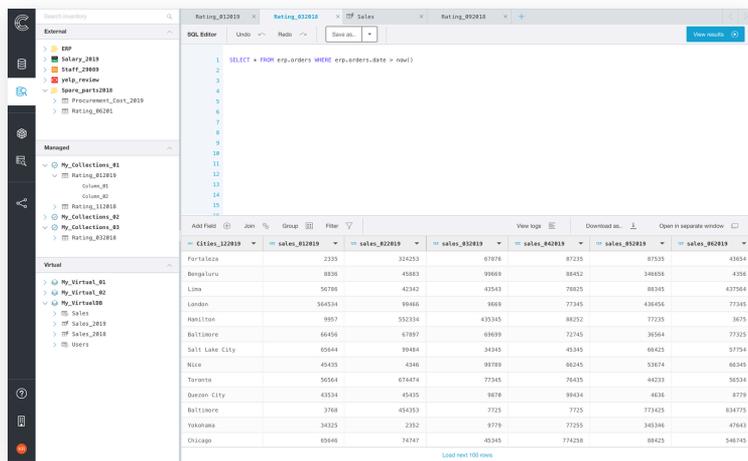




PRODUCT DATA SHEET

Contiamo is a data virtualization platform that radically accelerates data access to deliver agility and flexibility to data projects. Through its SQL-only setup, it is the fastest way for data engineers to deliver data. It also offers comprehensive data catalog functionality, resulting in the simplest way for data consumers to search and access data.

THE FASTEST WAY TO DELIVER DATA



SQL-only setup

Get started in minutes, utilizing all features without learning another proprietary language.



Intelligent materializations

Intelligent materializations ensure maximum query speeds with the minimum load on source systems.



Native connectors

One-click integration for the most popular traditional and modern enterprise data sources

THE SIMPLEST WAY TO ACCESS DATA



Search and discover data sets

Browse integrated data sources and virtual tables to find data relevant to your use case.



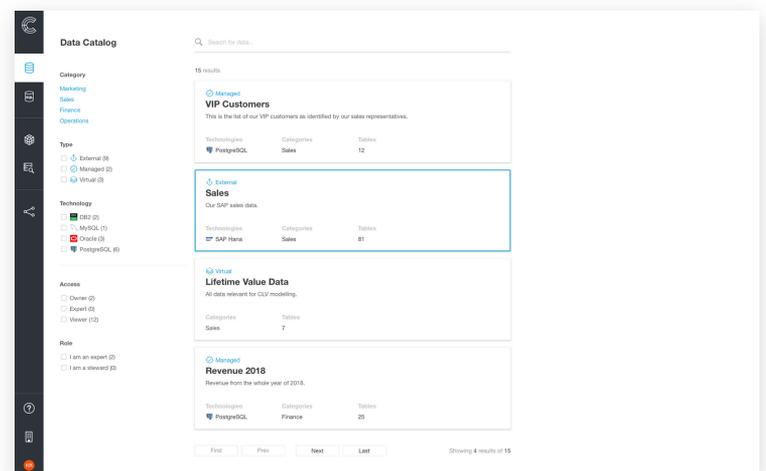
Explore data structures

Dig in to data structures to understand what the data set contains before requesting full access.

