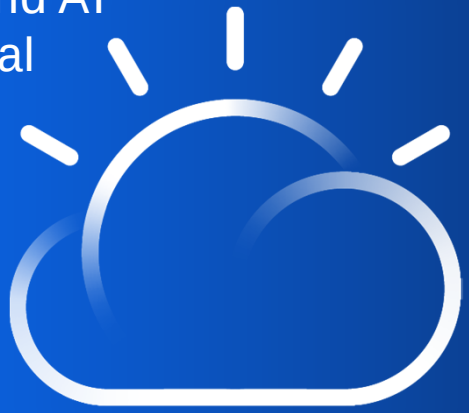
Release the breaks and accelerate your Journey to AI

Use evolving trends and lessons learnt to reinvent your Data and AI Strategy to support your Digital Transformation Journey



Wolfgang Knupp
Data and AI Architect
IBM South Africa

IBM Cloud

Data and AI Strategy

Quote:

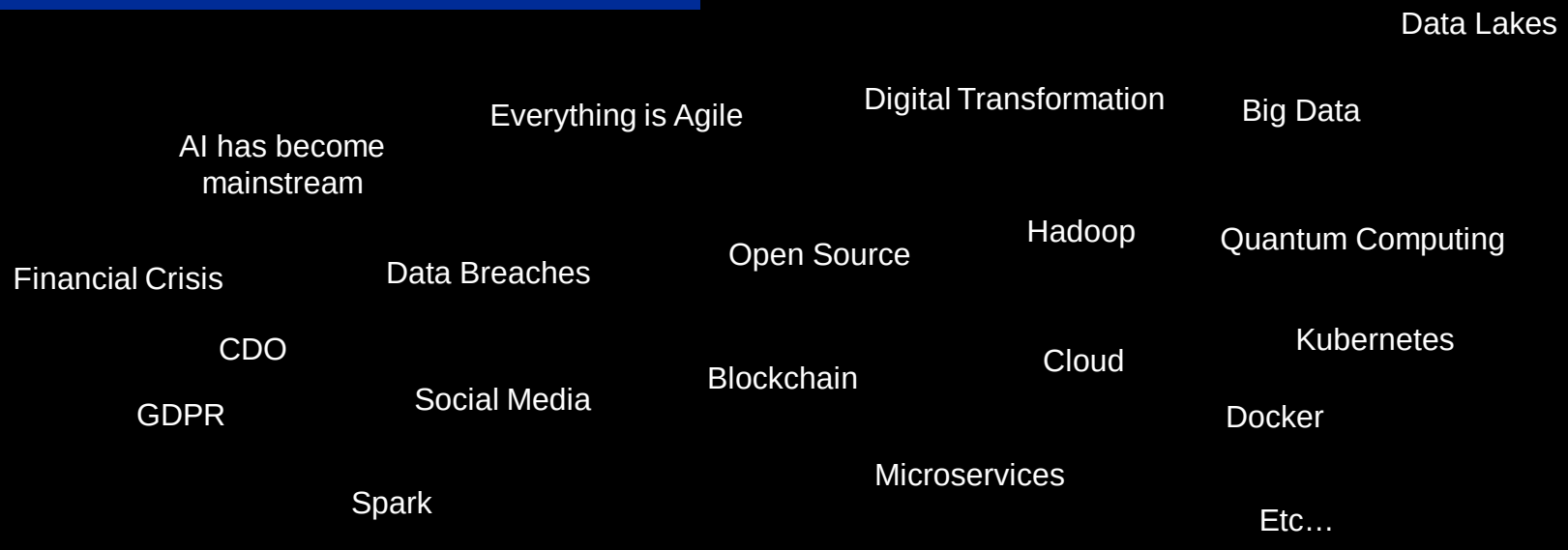
"Not having a data strategy is analogous to allowing each person within each department of your organization to develop their own chart of accounts and use their own numbering scheme. "

Sid Adelman, Data Warehousing Expert (Adelman, Moss, & Abai, 2005)

Data and AI Strategy

So, what has changed in the Data and AI world since 2005 ?

Not much ?... except...



Data and AI Strategy

Your Data Strategy is intended to support your Business Initiatives.

and must :

- Be owned by Business (Chief Data Officer)
- Align with the Business Strategy
- Align with changes in technology
- Be revised regularly
- Include applicable Regulations
- Include Ethical Guidelines for AI
- Encourage ambitious and achievable goal setting (be testable / defensible)
- Include the standard elements as in 2005
    - Operational Model
    - Data Governance
        - Data Ownership
        - Common Business Glossary of Terms, Policies and Rules
        - Security, Data Privacy, Lifecycle, Data Quality, etc...
- CDO, Data and AI Reference Architecture, DevOps, Cloud

Data and AI Strategy

Your Data Strategy is intended to support your Business Initiatives.

Lessons learnt:

- Your data strategy must reflect your own Business Strategy
  - not be dictated by vendors or consulting houses
- Best of Breed technology is too expensive
  - time to value too long
  - cost of skills to integrate adjacent best of breed technologies
- Culture needs to adapt -> Change management is vital
- IT overload -> Business User Data / Insights Self Service to encourage innovation, fail fast

- And last but not least... There is No AI without IA

There is no AI without an IA

(information architecture)

80%

of data is either inaccessible, untrusted or unanalyzed

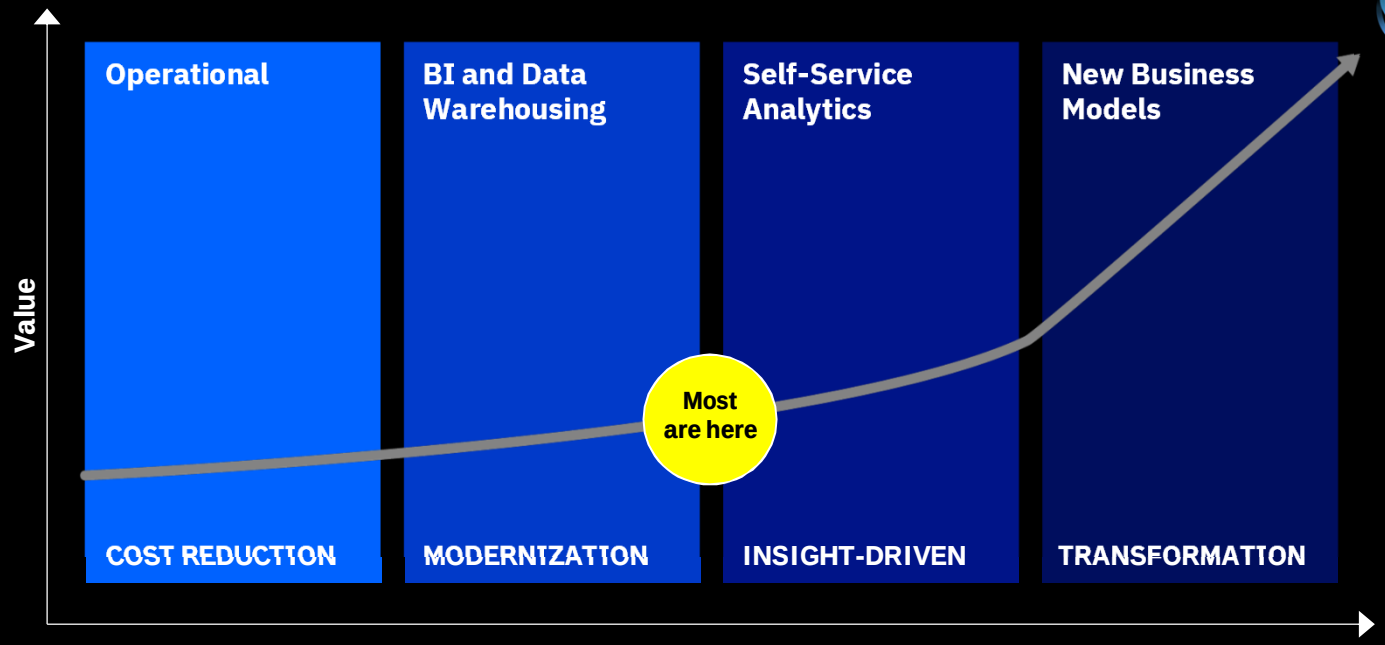
81%

do not understand the data required for AI

> *"No amount of AI algorithmic sophistication will overcome a lack of data [architecture] ... bad data is simply paralyzing*

