

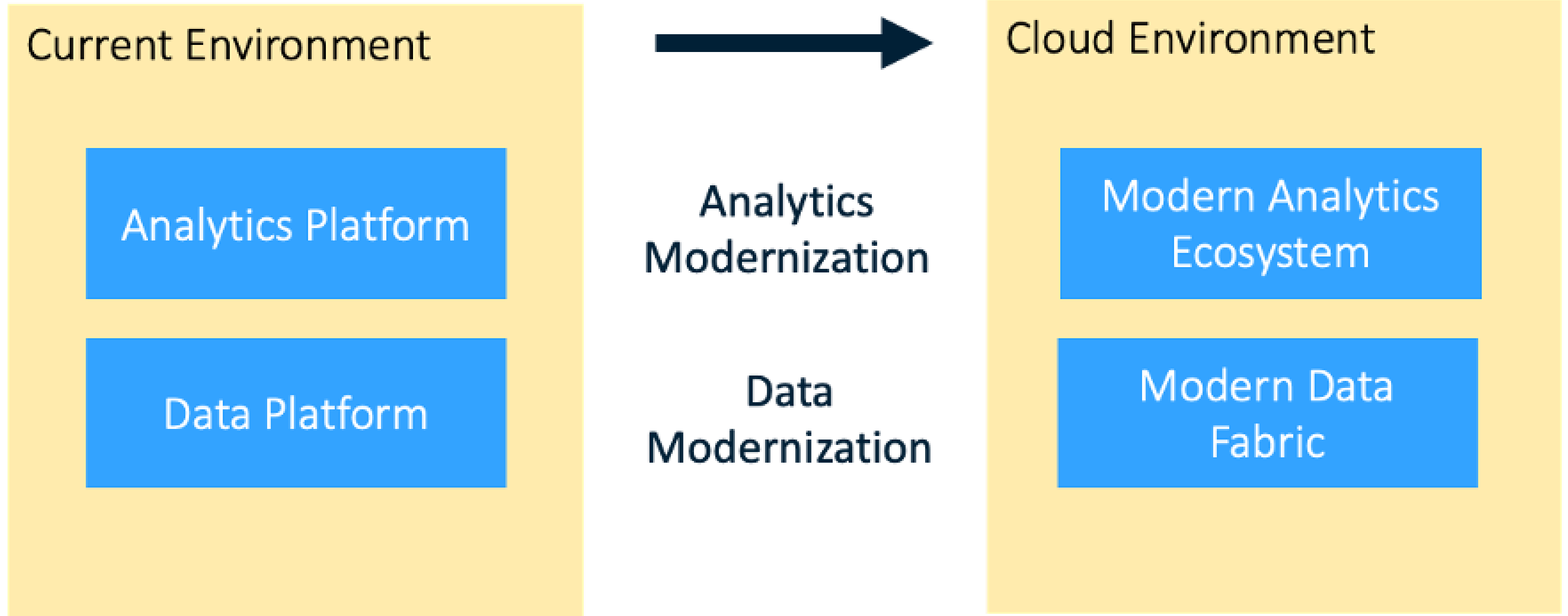
# Governance and data quality while moving to the cloud with SAS Viya

