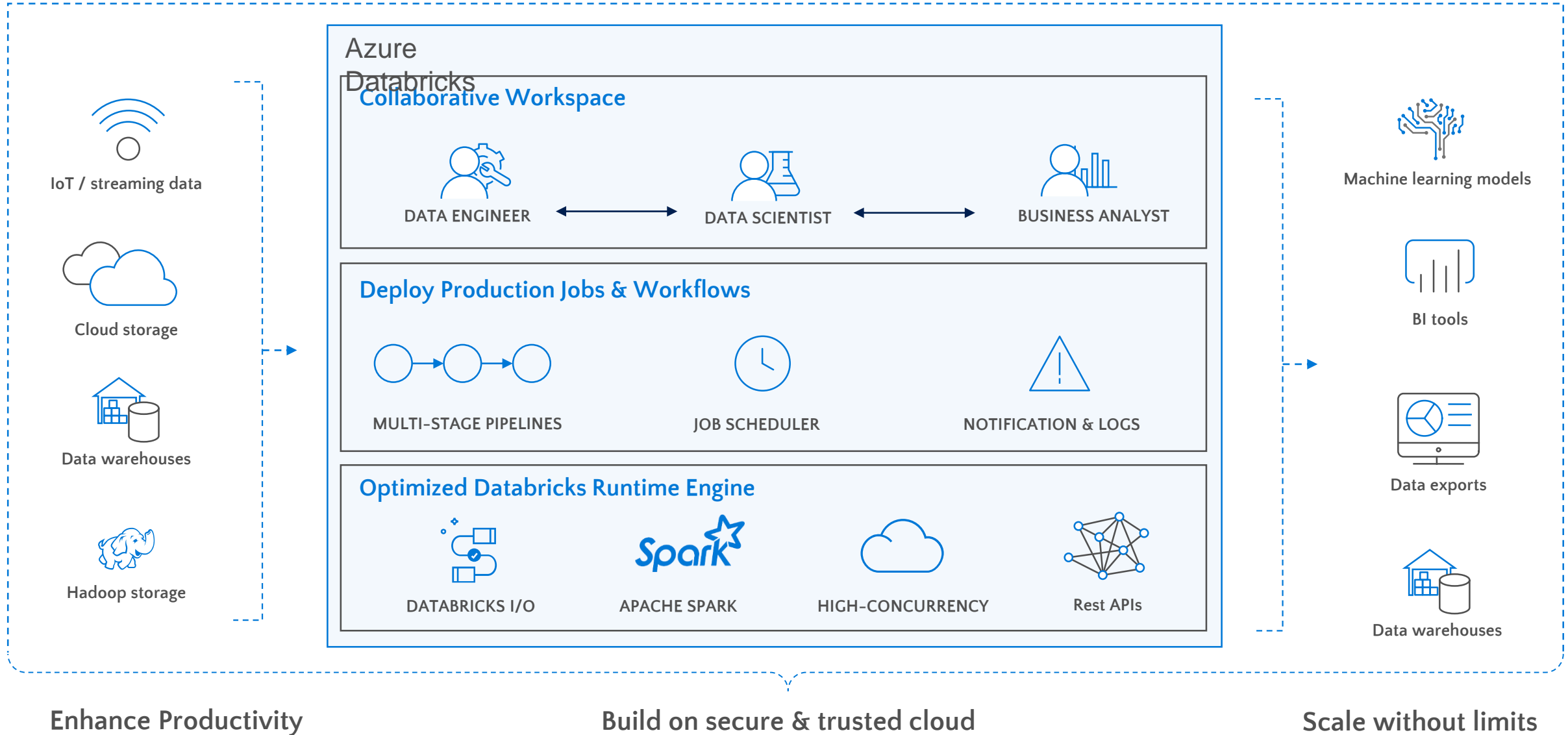


CDO, VP of analytics

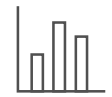
- Fast, collaborative analytics platform accelerating time to market
- No dev-ops required
- Enterprise grade security
 - Encryption
 - End-to-end auditing
 - Role-based control
 - Compliance