MITSloan Management Review

Clients are advancing their use of data for business value



83%
view AI as a strategic opportunity

MITSloan Management Review

Barriers to successful AI transformations

There is no Artificial Intelligence (AI) without Information Architecture (IA)

Data Ecosystem

- Data in silos
- Difficult to access
- No lineage

Workflow

- Not integrated
- Not governed
- Lack dev/prod parity

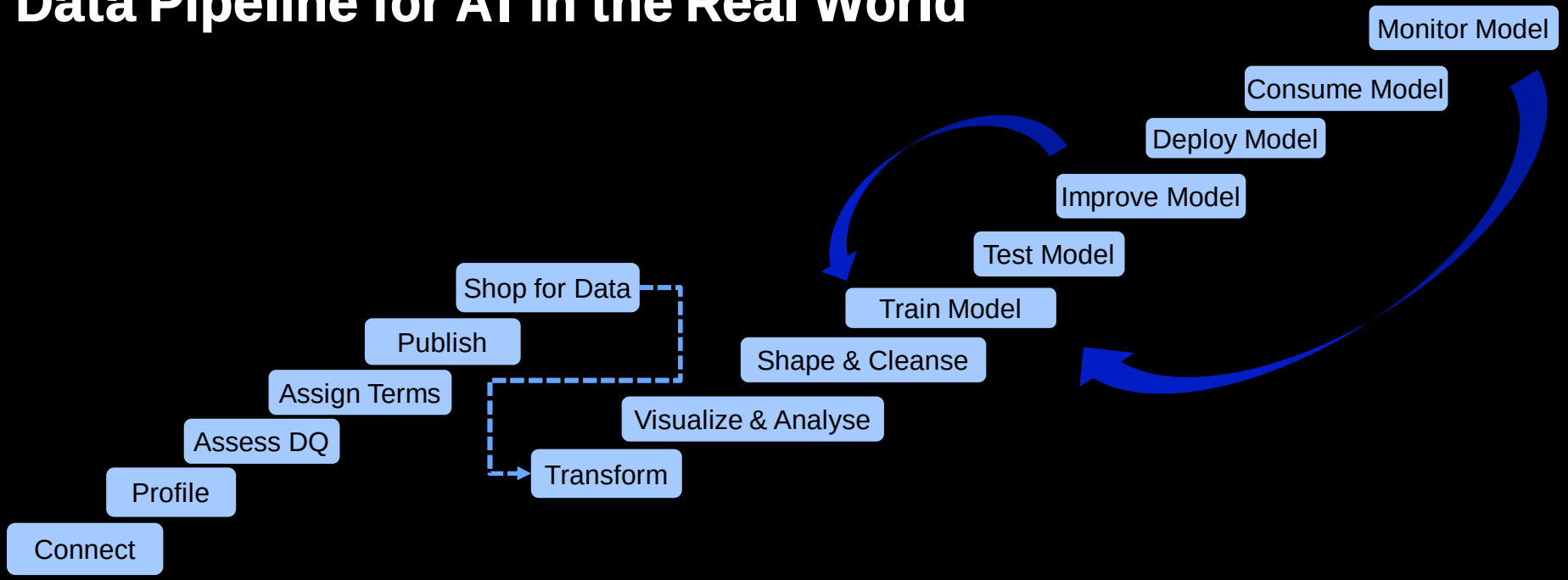
Analytics Tools

- Discrete tools
- Different preferences
- Difficult to manage

Culture

- Not collaborative
- Slow provisioning
- Lack trust in AI

Data Pipeline for AI in the Real World



The AI Ladder

A prescriptive approach to accelerating your journey to AI

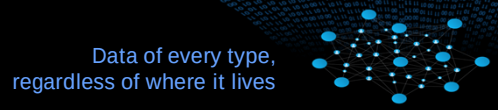
INFUSE - Automate and scale across your processes

