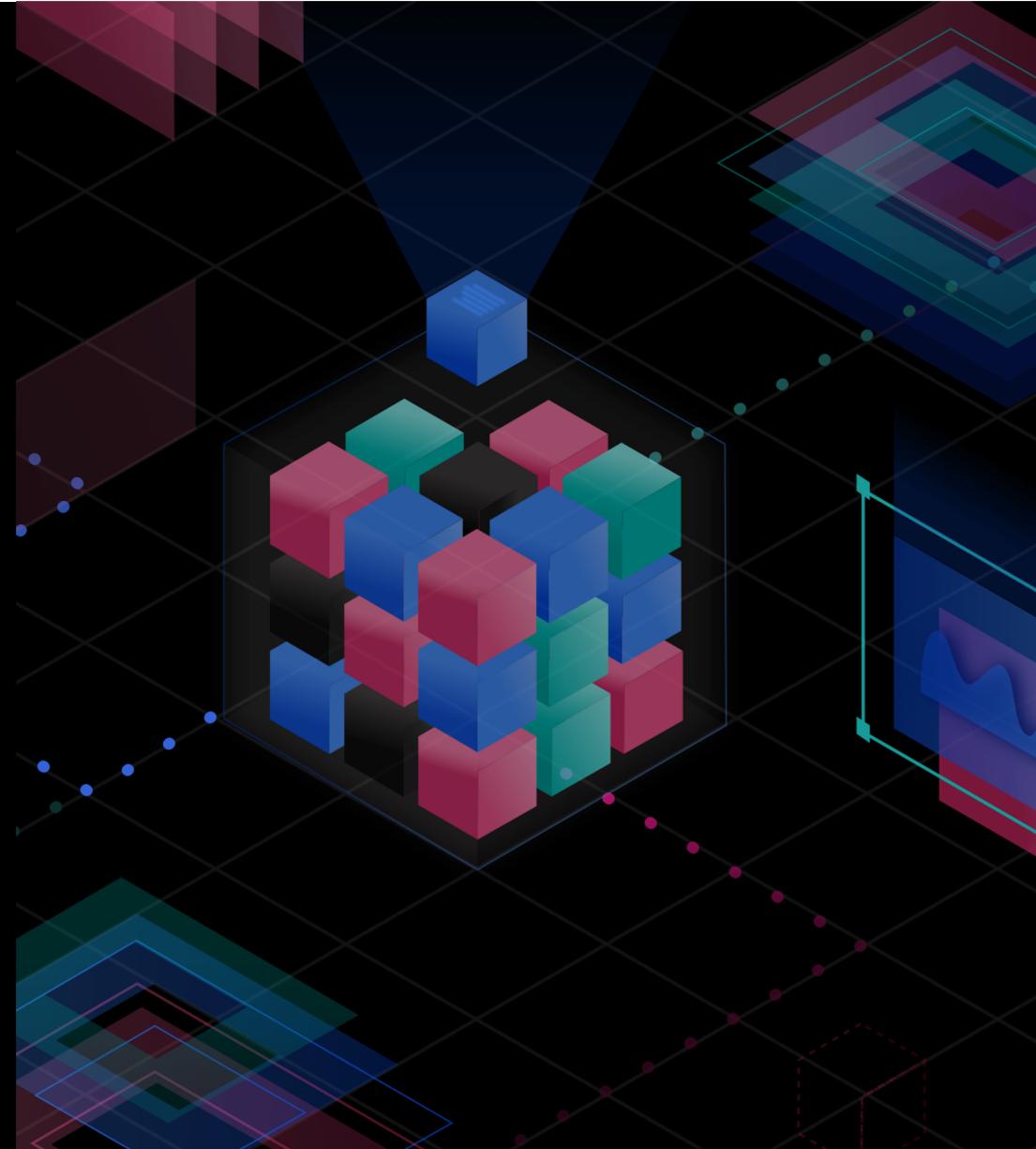


IBM Cloud Pak for Data

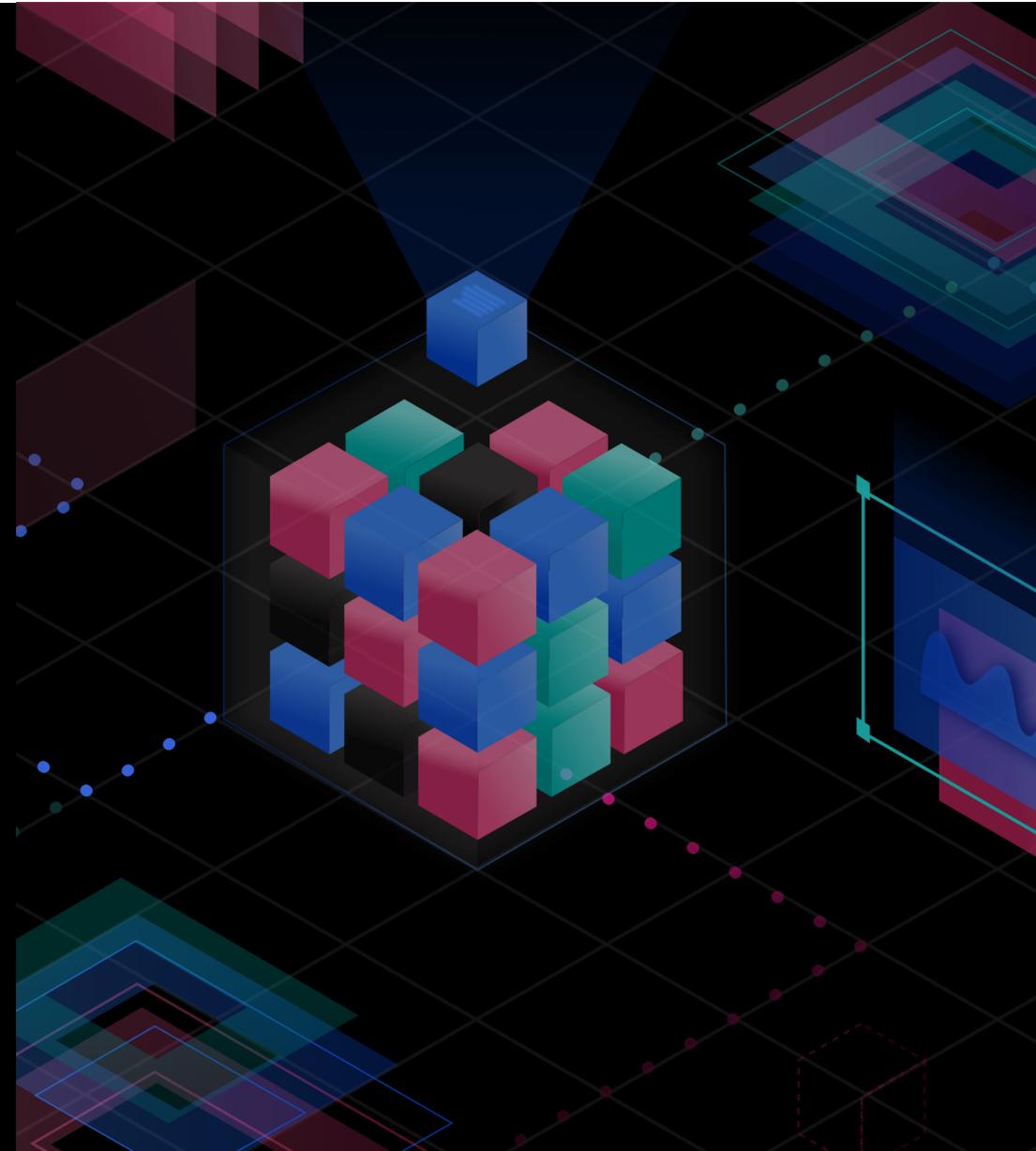