← Unified analytics platform →

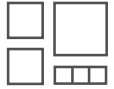
Azure Databricks



BIG DATA & ADVANCED ANALYTICS AT A GLANCE



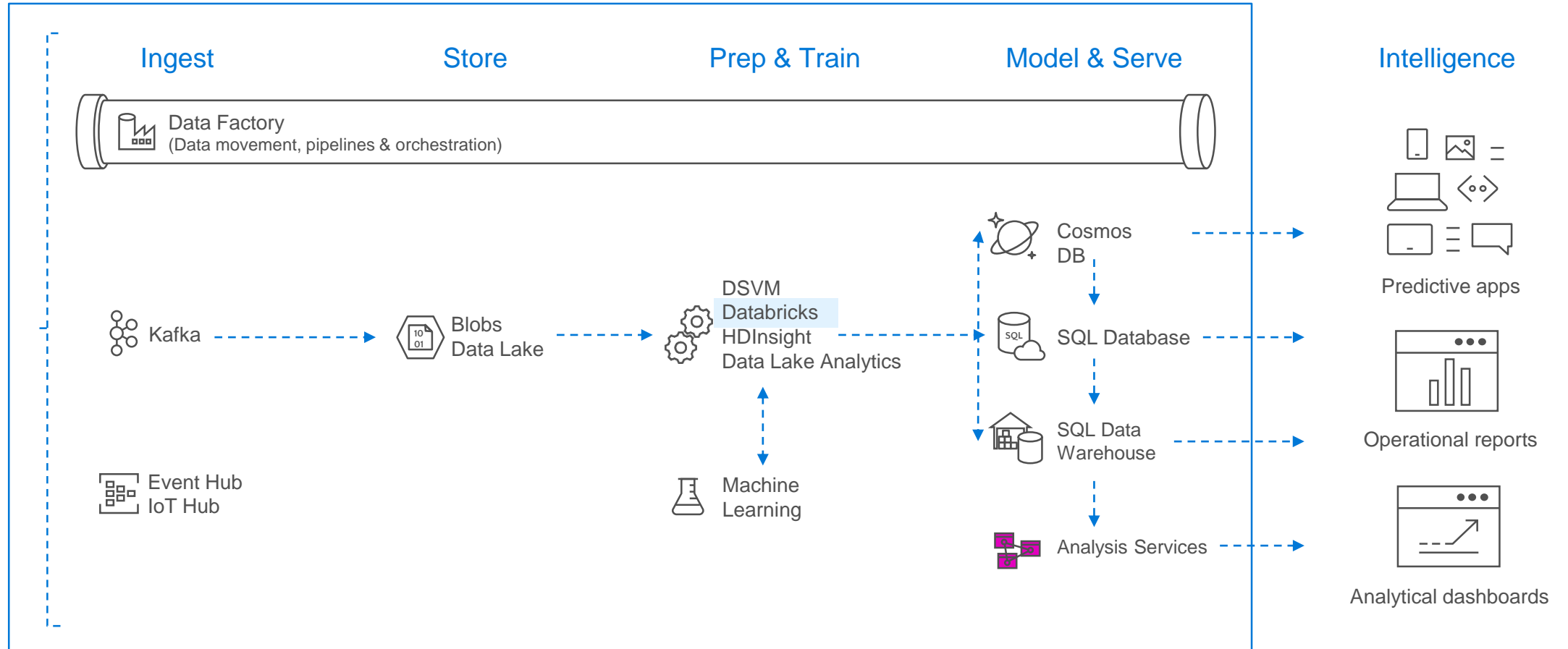
Business apps



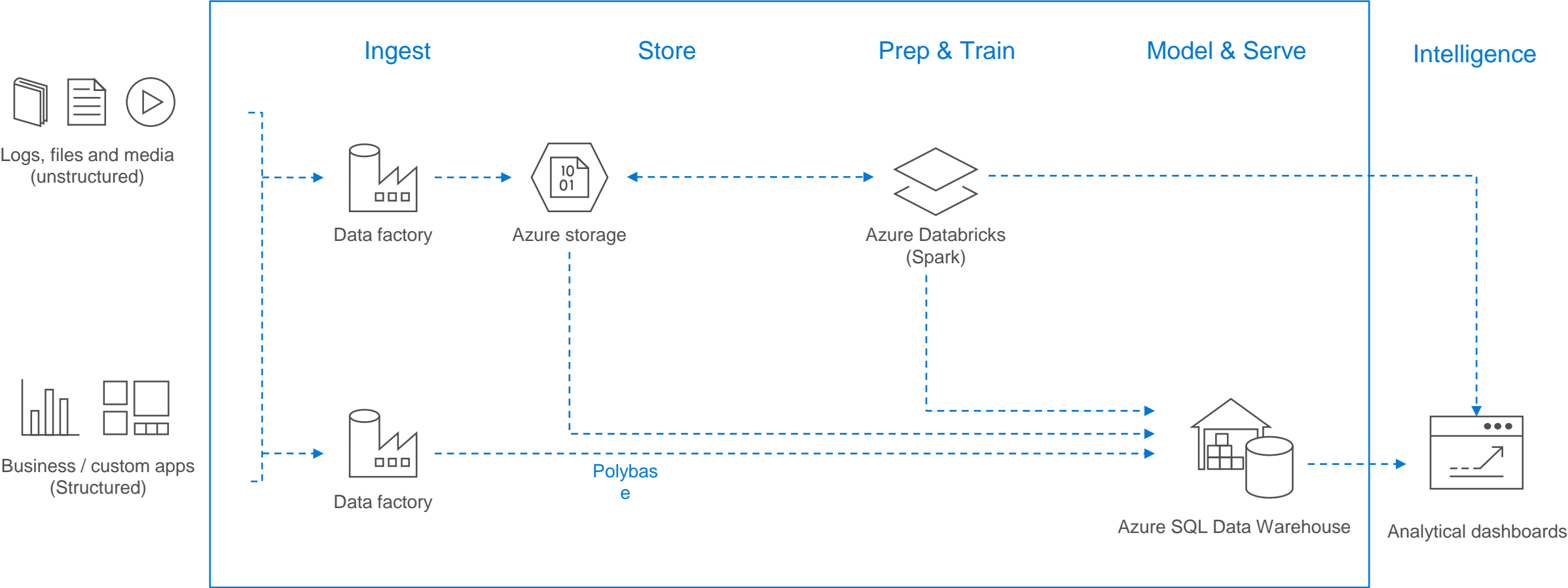
Custom apps



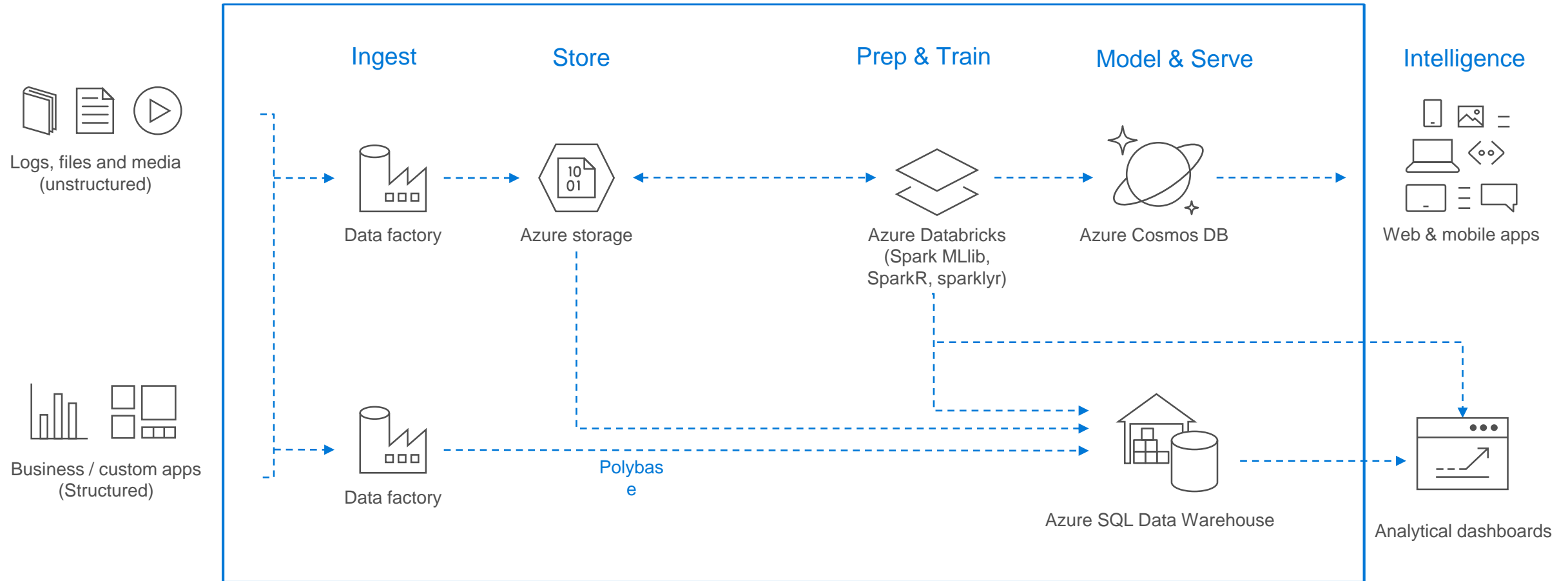
Sensors and devices



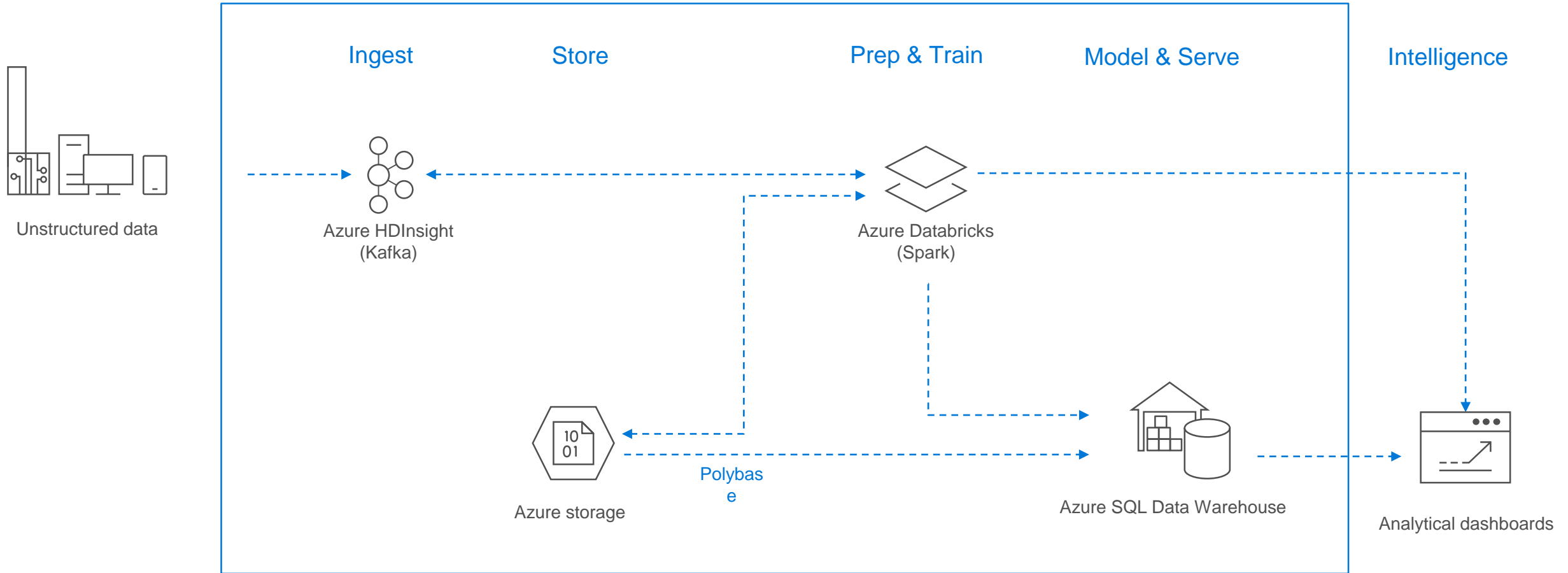
Modern Big Data Warehouse