TRUST - Achieve trust and transparency in outcomes

ANALYZE - Scale insights with ML everywhere

ORGANIZE - Create a trusted analytics foundation

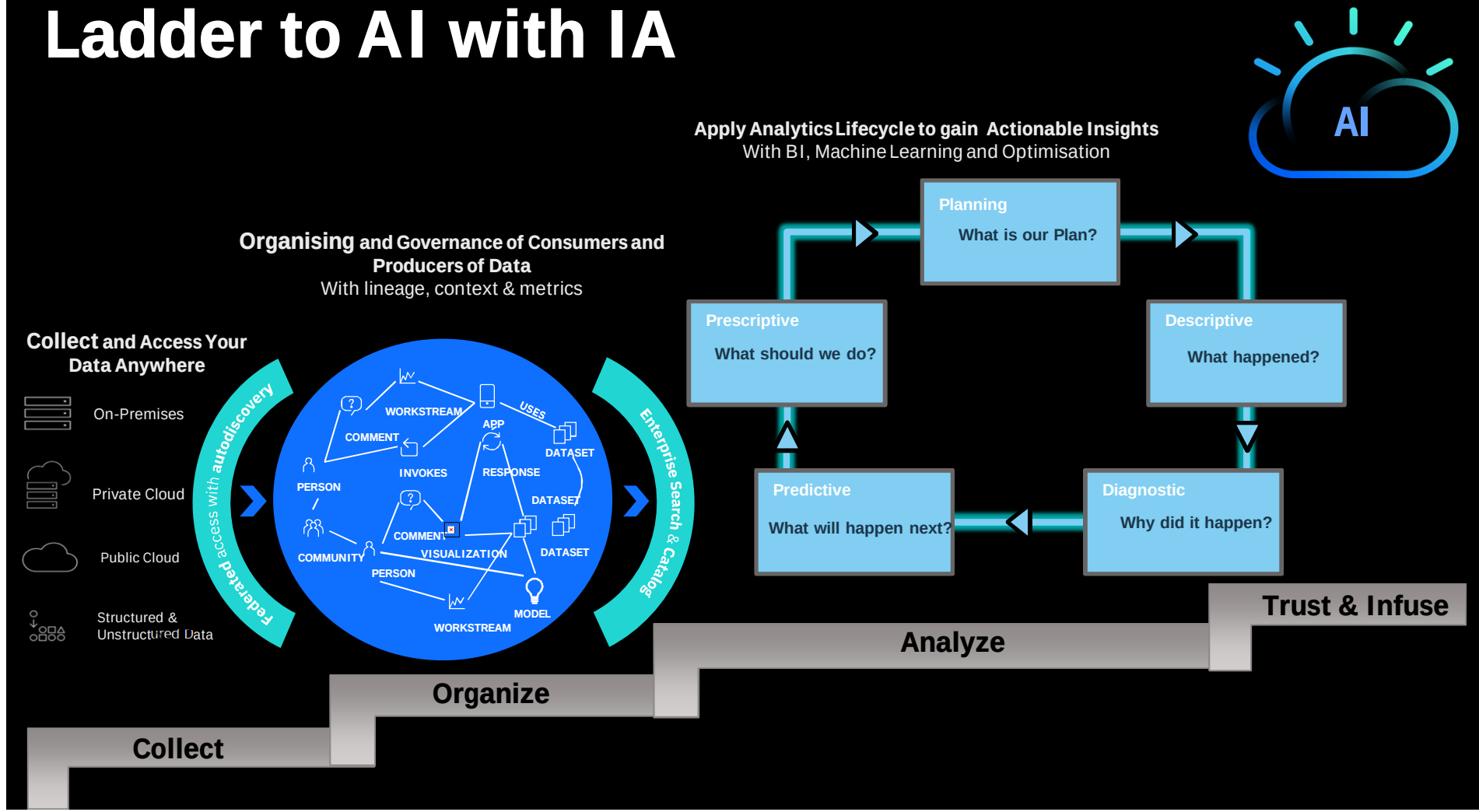
COLLECT – Simplify access to data



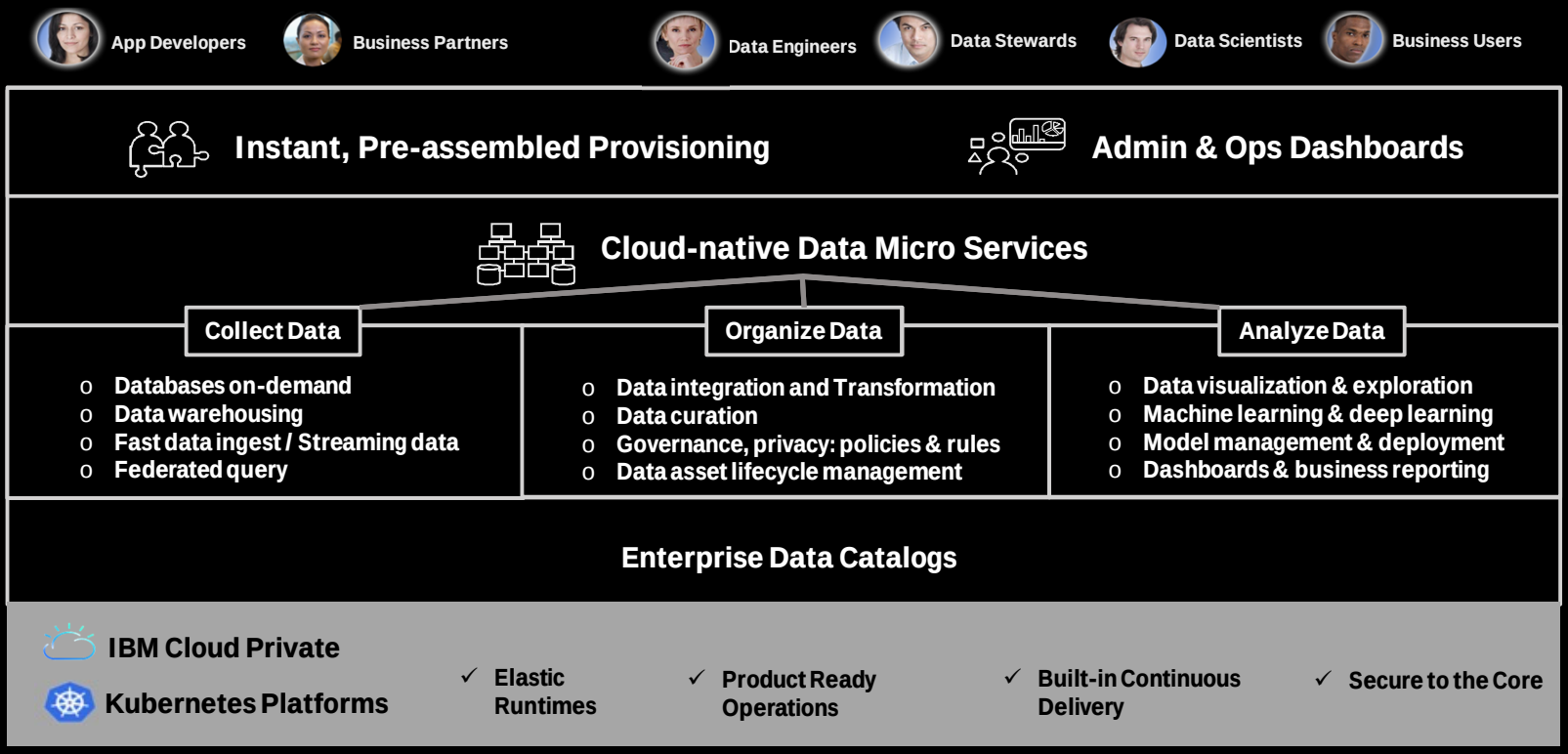
Data of every type, regardless of where it lives

MODERNIZE your data estate for an AI and multicloud world

Ladder to AI with IA

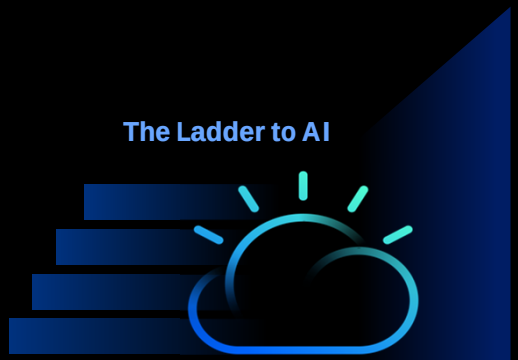


IBM Solution to accelerate the Journey to AI:
IBM Cloud Private for Data



IBM Cloud Private for Data

Unify on an open, multicloud Data & AI platform



Extensible “add-ons” from IBM’s Data & AI Portfolio & 3rd Party Services

IBM | IBM | IBM | IBM | IBM | ISV | OS | Custom | ...

Collect Data

- Data virtualization
- Data warehousing
- Databases on-demand
- Data source ingestion
- Distributed processing

***Powered by**: Db2 Family*

Organize Data

- Discovery & search
- Data transformation
- Data cataloging
- Business glossary
- Policies, rules & privacy

***Powered by**: InfoSphere*

Analyze Data

- AI Machine learning learning
- AI Model build & deploy
- AI Model management
- Dashboards & reporting
- Analytical visualization