Any Data. Any Cloud. Anywhere.

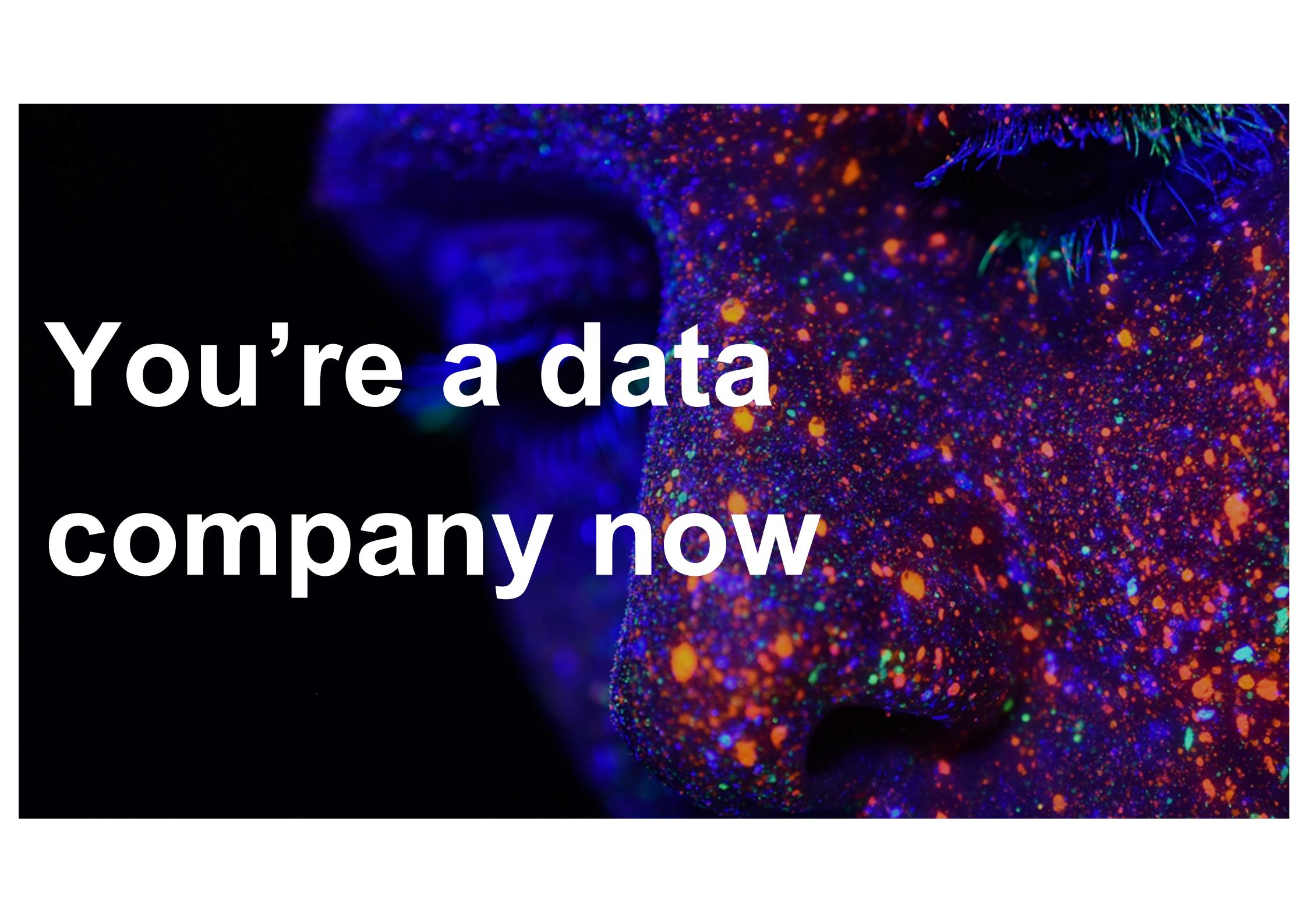
Rachana Vishwanathula
Data Scientist
Hybrid Cloud Build Team