FANS Kristiansand

Lars Arne Skår, solution architect

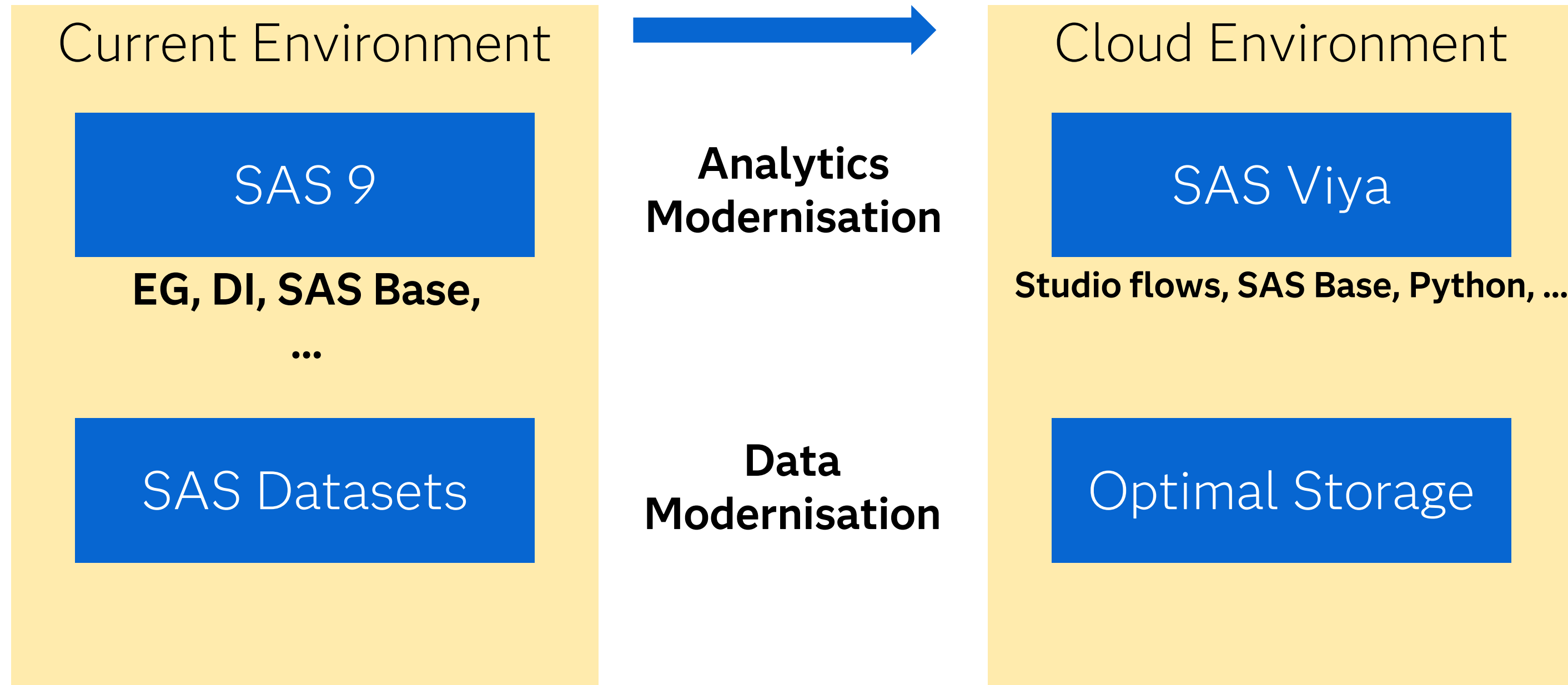


# Moving to Cloud



Source: Jeff Stander @ SAS

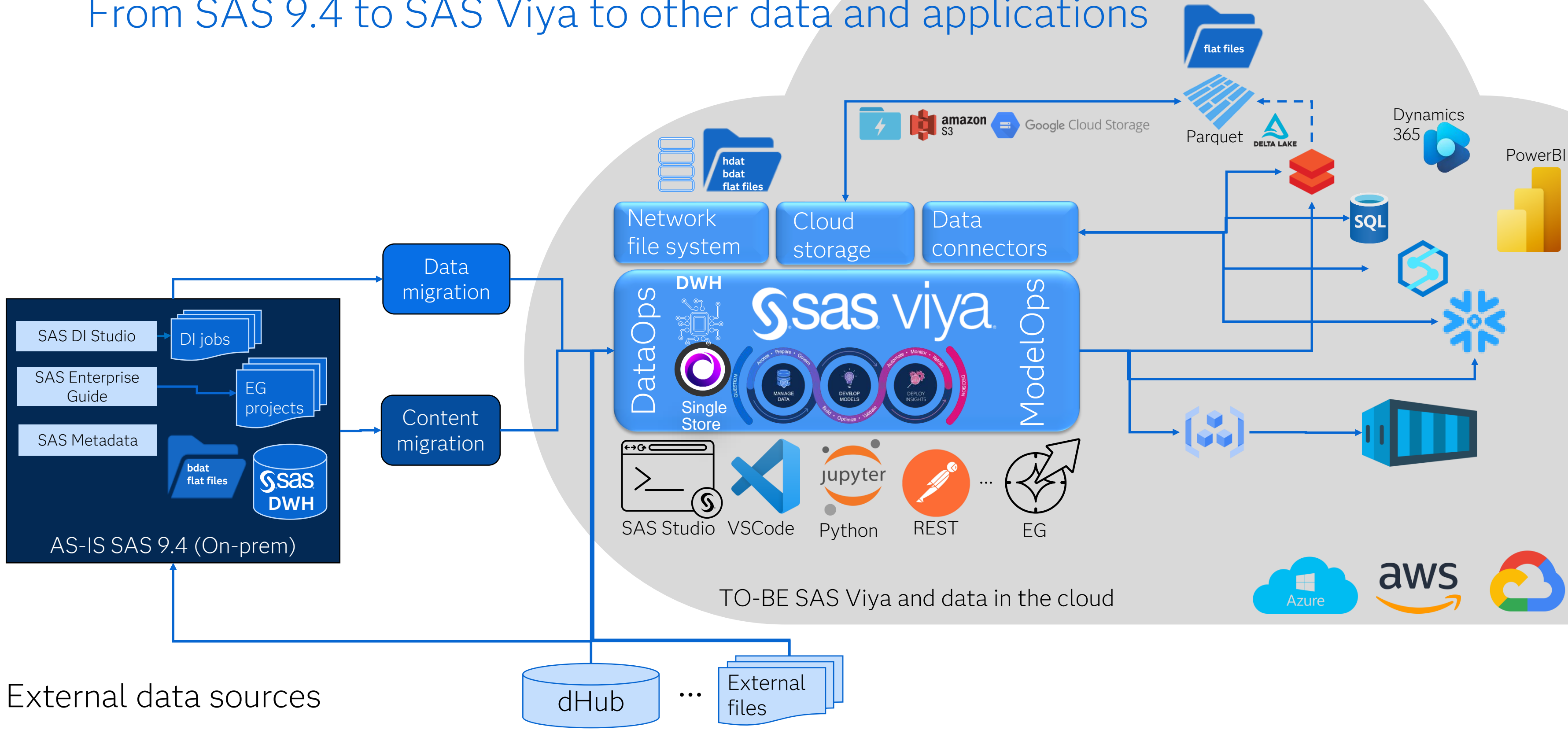