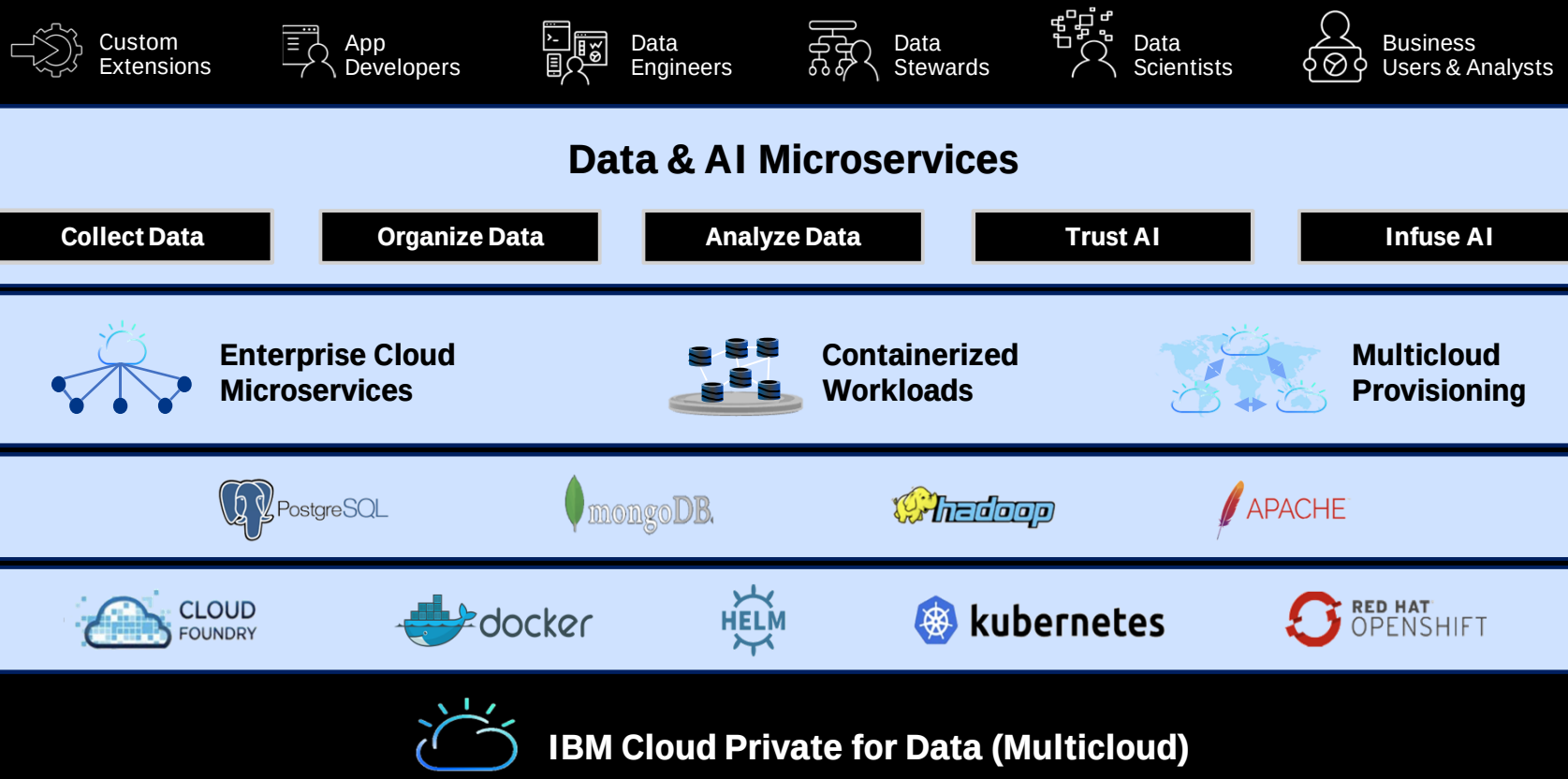
***Powered by**: Watson & Cognos*

Multicloud Services

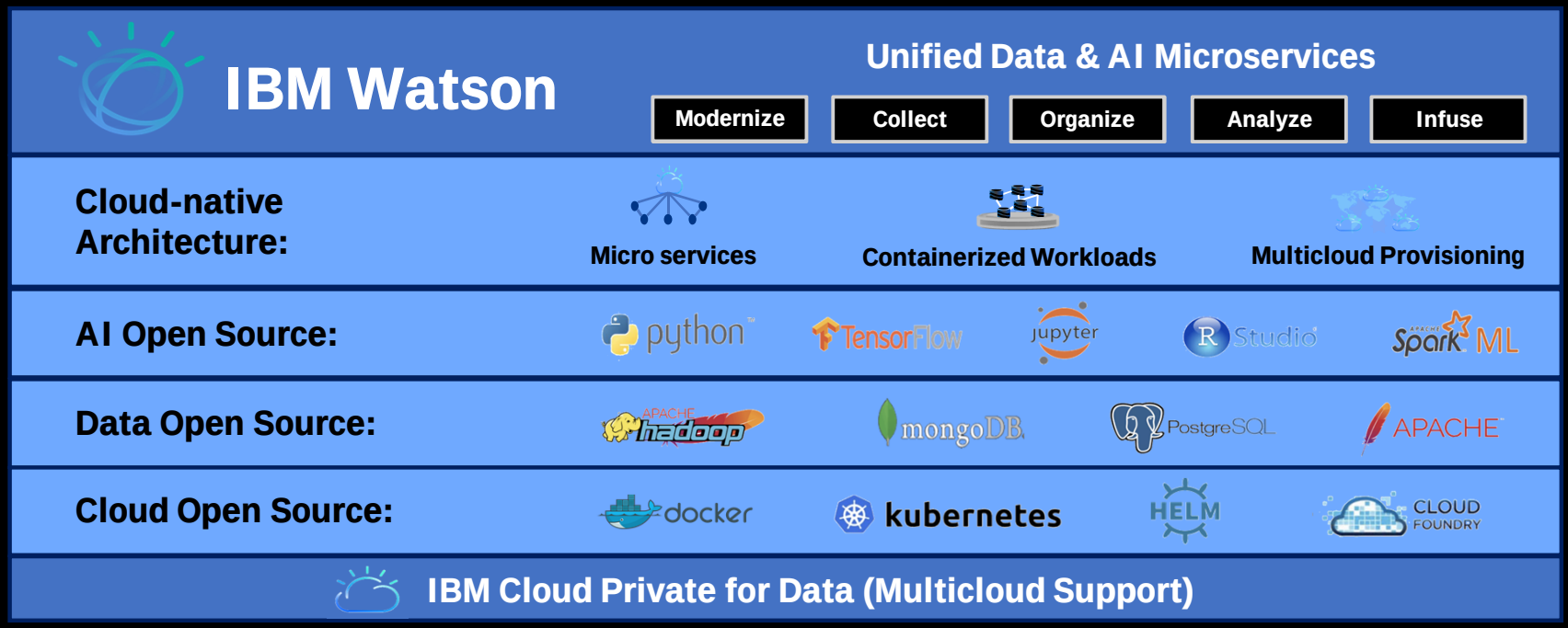
- Logging
- Monitoring
- Metering
- Persistent Storage
- Kubernetes
- Security
- Identity Access Mgmt.
- Docker Registry / Helm

IBM Cloud | RED HAT OPENSHIFT | amazon web services | Azure | openstack | Google Cloud

1) Deploys an Open Information Architecture for AI



Open Source Meets Multicloud, Working as One



IBM Cloud | RED HAT OPENSHIFT | amazon web services | Azure | openstack | Google Cloud

2) Makes your data ready for AI

Deploy a governed platform unifying team workflows and AI models with the data they rely on

- Automates the steps on the AI ladder
- Prepares and governs business-ready data for use in ML & AI models
- Built-in machine learning design, creation, train and deployment