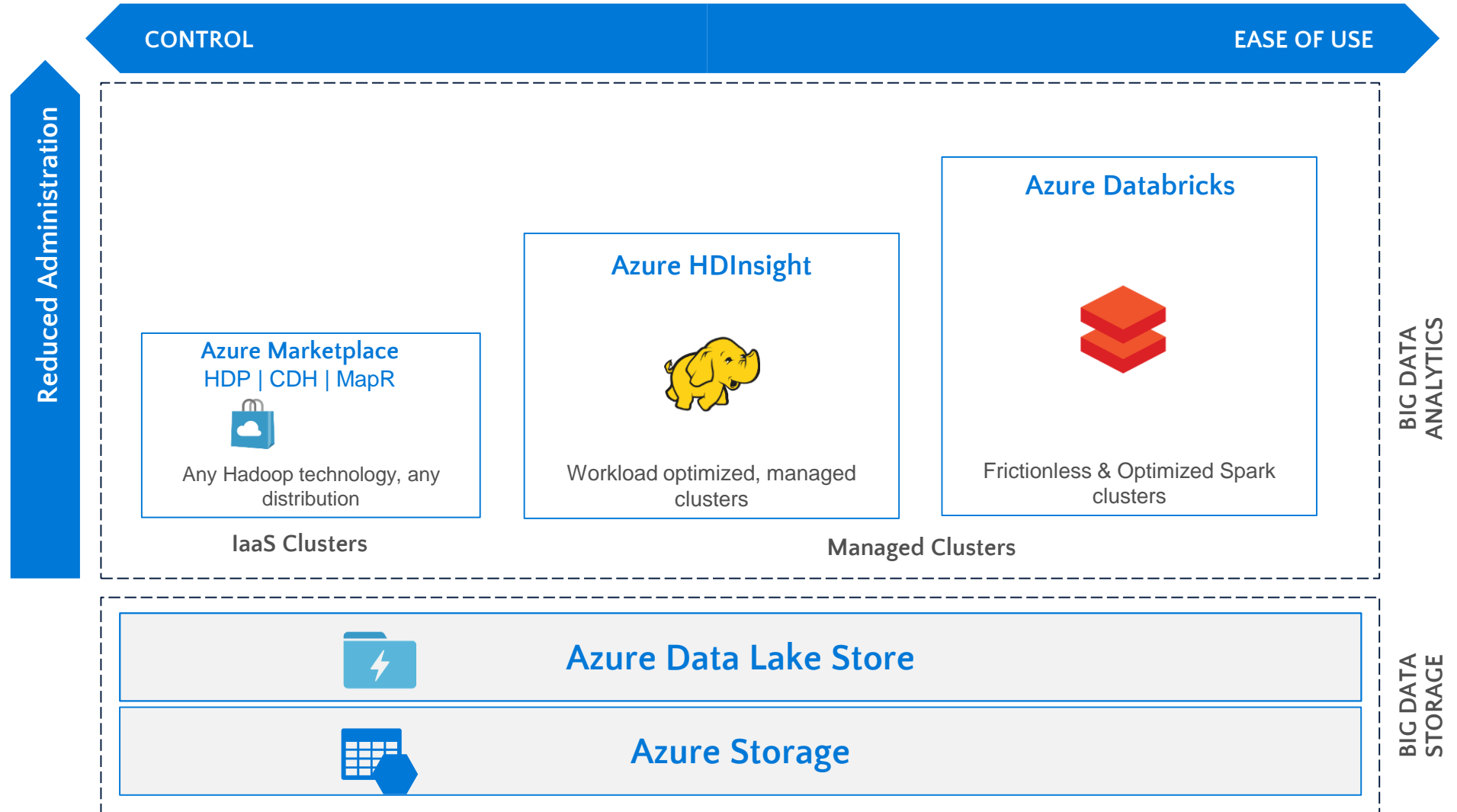
Advanced Analytics on Big Data



Real-time analytics on Big Data



KNOWING THE VARIOUS BIG DATA SOLUTIONS



Use Cases

“At Lennox International, we have 1000’s of devices streaming data back into our IoT environment. With Azure Databricks, we moved from 60% accuracy to 94% accuracy on detecting equipment failures. Using Azure Databricks has opened the flood gates to all kinds of new use cases and innovations. In our previous process, 15 devices, which created 2 million records, took 6 hours to process. With Azure Databricks, we are able to process 25,000 devices – 10 billion records – in under 14 minutes.”

Sunil Bondalapati, Director of IT, Lennox International



Lennox International is a global leader in the heating, air conditioning, and refrigeration markets.