Agenda

- Becoming a data-driven organization
- The ladder to AI
- What is Cloud Pak for Data?
- Deployment patterns
- Use Cases





You're a data
company now

What it means to deliver a **data-centric** customer experience

“A critical component of continuous customer experience includes evaluating where AI resources can extend or support your customer-facing associates, like chat bots, virtual customer assistants, or real-time analytics at key points in the customer journey.”

Gartner

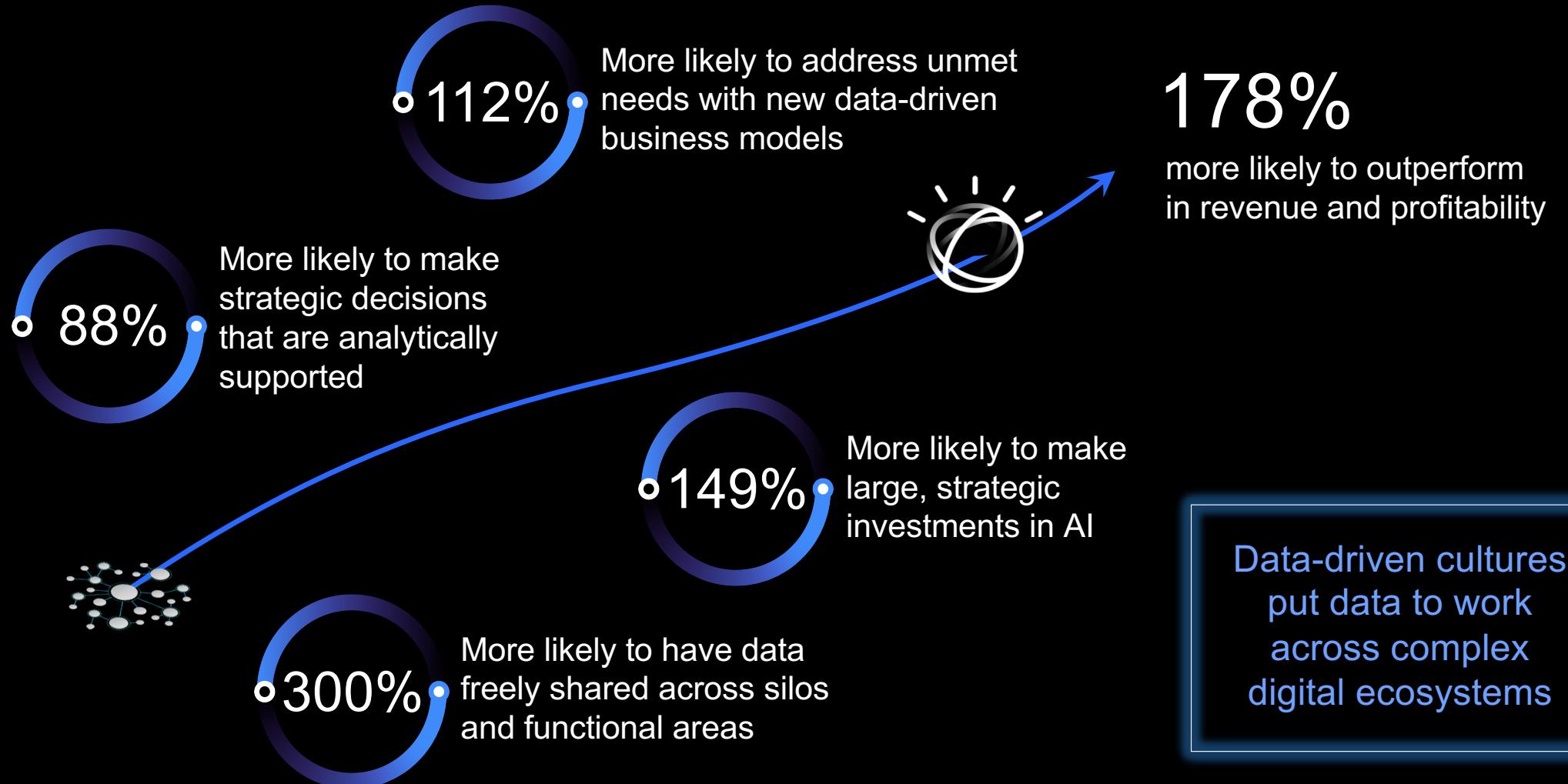
“Smart data will take priority over big data... for too long, the industry relied on sheer quantity of data with no quality metrics for making key audience assumptions.”