Collect Data

Collect data of every type, regardless of where it lives,

Organize Data

Organize your data into a trusted, business-aligned analytics foundation

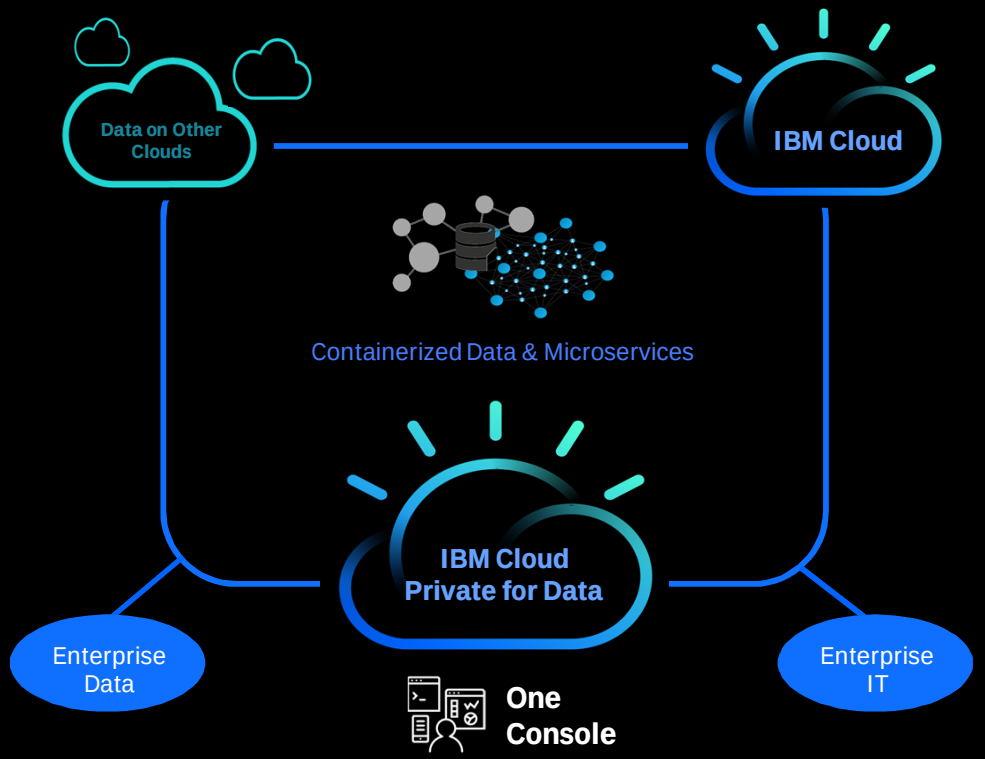
Analyze Data

Analyze your data in smarter ways, with Machine Learning everywhere

3) Modernizes data for a multi-cloud world

Leverage Cloud-native services to modernize and manage data in multicloud environments

- One integrated console for collecting, organizing and analyzing data
- Virtualize all your data of every type, regardless of where it lives
- Build once, run anywhere across dynamically changing environments



Watson on IBM Cloud Private for Data

Extensible Add-on AI Services

Watson Studio & Machine Learning

Data Science Premium
Tools to design, build and deploy AI models


SPSS Modeler


Data Refinery


Model Builder


Decision Optimization


Watson Explorer


Streams Designer

Watson OpenScale ▪ *Operational optimization services*

Watson Assistant ▪ *Conversational AI services*

Watson APIs ▪ *Interactive AI services (e.g., speech)*

Foundational Data & AI Services

Cloud Private for Data

Open Source frameworks to build, train and deploy AI models powered by **Watson Studio** and **Machine Learning**

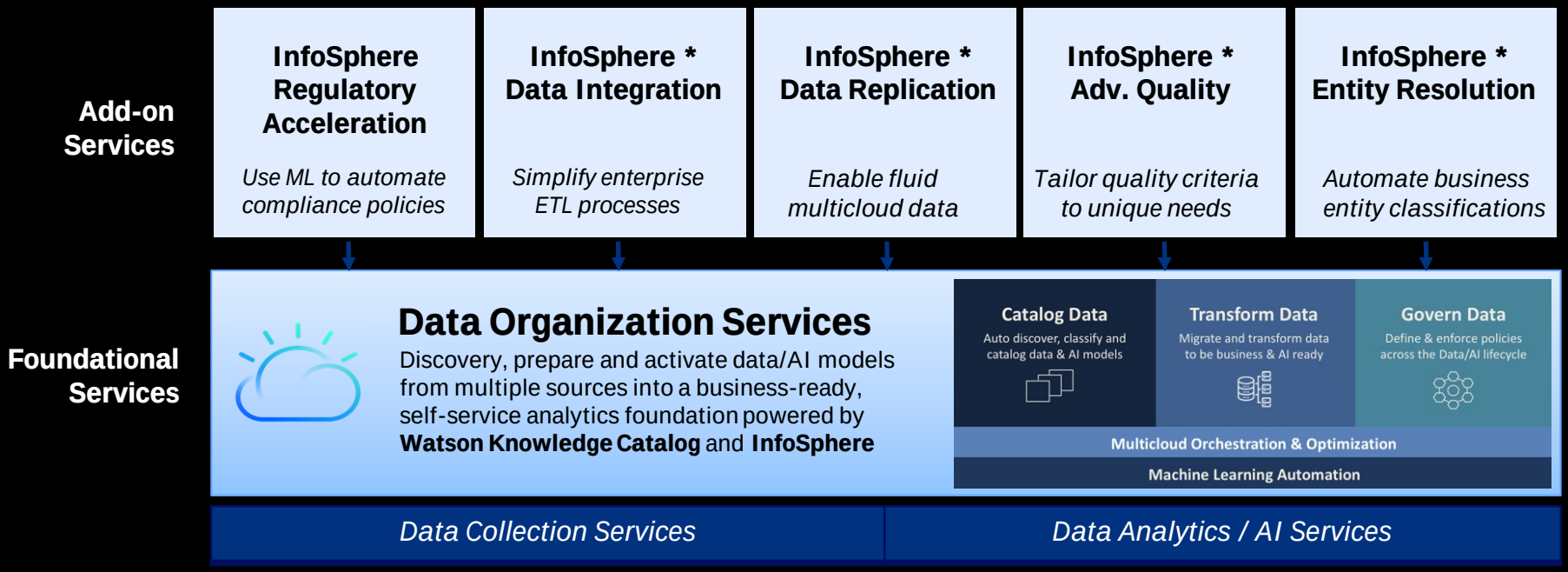

- **Data visualization**
- **Machine learning learning**
- **Model build & deploy**
- **Model management**
- **Dashboards & reporting**

Data Collection Services | *Data Organization Services*


IBM Cloud


Google Cloud

Unified Governance & Integration on Cloud Private for Data



IBM Cloud · RED HAT OPENSHIFT · amazon web services · Azure · openstack · Google Cloud

* On Roadmap Q2/Q3 2019

Thank you !

Please visit the IBM Stand to see how

IBM Cloud Private for Data

Can help you to accelerate your Journey to AI

IBM Cloud Private for Data

Top 4 Use Cases

Use Cases

 **Manage your Data Anywhere**	 **Operationalize Data Science & AI**	 **Shift to Next-Gen workloads**	 **Smarter Governance**
1. Manage all your enterprise data regardless of where it lives (*Data Virtualization*) **2. Gain control & leverage your data from connected devices** (Fast data & *Streaming analytics*)	**Build, deploy, manage & govern models & data @ scale to improve business outcomes** **(see following slides for details)** e.g. a. Customer Churn b. Cross Sell / Up Sell c. Predictive Maintenance	**Shift to Cloud Native** a. Provision & scale in minutes b. Build once, deploy anywhere – multi cloud support c. Built in automation & collaboration to increase productivity	**Governance to enable self service analytics** Auto-discover meta data, manage governance rules & policies, enforce privacy etc. to mitigate risk & ensure compliance e.g. accelerate GDPR Compliance

Collect Data

Solution Overview

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

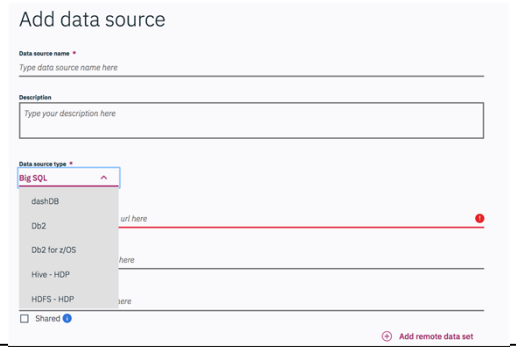
Data sources

- **Db2 Event Store** — IBM, Premium
  In-memory data store capable of extremely high speed ingest and deep, real-time analytics.
- **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** — Partner
  Scalable, open source NoSQL database.