Energy intelligence: renewables.AI uses Azure and Spark to help build a stable and profitable solar energy market

Solar power producers need to predict when they will generate electricity and coordinate distribution with multiple energy markets. That's why renewables.AI uses Microsoft Azure and Apache Spark to deliver a data analytics service that helps solar energy providers manage production and identify high-value energy markets. With Azure Databricks and other Azure services, renewables.AI can streamline product development, drive solar industry revenue, and promote a stable energy future for everyone.



Products and Services

Azure Databricks
Azure SQL Database
Azure Service Fabric
Azure Event Hubs
Azure Data Lake

Organization Size

12 Employees

Industry

Energy

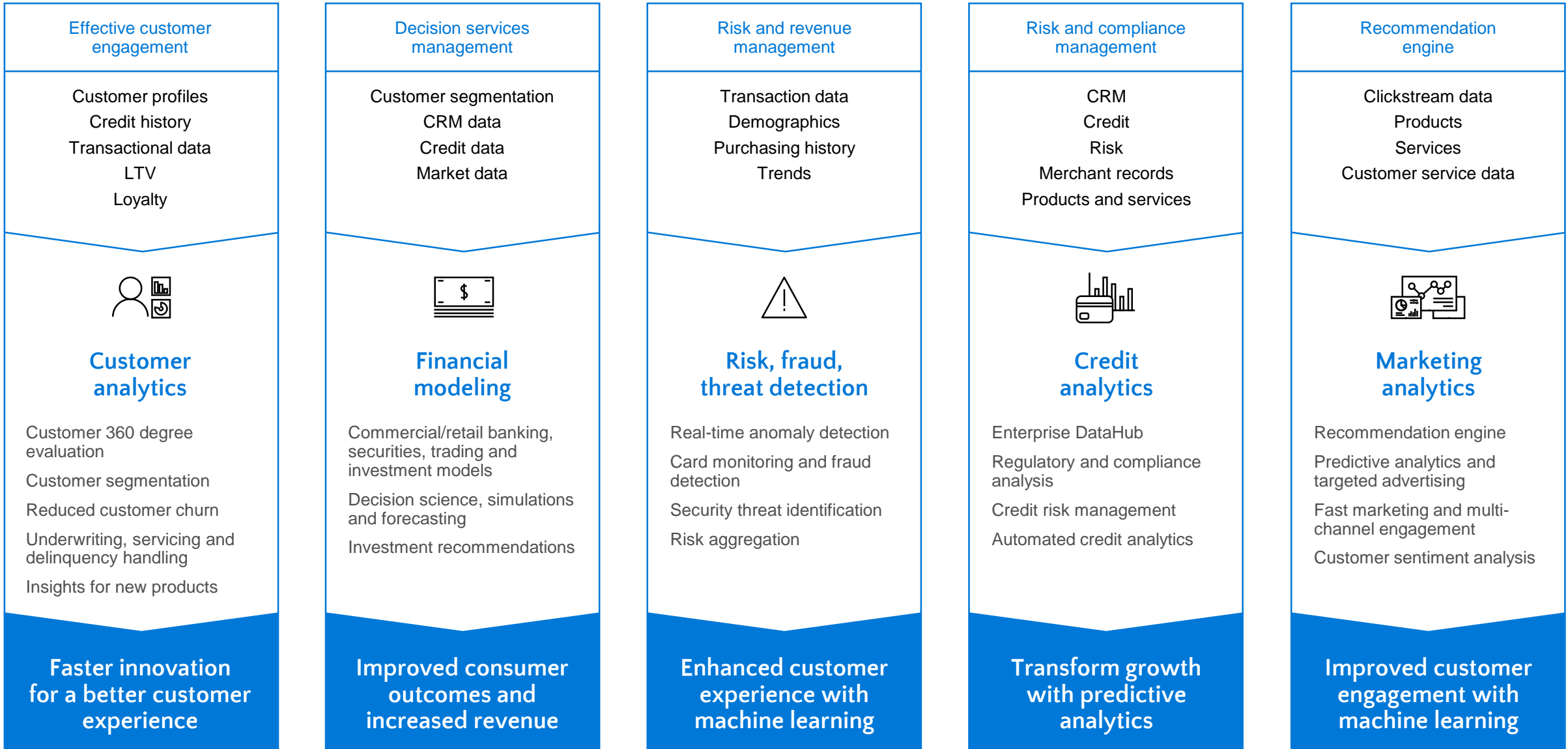
Country

United Kingdom



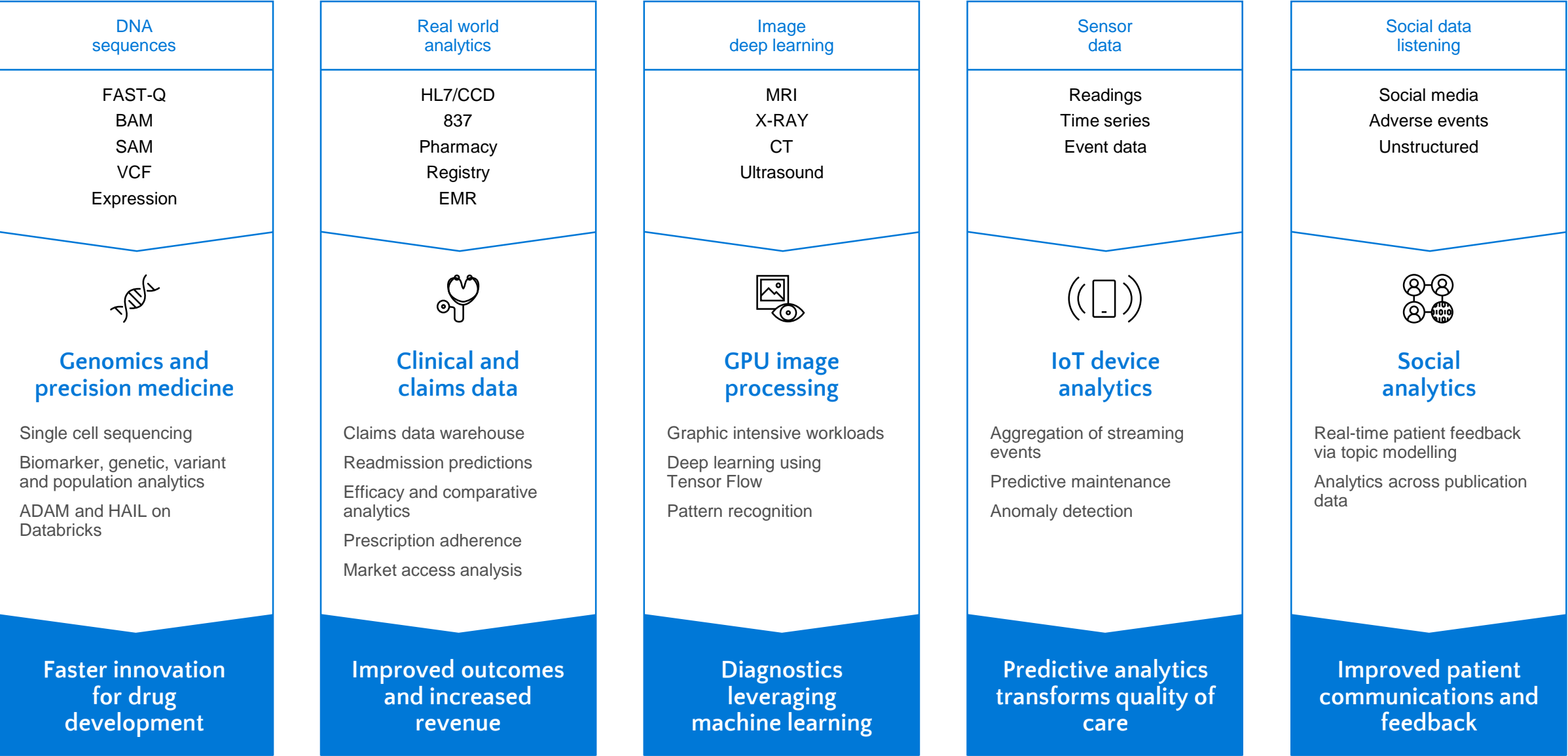
FINANCIAL SERVICES

Use cases








HEALTH & LIFE SCIENCES

Use cases



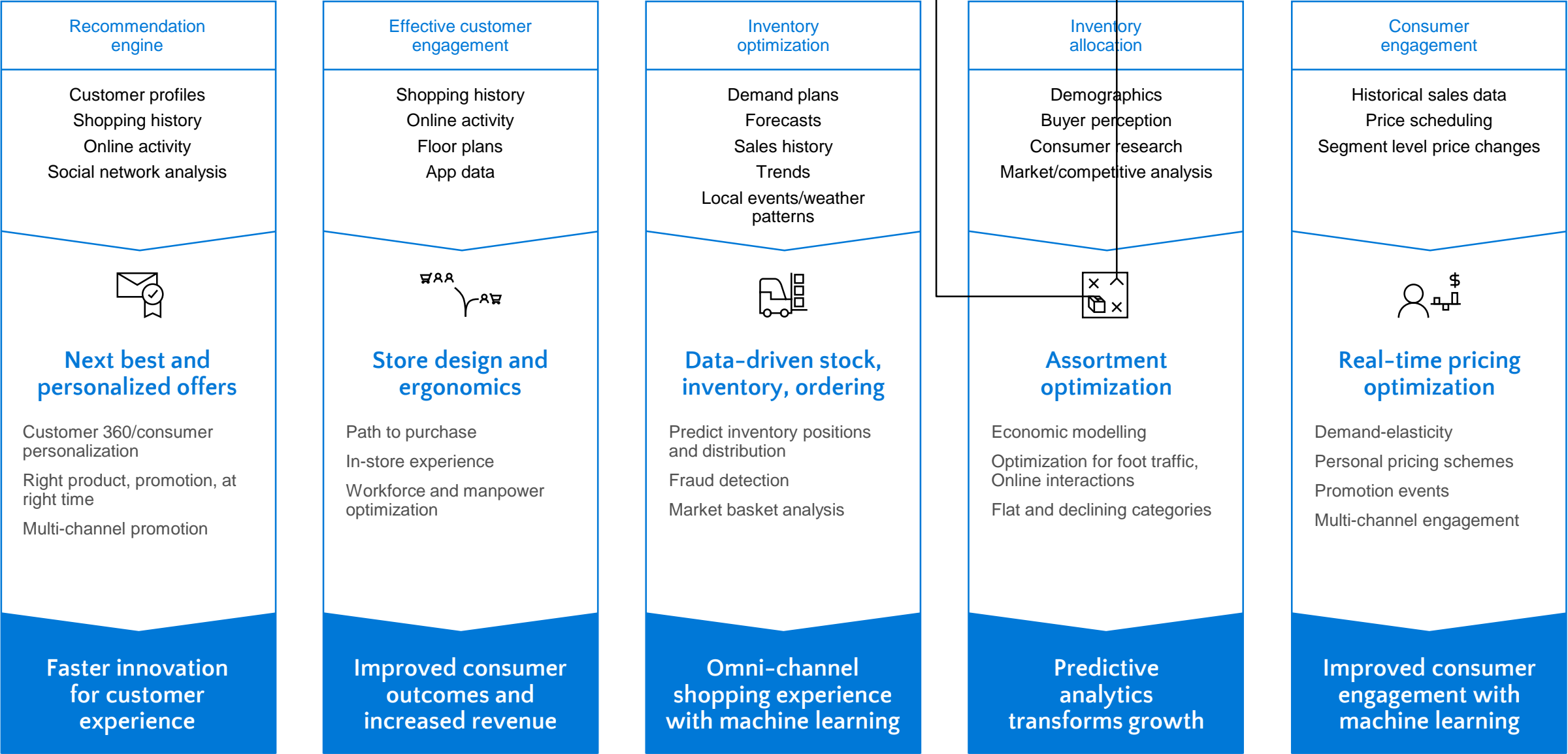
MEDIA & ENTERTAINMENT

Use cases

<div>Personalized recommendations</div> <div>Customer profiles Viewing history Online activity Content sources Channels</div> <div></div> <div>Content personalization</div> <div>Personalized viewing and engagement experience Click-path optimization Next best content analysis Improved real time ad targeting</div> <div>Faster innovation for customer experience</div>	<div>Effective customer retention</div> <div>Customer profiles Online activity Content distribution Services