TC

“If you fail to keep up with your clients’ increasing demands, the chances they will reach out to your competitors increases.”

Forbes

The best performers are data-driven



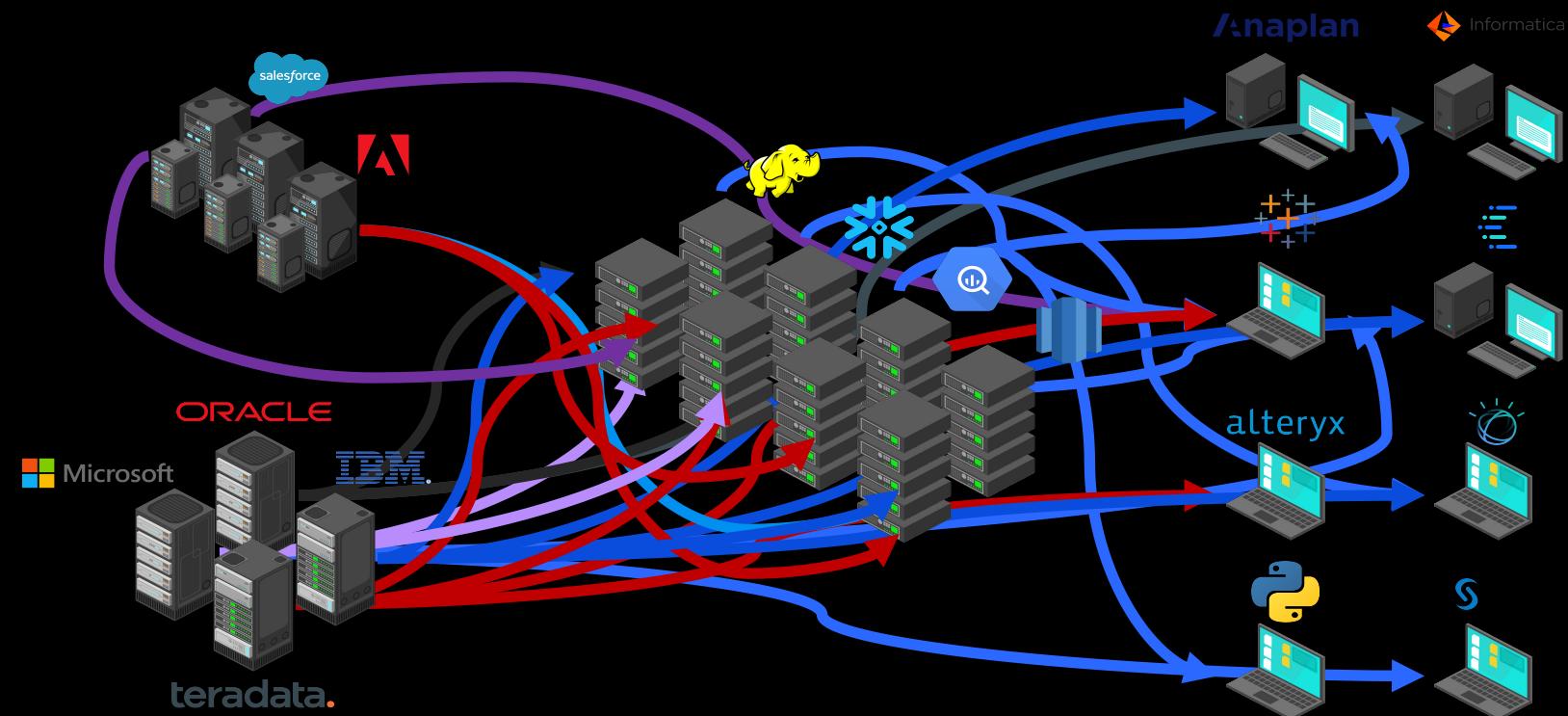
Data management complexity is the great inhibitor



- Data is spread across expansive digital ecosystems as cloud apps proliferate
- Overly manual efforts to move and unify data are slow, costly, and have limited impact
- Rapid growth in data with decreasing lifespans will only make things worse