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

Sample Data Sources

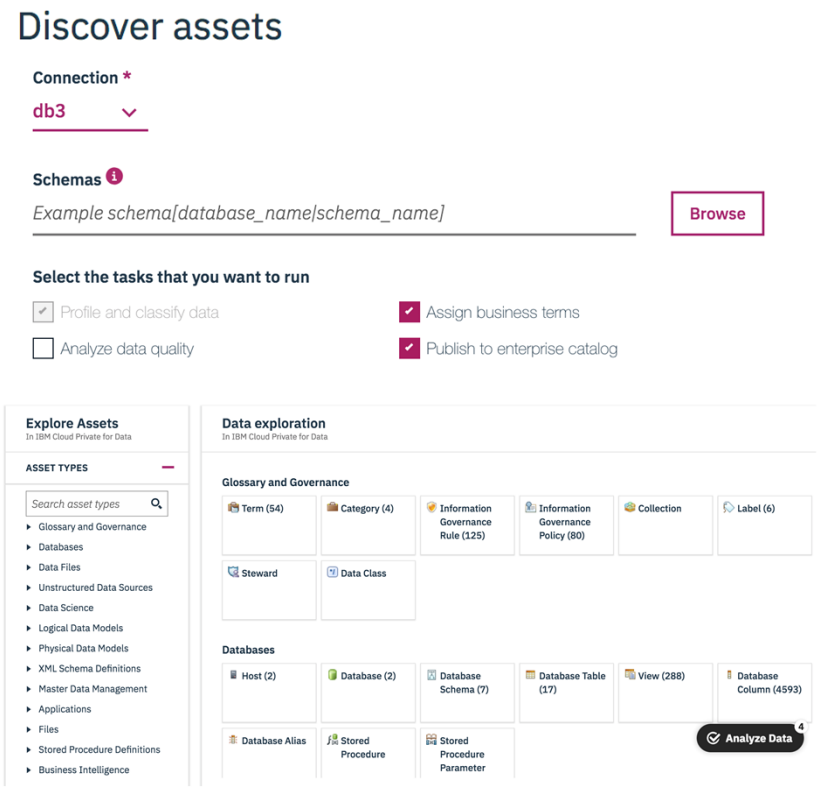
- Db2
- HDFS
- Custom JDBC
- Cloudera
- Oracle
- Teradata
- Hortonworks
- Netezza