# Moving SAS to the cloud



Source: Jeff Stander @ SAS

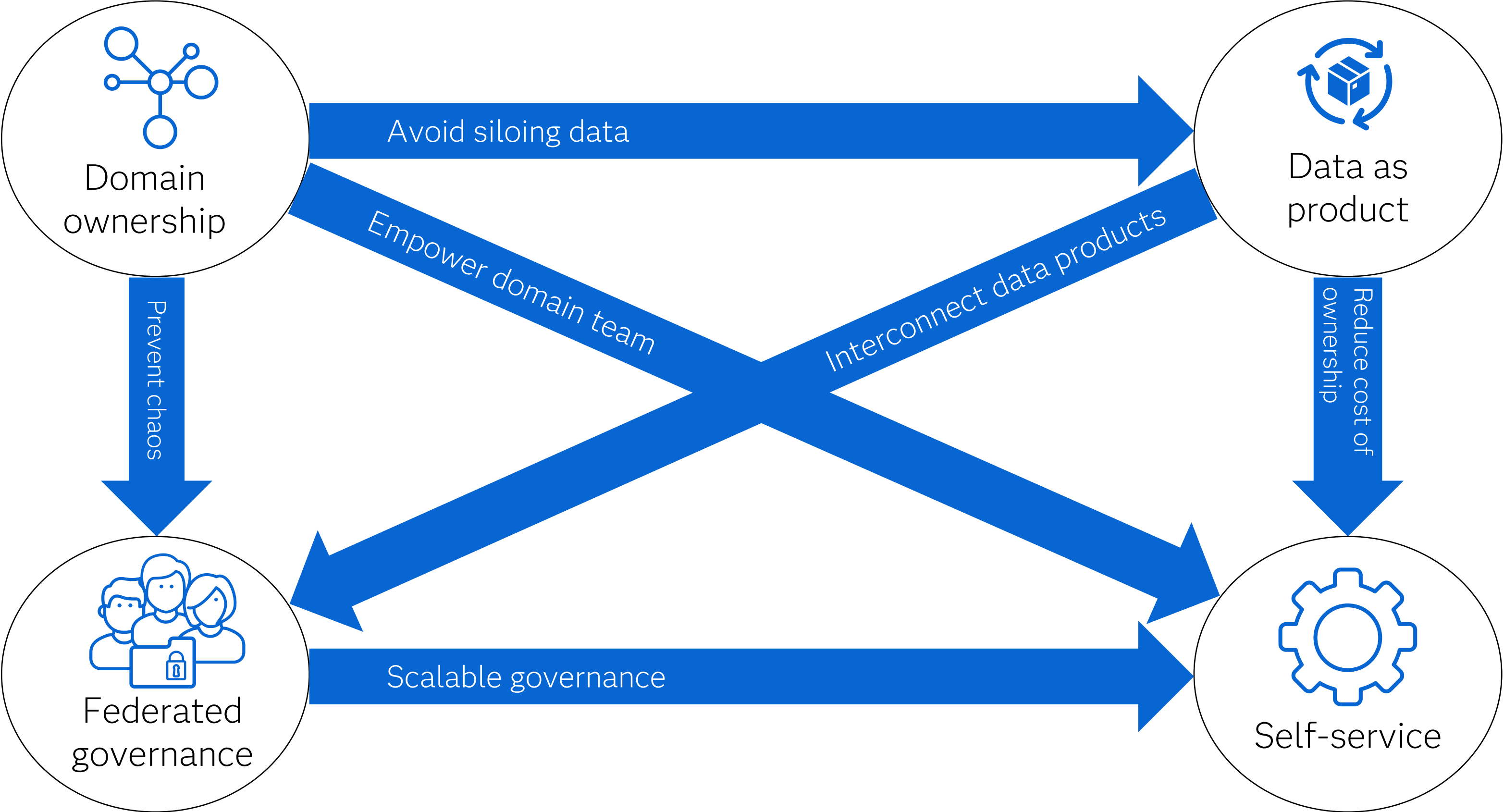
# Data, analytics and AI architecture to the cloud