42%
annual growth
in enterprise
data volumes

Common reality:
Tool oriented, little value, unnecessarily complex, and costly



The AI Ladder

IBM's prescriptive approach to the journey to AI

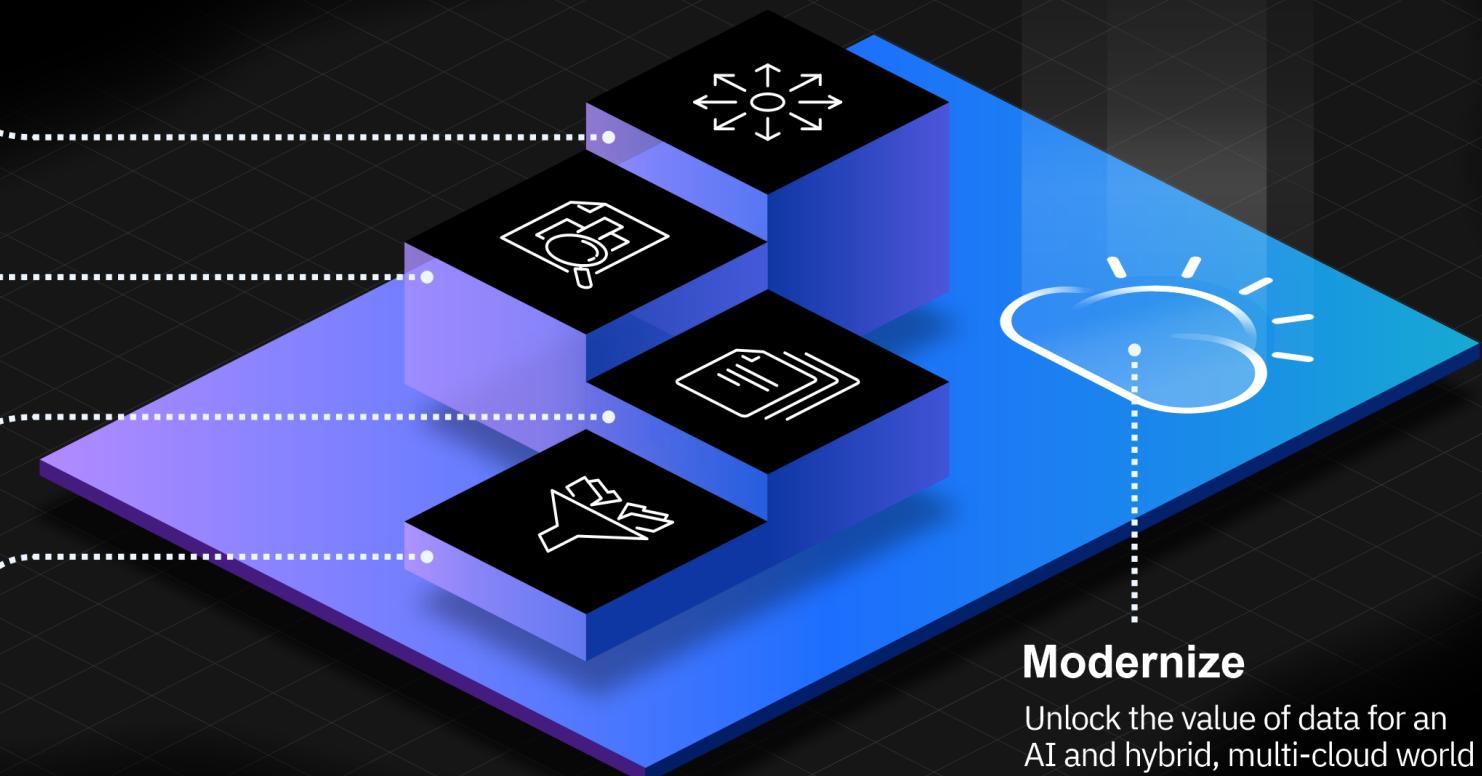
Infuse
Operationalize AI
throughout the enterprise

Analyze
Build and scale AI with
trust and transparency

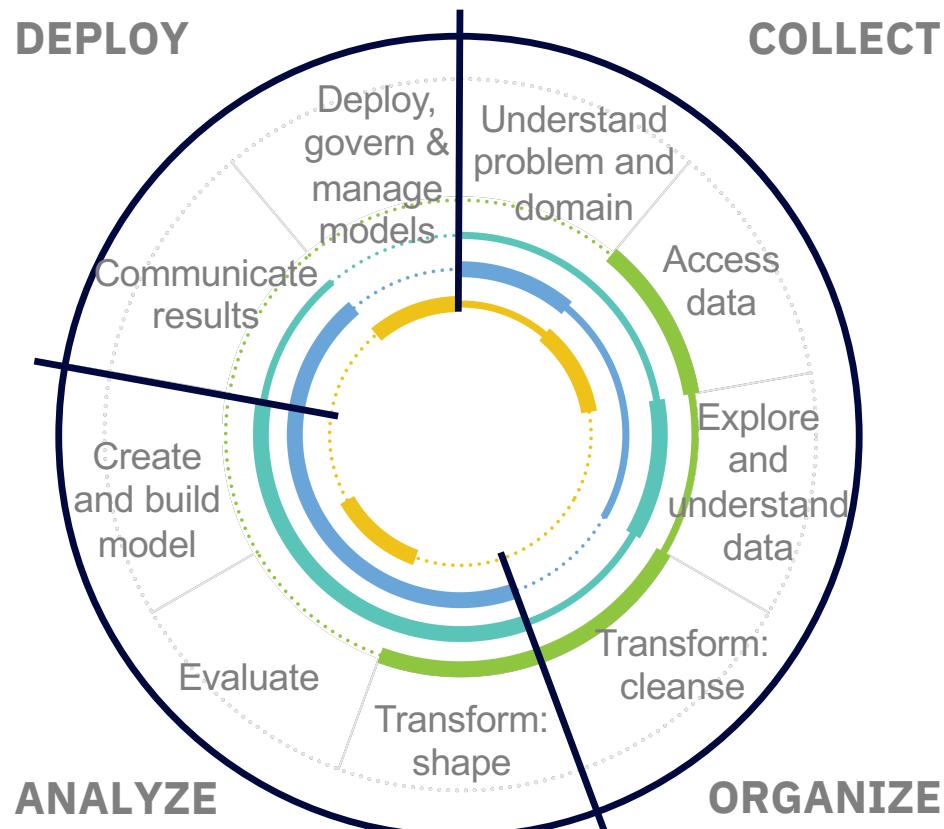
Organize
Create a business-ready
analytics foundation

Collect
Make all data simple
and accessible

Modernize
Unlock the value of data for an
AI and hybrid, multi-cloud world



Methodology



Data Engineer

Architects data pipelines & ensures operability



Data Steward

Governs data and ensures regulatory compliance



Data Scientist

Gets deep into the data to draw insights for the business



Business Analyst

Works with data to apply insights to business strategy



App Developer

Plugs into analysis and code to build apps

Decision Maker(s)