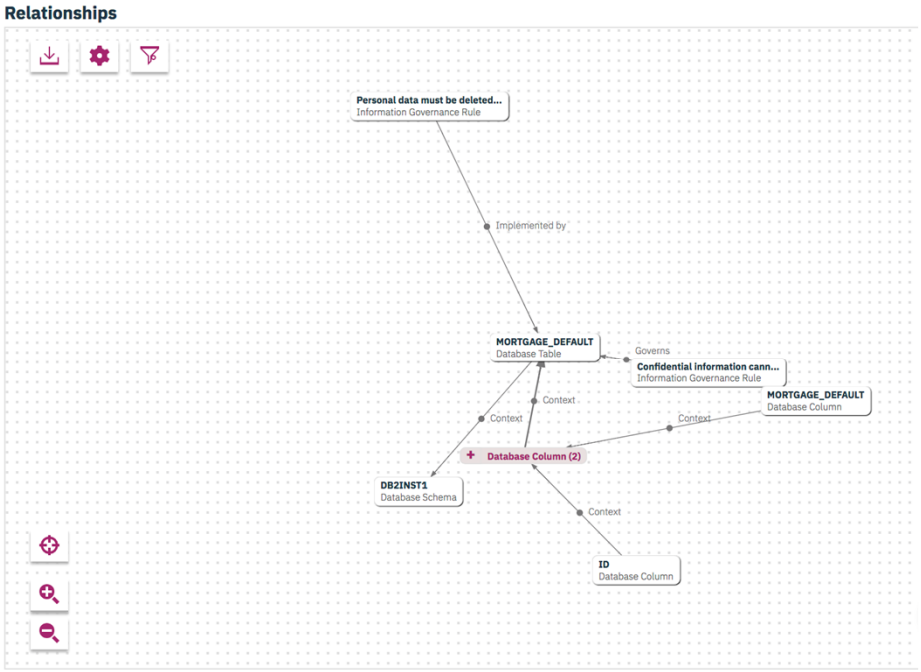


Organize Data

Solution Overview

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

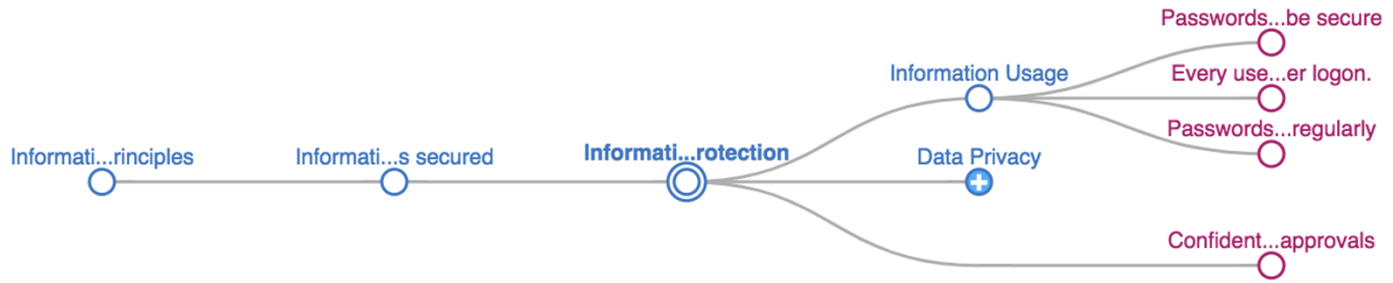


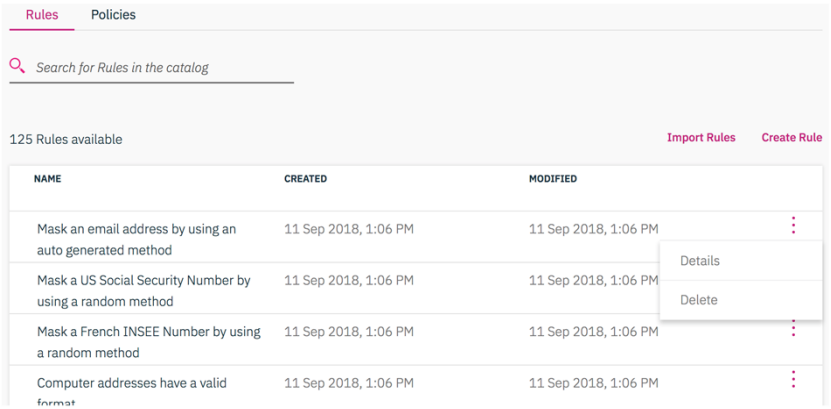


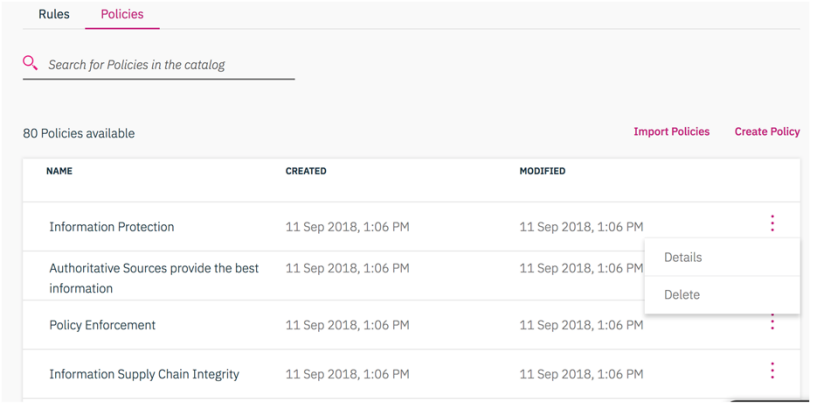
Organize Data

2. Solution Overview

2. Define & Enforce Governance Policies / Rules



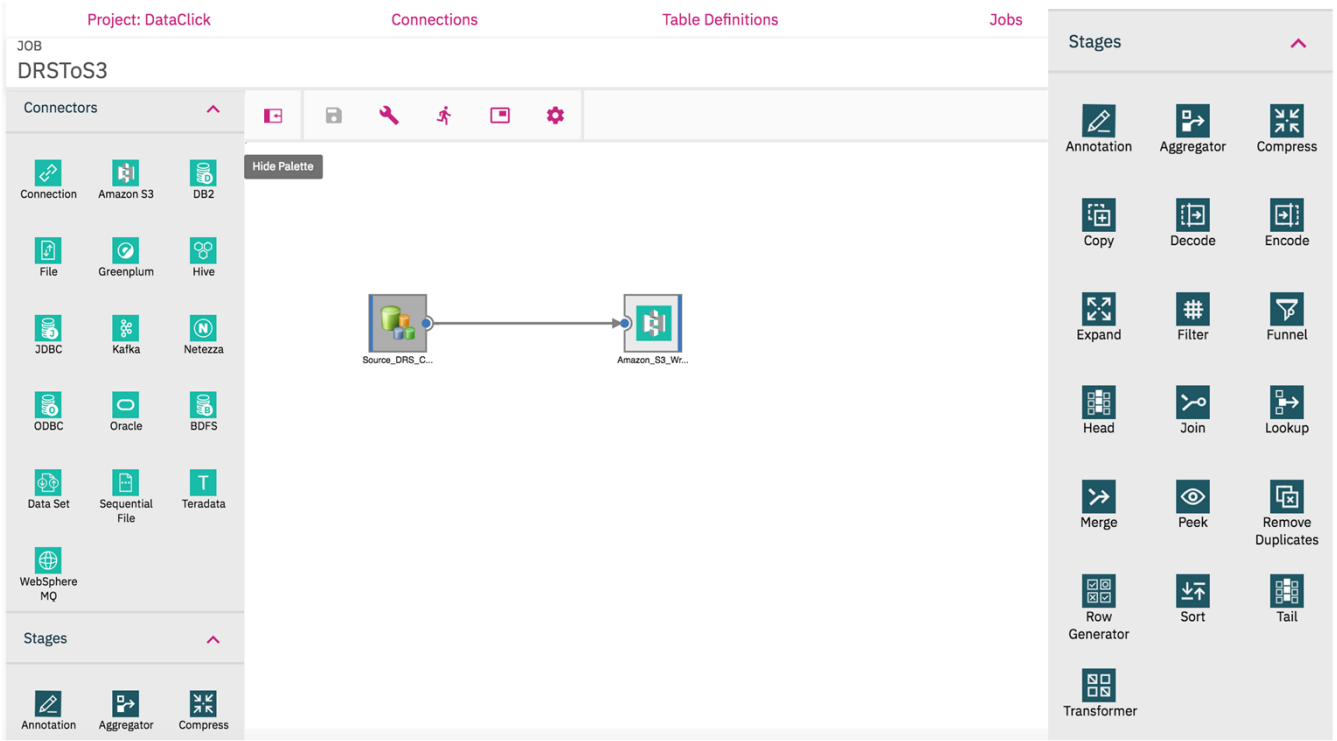




Organize Data

Solution Overview

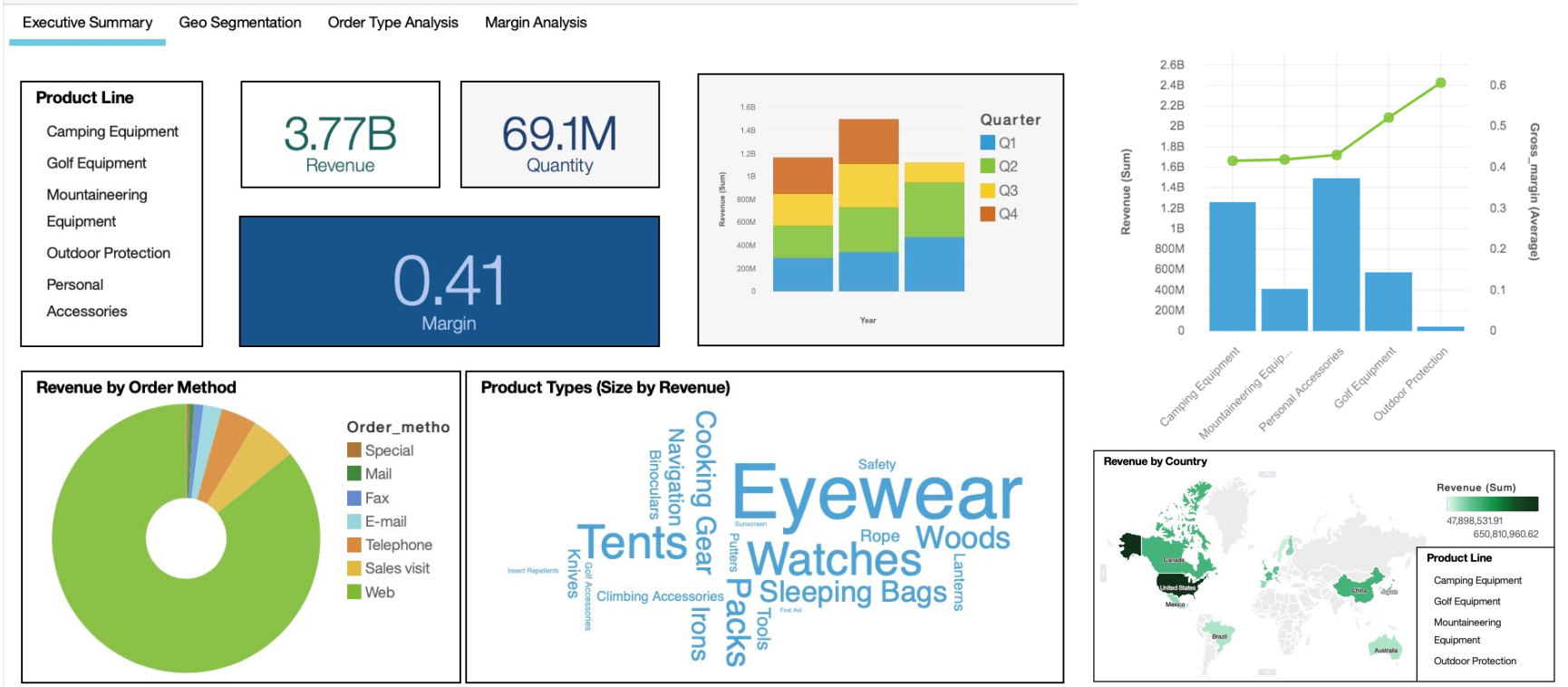
3. Transform & Migrate Data – Build & Execute ETL jobs @ scale