From SAS 9.4 to SAS Viya to other data and applications



External data sources

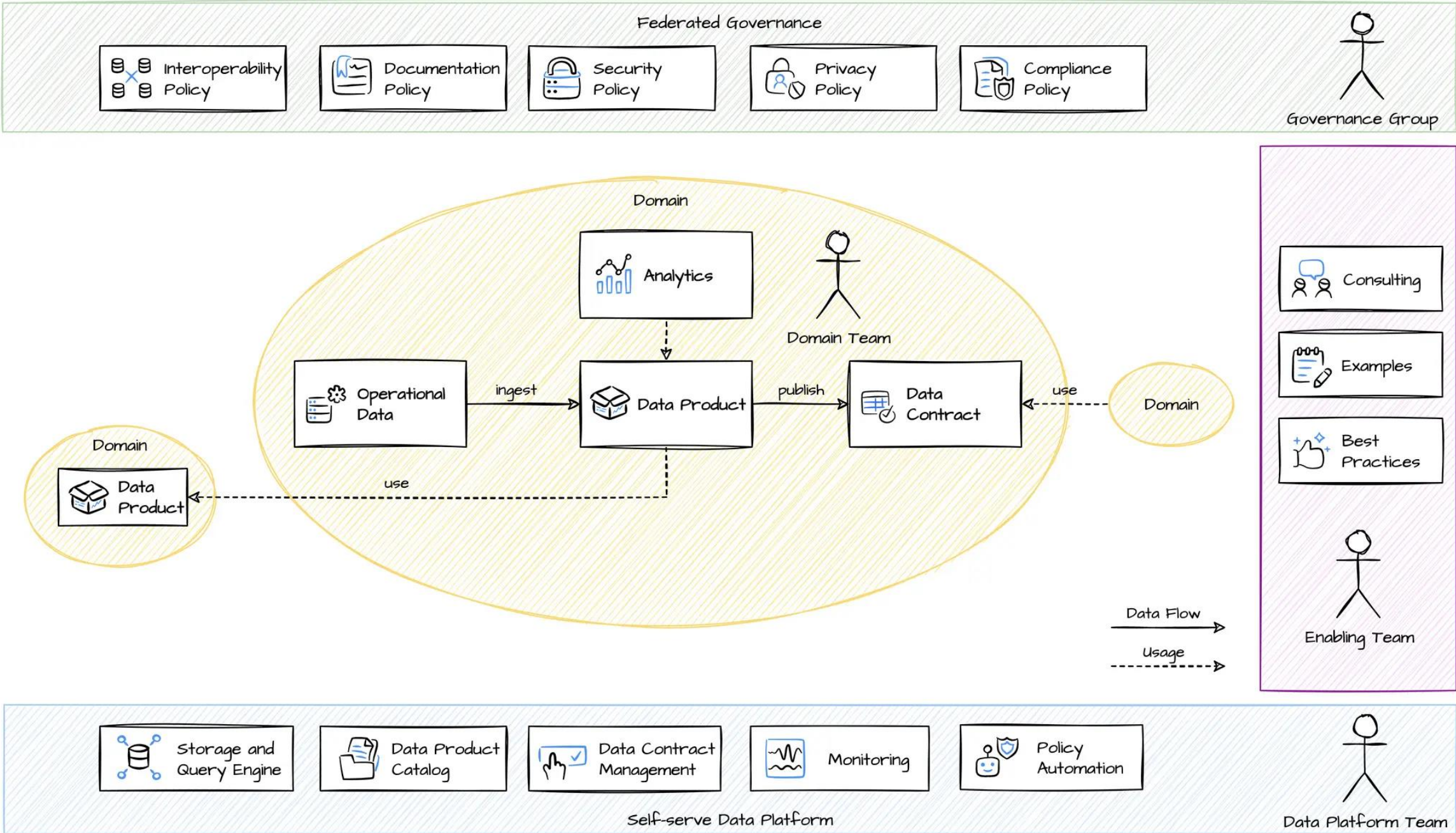
# Data mesh – what is it and why do we want it?



# Key capabilities from SAS for Data Mesh

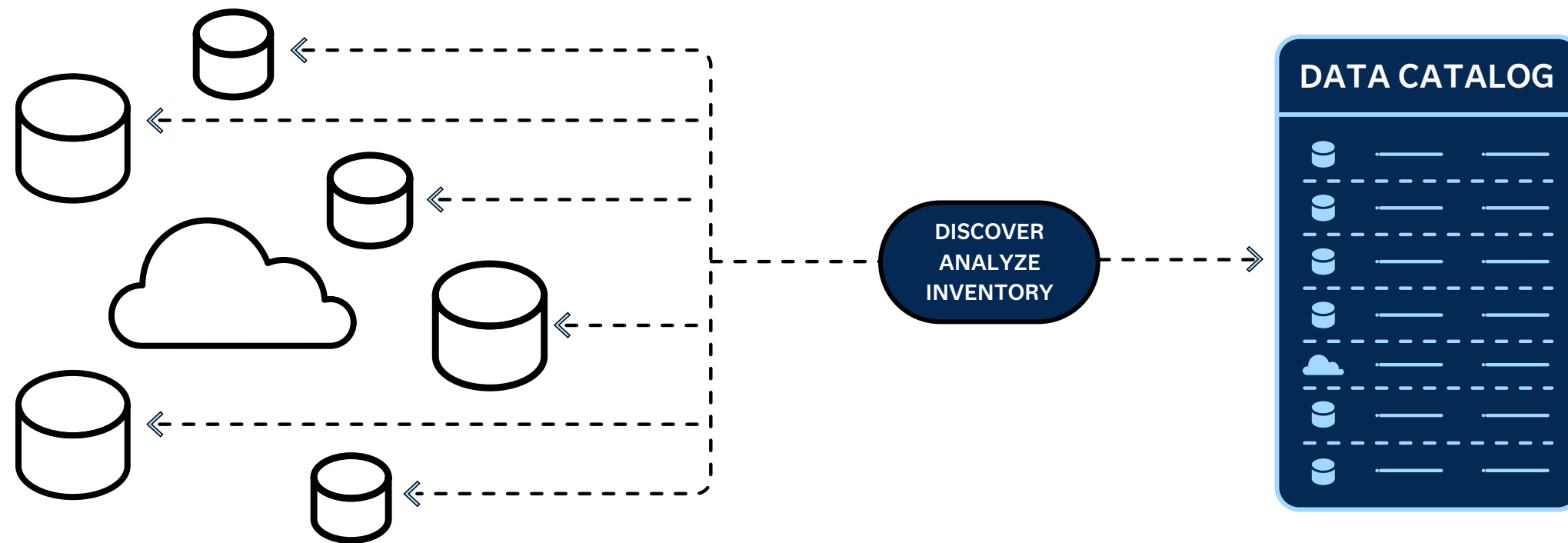
Federated governance, data access on-prem and cloud, self-service