CXO

CIO / CDO / CAO
LOB Leader

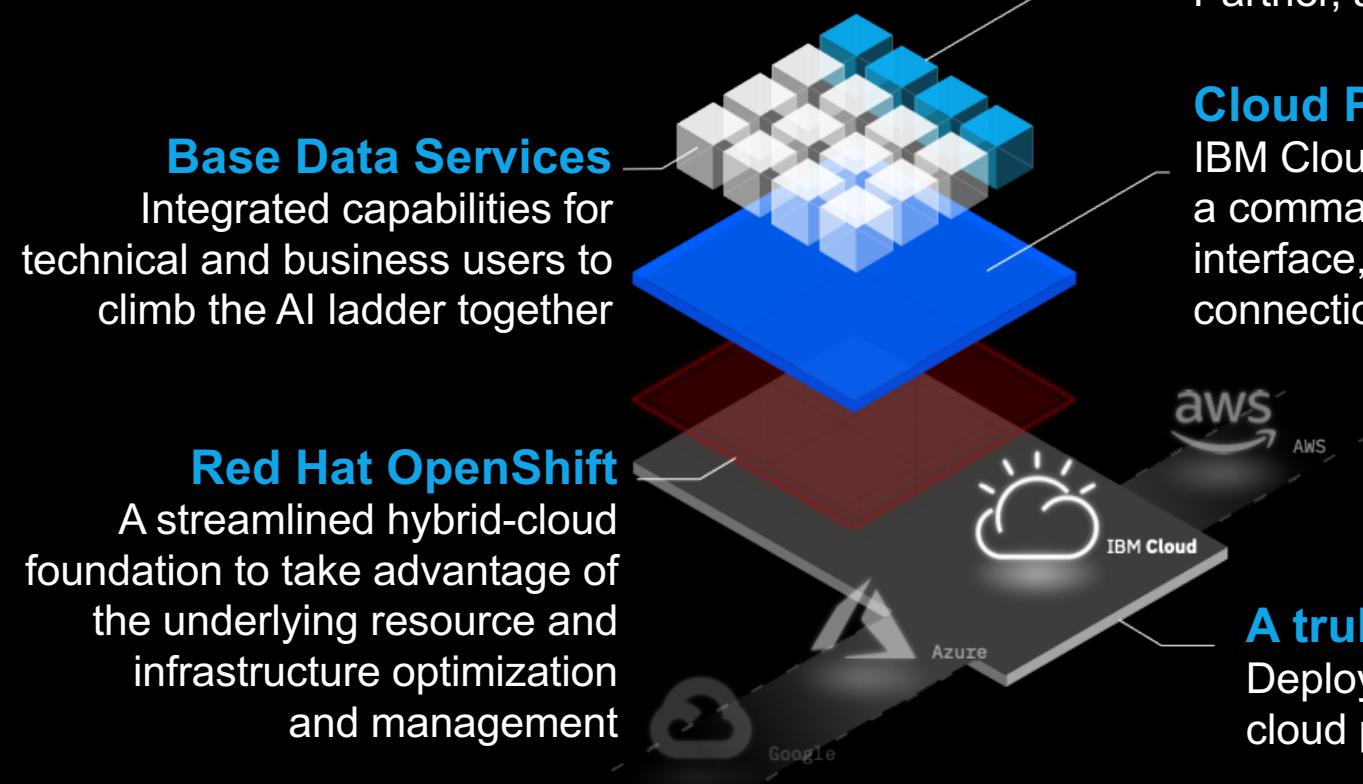


Sys Admin

IBM.

Cloud Pak for Data

Simplifies, unifies, and automates your journey to AI



Base Data Services

Integrated capabilities for technical and business users to climb the AI ladder together

Red Hat OpenShift

A streamlined hybrid-cloud foundation to take advantage of the underlying resource and infrastructure optimization and management

Extended Services

A growing ecosystem of Open Source, Partner, and IBM services

Cloud Pak Control Plane

IBM Cloud Pak Foundational Services provide a command-line interface (CLI), administration interface, services catalog, central list of connections, and the central user experience

A truly Hybrid-Cloud world

Deployable on-premises or on any vendor's cloud platform; also available as-a-Service

Cloud Pak for Data

Unified, modular, deployable anywhere

App Developers | Business Analysts | Data Engineers | Data Stewards | Data Scientists | Business Users

Integrated User Experience

Extensible: APIs, partner ecosystem, accelerators, and solutions

Collect

- Data virtualization
- SQL and NoSQL databases
- Event ingestion
- Streaming Analytics
- Apache Spark

Organize

- Data transformation
- Data quality and classification
- Policies and rules
- Data cataloging
- Self-service discovery and search

Analyze

- Data science and visualization
- AutoAI
- Model trust and explainability
- Model optimization