data</div> <div></div> <div>Customer churn prevention</div> <div>Quality of service and operational efficiency Market basket analysis Customer behavior analysis Click-through analysis</div> <div>Improved consumer outcomes and increased revenue</div>	<div>Information optimization</div> <div>Consumption logs Clickstream and devices Marketing campaign responses</div> <div></div> <div>Recommendation engine</div> <div>Ad effectiveness Content monetization Fraud detection Information-as-a-service High value user engagement</div> <div>Enhance user experience with machine learning</div>	<div>Inventory allocation</div> <div>Transactions Subscriptions Demographics Credit data</div> <div></div> <div>Predictive analytics</div> <div>Predict audience interests Network performance and optimization Pricing predictions Nielsen ratings and projections Mobile spatial analytics</div> <div>Predictive analytics transforms growth</div>	<div>Consumer engagement analysis</div> <div>Content metadata Ratings Comments Social media activity</div> <div></div> <div>Sentiment analysis</div> <div>Demand-elasticity Social network analysis Promotion events time-series analysis Multi-channel marketing attribution</div> <div>Improved consumer engagement with machine learning</div>
---	---	--	---	--






RETAIL

Use cases








ADVERTISING AND MARKETING TECH

Use cases

<div>Effective customer engagement</div> <div>Customer profiles Online history Transaction data Loyalty</div>	<div>Recommendation engine</div> <div>Customer segmentation CRM data Credit data Market data</div>	<div>Risk and fraud analysis</div> <div>Transaction data Demographics Purchasing history Trends</div>	<div>Campaign reporting analytics</div> <div>CRM Merchant records Products Services Marketing data</div>	<div>Brand promotion and customer experience</div> <div>Social media Online history Customer service data</div>
<div></div> <div>Customer value analytics</div> <div>Customer 360, segmentation aggregation and attribution Audience modelling/index report Reduce customer churn Insights for new products Historical bid opportunity as a service</div>	<div></div> <div>Next best and personalized offers</div> <div>Right product, promotion, at right time Real time ad bidding platform Personalized ad targeting Ad performance reporting</div>	<div></div> <div>Risk and fraud management</div> <div>Real-time anomaly detection Fraud prevention Customer spend and risk analysis Data relationship maps</div>	<div></div> <div>Sales and campaign optimization</div> <div>Optimizing return on ad spend and ad placement Multi-channel promotion Ideal customer traits Optimized ad placement</div>	<div></div> <div>Sentiment analysis</div> <div>Opinion mining/social media analysis Deeper customer insights Customer loyalty programs Shopping cart analysis</div>
<div>Faster innovation for customer growth</div>	<div>Improved outcomes and increased revenue</div>	<div>Risk management with machine learning</div>	<div>Predictive analytics transforms growth</div>	<div>Improved customer engagement with machine learning</div>