Data Mesh Architecture



datamesh-architecture.com

- **Federated Governance** through information catalogue and data lineage as well as embedded security enforcement in the platform
- **Self-service data** flow with SAS Studio flow and enablement through custom steps and popular languages for analytics and data prep (SAS, Python and R)
- **Operational and analytics processing** (in-DB, container, in-memory, on the edge, events)
- Providing **access to data** through external data sources, singlestore or REST-based access through MAS
- **Storage and query engine** through our access and connect modules and compute/CAS
- Data contract management through our **information catalogue and business glossary**
- **Data exploration and visualization** on metadata and content
- **Data product catalog** with our information catalogue using our discovery agent
- **Monitoring and policy automation** through DQ monitoring with business rules using QKB and discovery agents



# SAS Data Engineering Platform Overview

One single centralized platform for Data Citizens

Data Preparation

Data Quality

Data Profiling

Data Matching

Data Survivorship



Data Monitoring

Data Stewardship

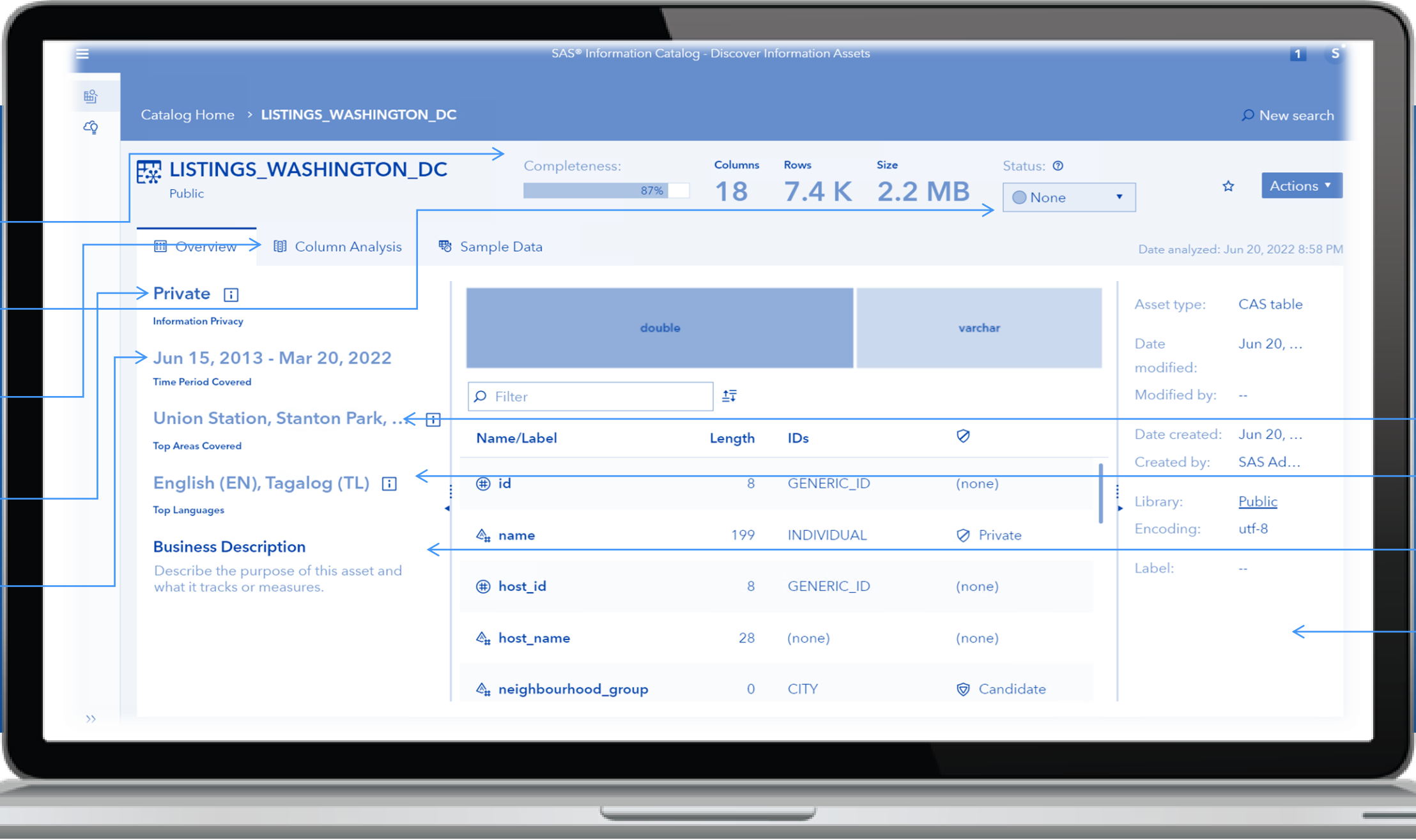
Data Governance

Pre-built Rules(QKB)