Infuse

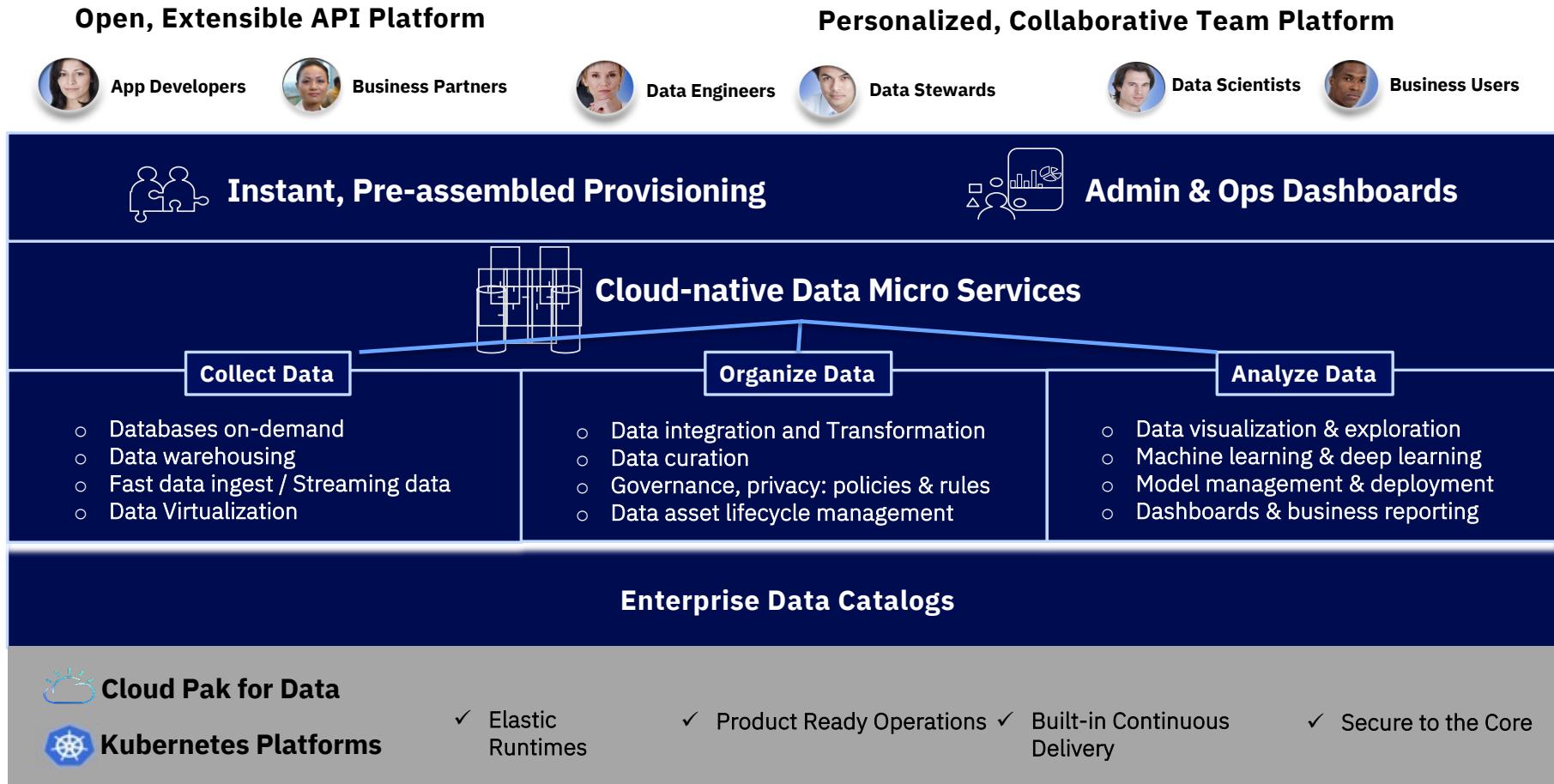
- Business reporting and visualization
- Financial planning and analysis
- Cloud native AI services
- RegTech and Financial Crimes Insight

Cloud Pak core services
Security, Administration, Operations

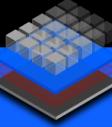
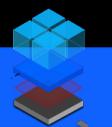
Red Hat OpenShift

IBM Cloud | Amazon Web Services | Microsoft Azure | Google Cloud | Hyperconverged system

Cloud Pak for Data



Cloud Pak for Data v4.0 Packaging

		AI Ladder Rung	Base Services	IBM Cartridges
	Collect			IBM Cartridges
Collect		Db2 Warehouse Data Virtualization Db2 BigSQL Netezza Performance Server IBM Streams Analytics Engine for Apache Spark Hadoop Execution Engine		Db2 AE/SE Informix
Organize		Watson Knowledge Catalog (including IGC) Information Analyzer (included in WKC) Data Management Console Data Privacy (Beta) IBM Match360 with Watson Guardium (Integration)		Master Data Management Product Master DataStage Information Server Knowledge Accelerators
Analyze		Watson Studio (includes Data Refinery) Watson Machine Learning (includes AutoAI) Watson Machine Learning Accelerator Watson OpenScale SPSS Modeler Decision Optimization		-
Infuse	Cognos Dashboards Embedded			Cognos Analytics Planning Analytics with Watson Watson Assistant Watson Discovery Watson Speech Services Financial Crimes Insights OpenPages with Watson Open Data for Industries Financial Services Workbench

Collect – Data management

Make data simple and accessible



Multi-Cloud

- Fast and flexible data service management
- Ease of portability across private and public clouds



Unified Console

- One experience across all data services
- Coherent, efficient, scalable data services experience
- Ease of provisioning, monitoring, and more



Comprehensive Ecosystem

- Netezza, Db2 Warehouse, AESE, and Event Store
- Partnerships with MongoDB, EnterpriseDB, and more
- Data Virtualization