Analyze Data

Solution Overview

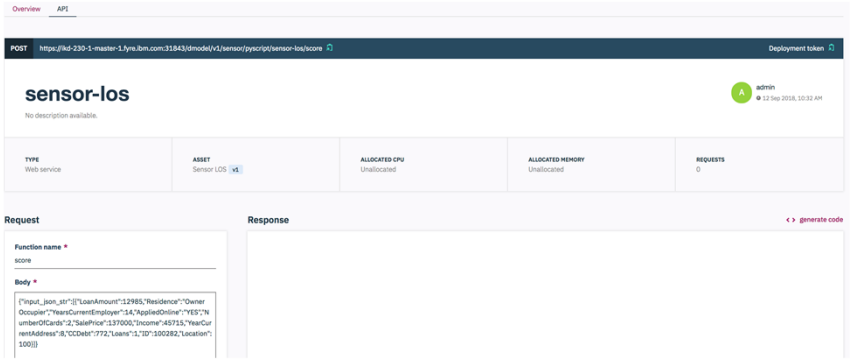
4. Build Visualizations & Dashboards

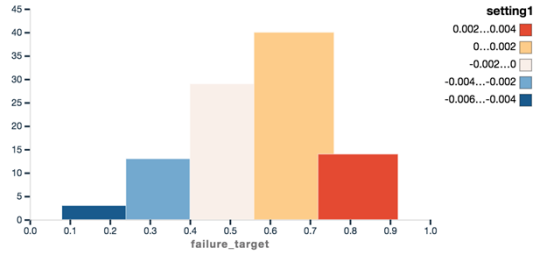


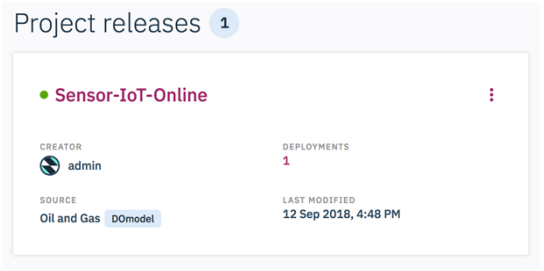
Analyze Data

Solution Overview

5. Build , Deploy and manage Data Science & Machine Learning Models



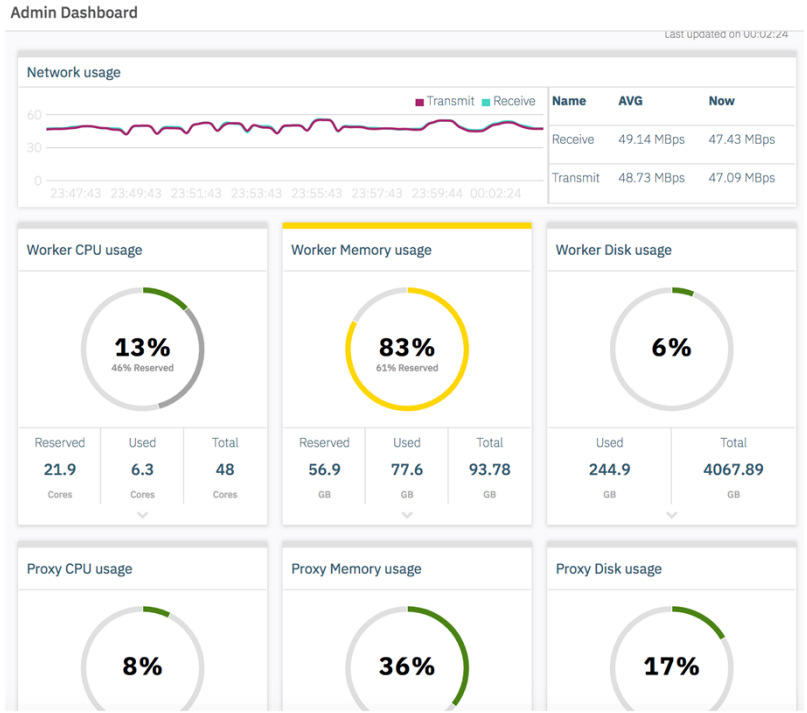


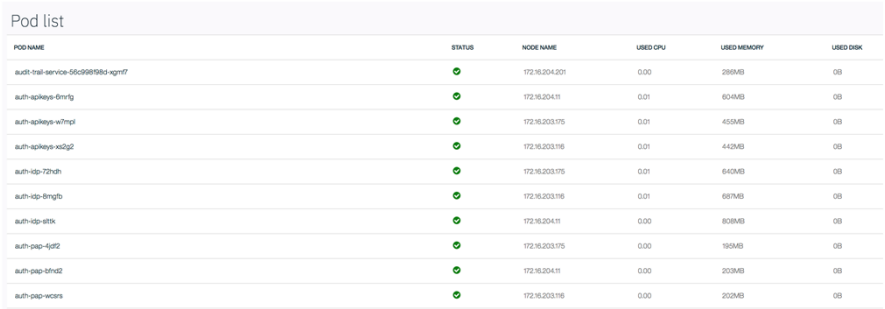


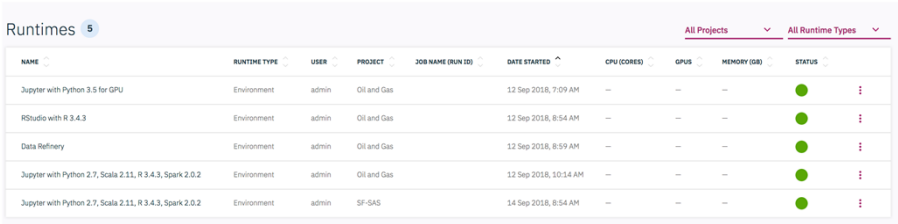
Administer

2. Solution Overview

6. A single Pane to administer & monitor your data & analytics workloads

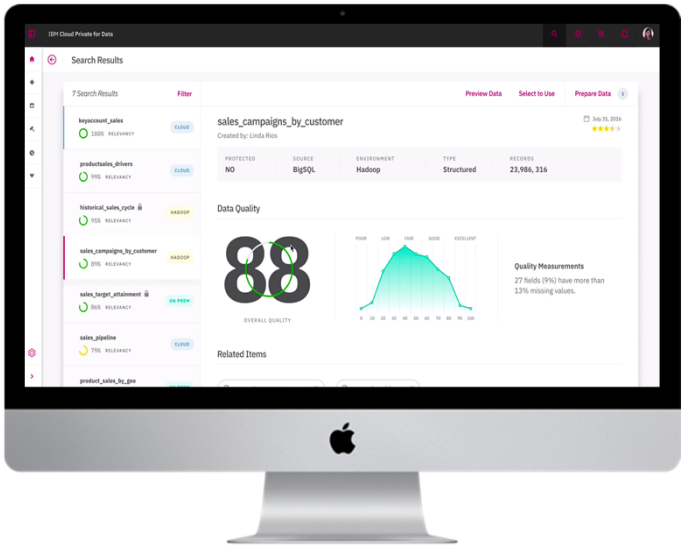






Product Demo

Solution Overview



http://ibm.biz/experienceICP4D