# SAS Information Catalog



Metrics

Status

Stats

Information Privacy

Time interval

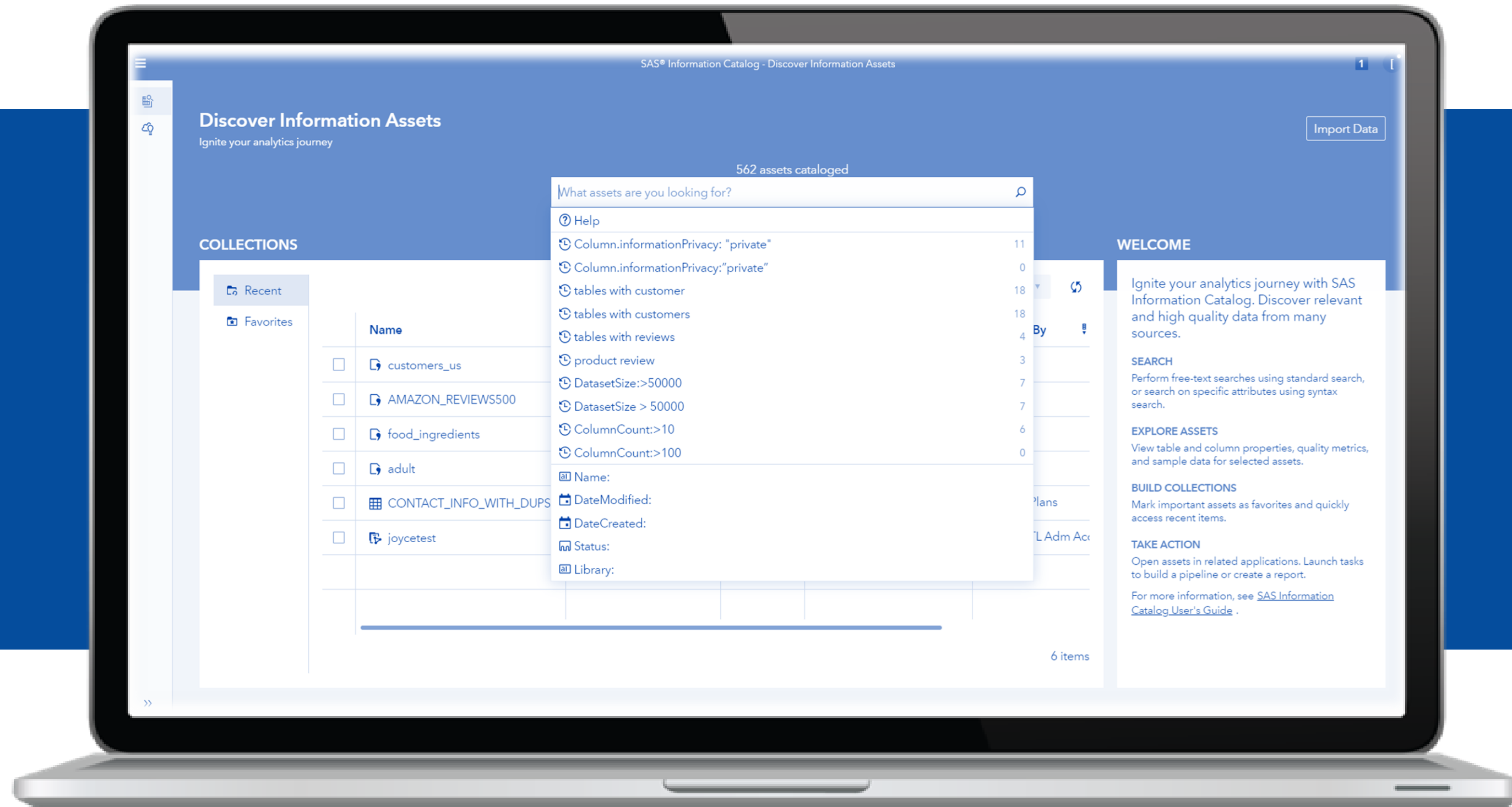
Area covered

Languages detected

Business Description

Additional information

# Faceted Search



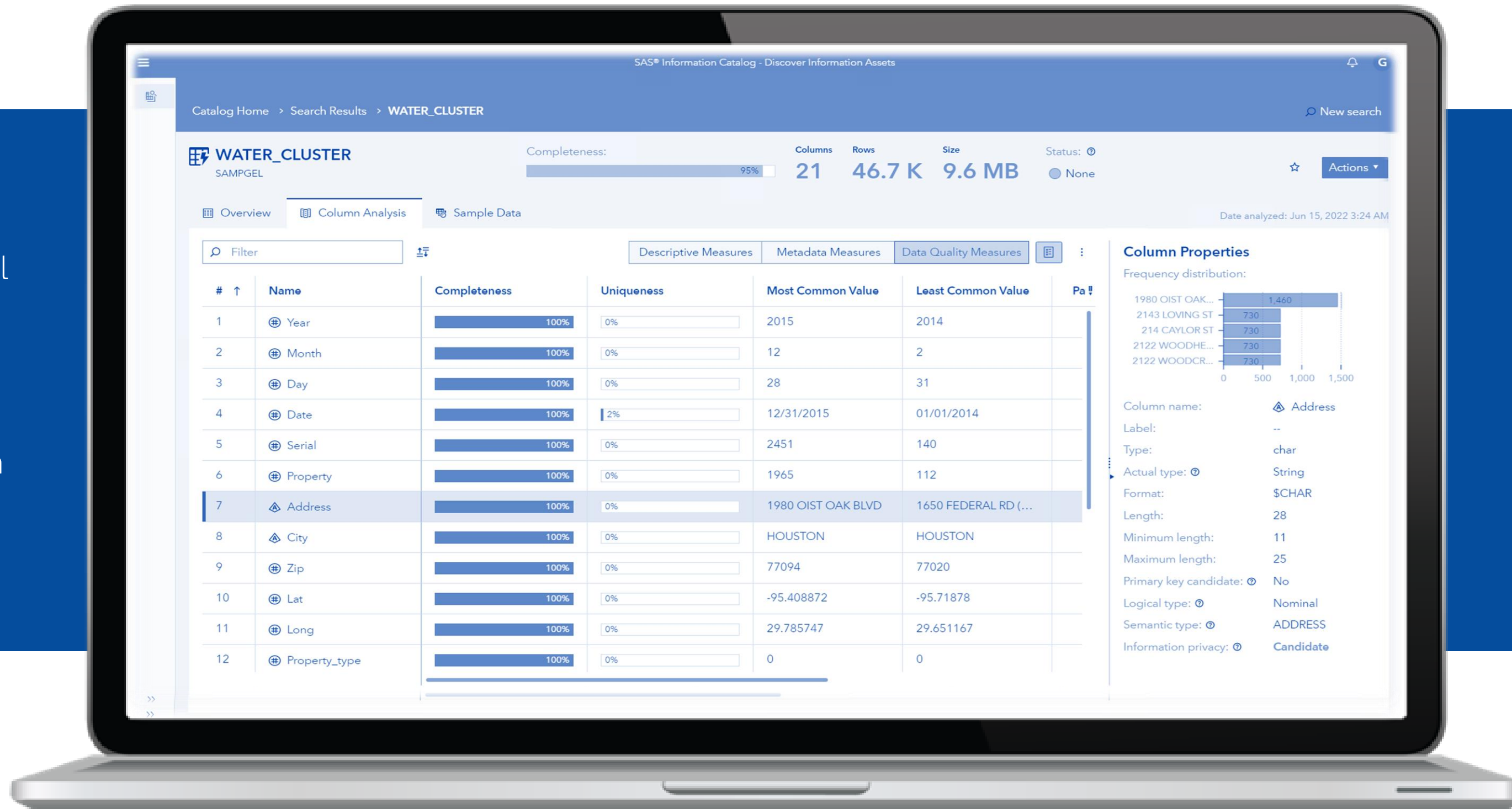
# Define, discover, and profile data

Automated & prebuilt Data Quality Statistics:

Missing, blank & null values etc.

Outliers

Pattern Recognition



Frequency Distribution

Referential Integrity

Redundant Analysis

# SAS Lineage

Provides complex business knowledge

Enables traceability of information

Visualize relations between business terms, data elements, data processes, controls and roles