Data Governance and Self-Service Analytics

- Data automatically integrated with data governance capabilities for self-service data discovery

Collect Data

1. Provision, host & manage data sources on the platform

Data sources

The screenshot shows a grid of data source cards:

- Db2 Event Store**: IBM Premium. In-memory data store capable of extremely high speed ingest and deep, real-time analytics.
- IBM Db2**: Db2 Advanced Enterprise Server Edition. IBM Premium. Relational database that delivers advanced data management and analytics capabilities for transactional and warehousing workloads.
- Db2 Warehouse**: IBM. Data warehouse designed for high-performance, in-database analytics.
- IBM Db2 for z/OS**: IBM. Create databases in Db2 for z/OS and work directly with the data from IBM Cloud Private for Data.
- Data Virtualization**: IBM Available. Query across many data sources as one.
- MongoDB**: mongoDB Partner. Scalable, open source NoSQL database.

2. Connect to Existing data sources and/or upload your data

Sample Data Sources



Add data source

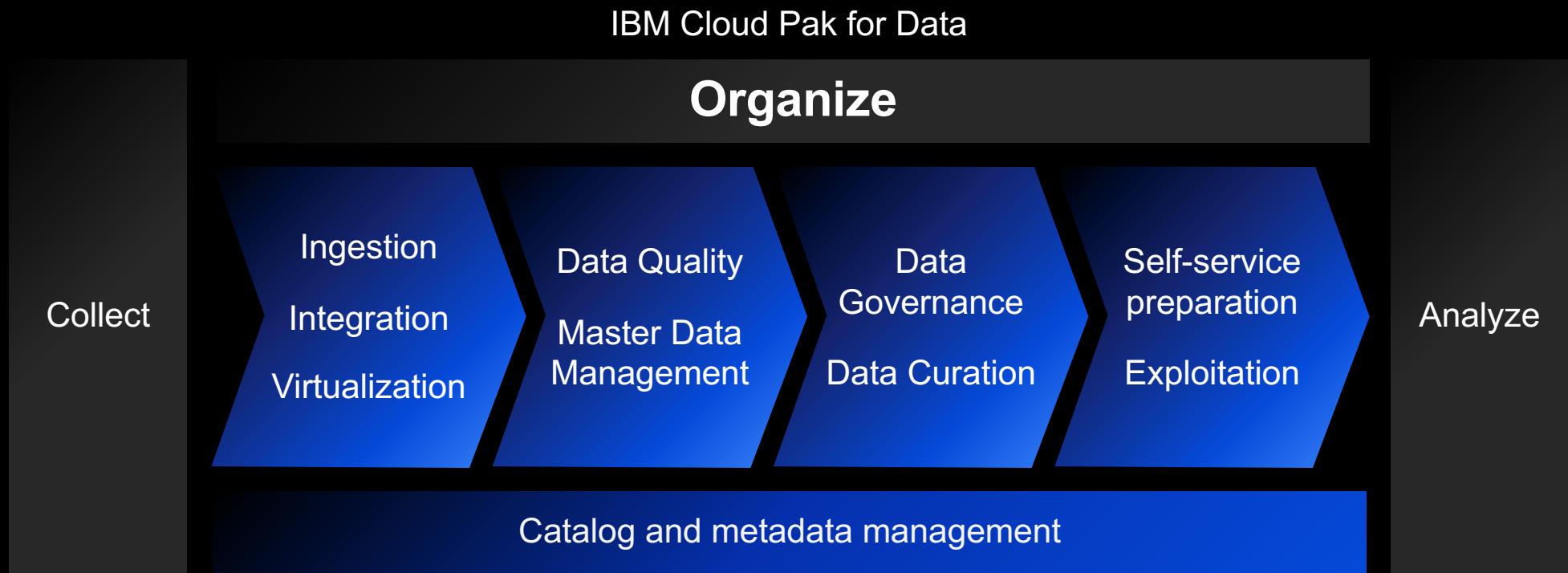
The form fields are as follows:

- Data source name ***: Type data source name here
- Description**: Type your description here
- Data source type ***: A dropdown menu is open, showing "Big SQL" selected, along with other options: dashDB, Db2, Db2 for z/OS, Hive - HDP, HDFS - HDP, and Shared.
- url here**: An input field with a red error message "here".
- here**: Another input field.
- here**: A third input field.

At the bottom right of the form is a link: [Add remote data set](#).

Organize - DataOps

Create a business-ready analytics foundation



Organize Data

Auto-discover, catalog and publish meta data, track lineage etc.

Discover assets

Connection *

db3

Schemas

Example schema[database_name|schema_name]

Browse

Select the tasks that you want to run

Profile and classify data

Assign business terms

Analyze data quality

Publish to enterprise catalog

Explore Assets
In IBM Cloud Private for Data

ASSET TYPES

Search asset types

- > Glossary and Governance
- > Databases
- > Data Files
- > Unstructured Data Sources
- > Data Science
- > Logical Data Models
- > Physical Data Models
- > XML Schema Definitions
- > Master Data Management
- > Applications
- > Files
- > Stored Procedure Definitions
- > Business Intelligence

Data exploration
In IBM Cloud Private for Data

Glossary and Governance

Term (54) Category (4) Information Governance Rule (125) Information Governance Policy (60) Collection Label (6)

Steward Data Class

Databases

Host (2) Database (2) Database Schema (7) Database Table (17) View (288) Database Column (4593)

Database Alias Stored Procedure Stored Procedure Parameter

Analyze Data

Relationships