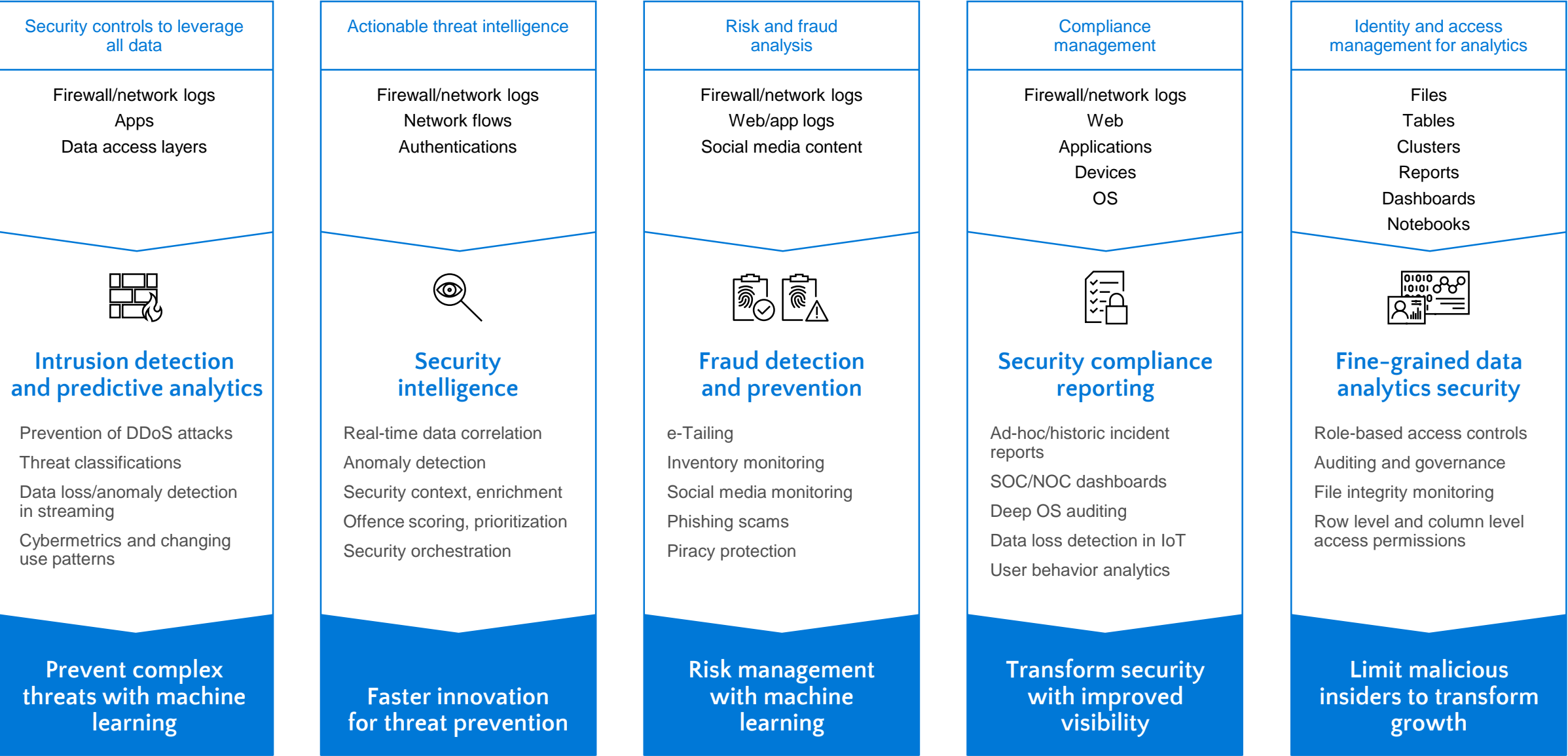
OIL & GAS AND ENERGY

Use cases

<div>Upstream optimization, maximize well life</div> <div>Field data Asset data Demographics Production data</div> <div></div> <div>Digital oil field/ oil production</div> <div>Production optimization Integrate exploration and seismic data Minimize lease operating expenses Decline curve analysis</div> <div>Faster innovation for revenue growth</div>	<div>Grid operations, asset inventory optimization</div> <div>Sensor stream data UAVs images Inventory data Production data</div> <div></div> <div>Industrial IoT</div> <div>Pipeline monitoring Preventive maintenance Smart grids and microgrids Grid operations, Field Service Asset performance as a Service</div> <div>Improved outcomes and increased revenue</div>	<div>Supply-chain optimization</div> <div>Transaction data Demographics Purchasing history Trends</div> <div></div> <div>Supply-chain optimization</div> <div>Trade monitoring, optimization Retail mobile applications Vendor management - construction, transportation, truck & delivery optimization</div> <div>Optimizing supply- chain with machine learning</div>	<div>Risk optimization</div> <div>Sensor stream data Transport Retail data Grid production data Refinery tuning parameters</div> <div></div> <div>Safety and security</div> <div>Real-time anomaly detection Predictive analytics Industrial safety Environment health and safety</div> <div>Predictive analytics transforms safety and security</div>	<div>Recommendations engine</div> <div>Clickstream data Products Services Market data Competitive data Demographics</div> <div></div> <div>Sales and marketing analytics</div> <div>Fast marketing and multi-channel engagement Develop new products and monitor acceptance of rates Predictive energy trading Deep customer insights</div> <div>Improved customer engagement with machine learning</div>
---	--	--	---	--

SECURITY

Use cases



Azure Databricks Next Step

Azure Databricks Home
Documentation, Pricing, Get Started Information

<https://azure.microsoft.com/en-us/services/databricks/>

Demo