Knowledge repository integrated in Information Catalog

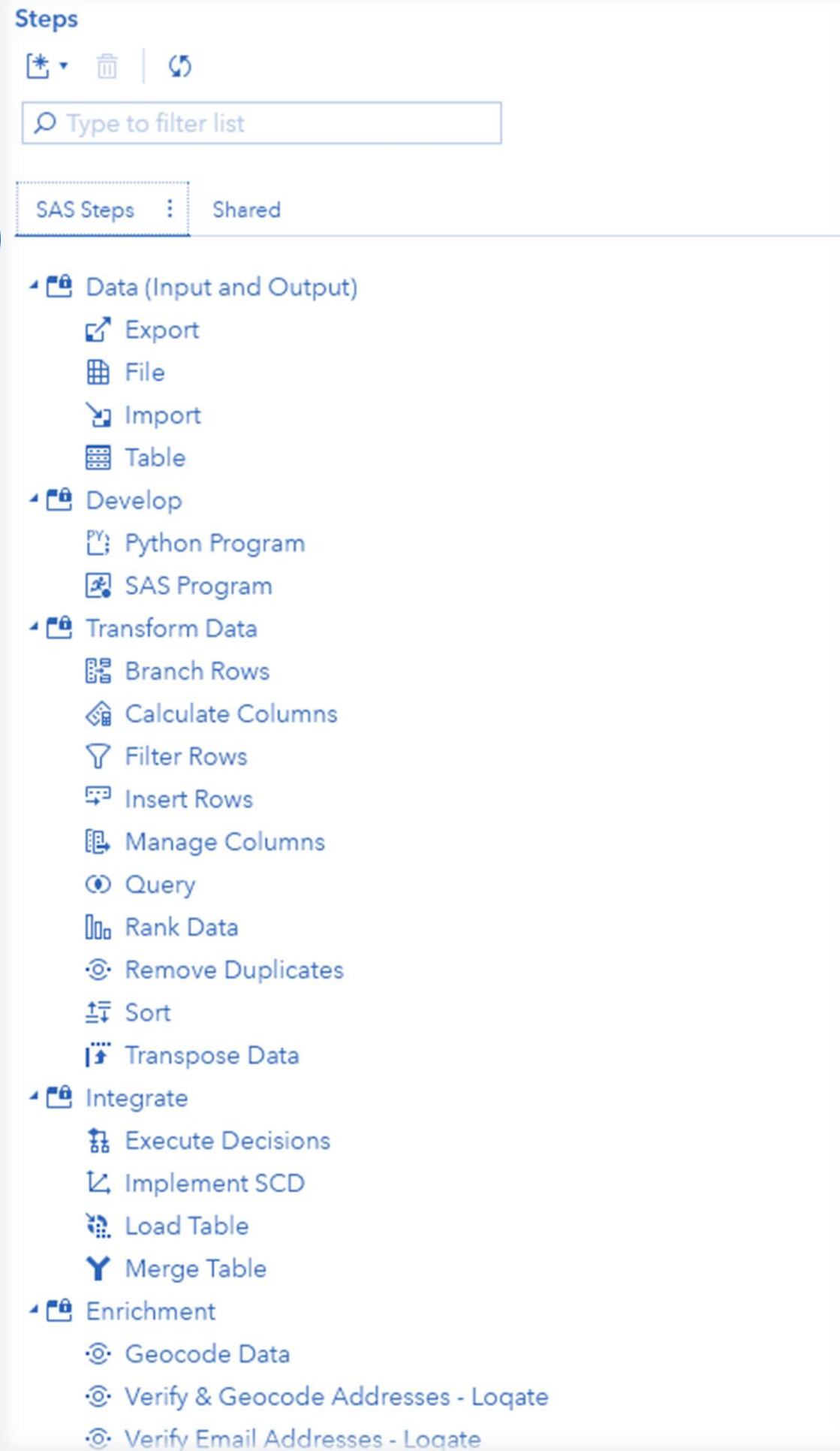


# SAS Studio Overview

- Access virtually all data sources
- Extract, cleanse, transform, conform, aggregate, load and manage data.
- Support data warehousing, migration, synchronization, federation and provisioning initiatives.
- Support both batch-oriented and real-time master data management solutions.
- Create reusable data integration services in support of service-oriented architectures and data governance.



# SAS Studio Overview

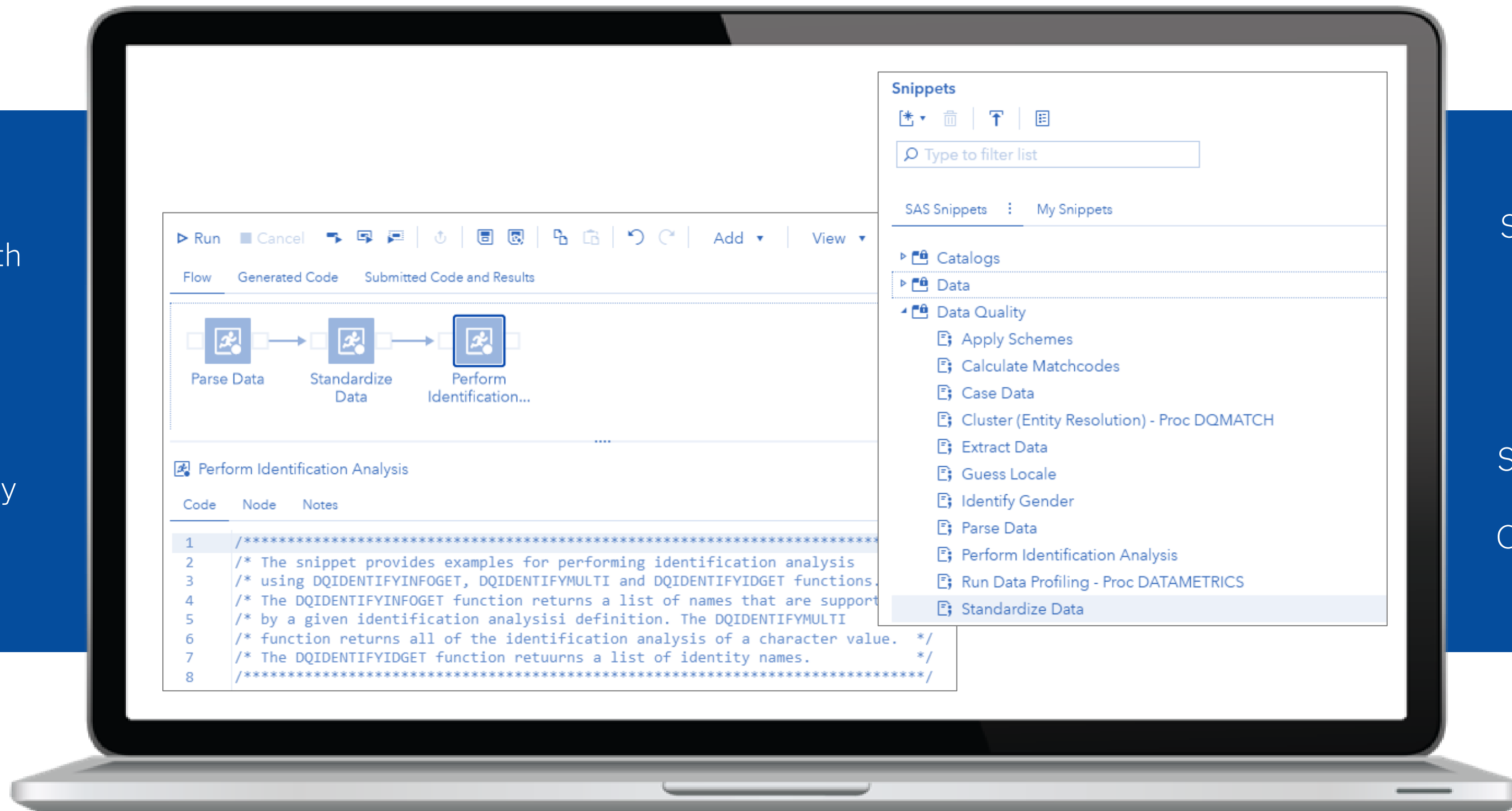


- Interactive data integration development environment
- Enterprise connectivity
- Data profiling
- Extraction, transformation and load (ETL) and (ELT) ready
- Metadata management
- Migration and synchronization
- Works with both CAS and SPRE service context
- Leverage/reuse work by others (GitHub repository)
- Work collaboratively in groups or teams

# Enable Data Quality with SAS Studio

The Data Management solution comes with world-class data cleansing and enrichment capabilities

Improve the quality of source data



Standardization

Matching

Clustering

Surviving record

Country specific