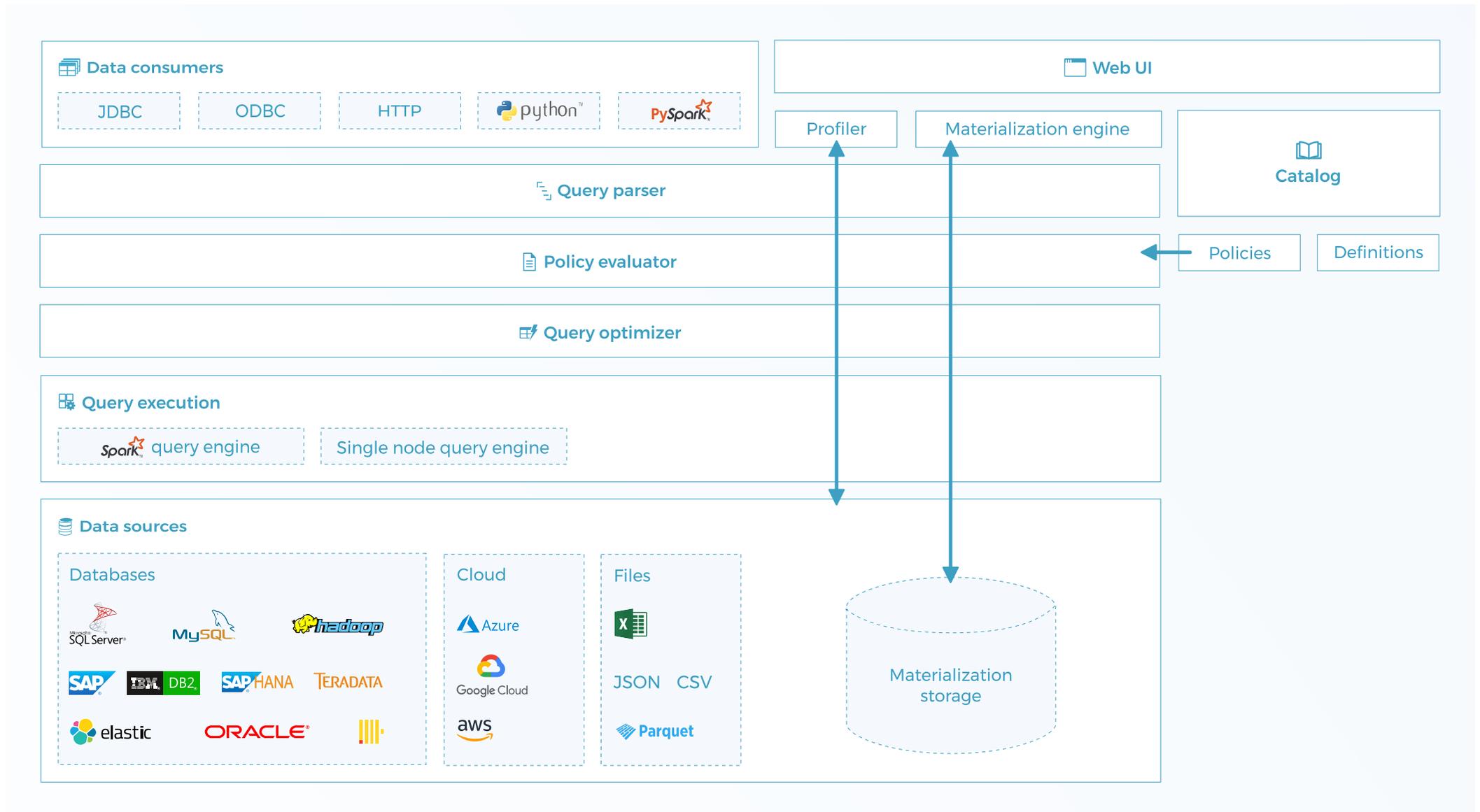


Easily request data access

At the click of a button, notify data stewards of your interest in gaining access to a data set.



THE CONTIAMO PLATFORM - ARCHITECTURE OVERVIEW



Data virtualization layer

Supported data sources

AWS

- S3
- Redshift
- (...)

Azure

- Blob storage
- Data Lake Gen 2
- (...)

Google Cloud

- Bigquery
- (...)

SAP HANA

Files

- CSV
- Excel
- Parquet
- JSON

Databases

- Oracle
- MySQL
- Microsoft SQL Server
- PostgreSQL
- MongoDB
- IBM DB2
- Redis
- Elasticsearch
- Microsoft Access
- SQLite
- Cassandra
- Splunk
- MariaDB
- Teradata
- Hive
- Solr
- HBase
- FileMaker
- SAP Adaptive Server
- Neo4j
- Couchbase
- Memcached
- Redshift
- Clickhouse
- Hadoop
- (...)

Supported endpoints

Endpoints for easy integration into all existing BI tools, data analytics platforms and operational systems.

- JDBC
- ODBC
- Postgres
- REST
- Python
- R
- Spark

Performance optimizations

Persistence & caching of data for optimised query performance

- Postgres
- AWS Redshift
- Clickhouse
- (...)

Advanced query planner

Load distribution

Data catalog

Intelligent search & filtering

- Search across hierarchies
- Search across multiple data sets

Advanced data set suggestions

- Identification of similar & joinable datasets

Data governance

- Flexible 5 role-based access management system
- Vertical and horizontal data lineage

Security

- Integration into Active Director, LDAP, Kerberos
- Query logs

Deployment

Deployment options

- On-premise, cloud and hybrid
- Cloud native deployment (AWS, Azure, Google Cloud)
- Any kind of container-based deployment

Cluster size

- Minimum: single node (4 CPUs, 32GB RAM)
- Maximum: tested on 50 nodes