Lenses 2.2
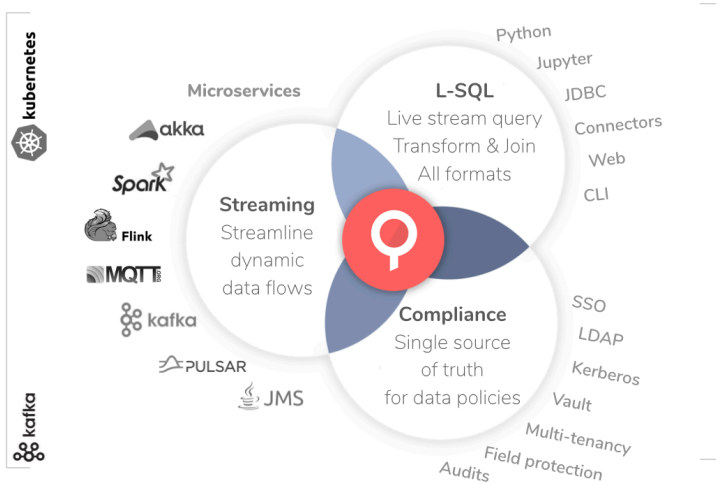# Enterprise Features

PRODUCT DATA SHEET

LENSES

Lenses ® is a DataOps platform that provides cutting edge visibility into data streams, integrations, processing, operations and enables ethical and secure data usage for the modern data-driven enterprise.

Accelerate time to value, open up data in motion to a wide audience, enable rapid construction and deployment of data pipelines at scale, for enterprises, with governance and security.

DataOps is currently transforming data management. Lenses is the answer to customised DIY solutions with in-house engineering teams spending most of their time building infrastructure and tooling.

Data should be in the hands of its users, as simple as their email, enabling innovation and minimising time to value. Building a data-driven culture allows all data personas to participate across the entire business. Lenses makes this possible.

OVERVIEW

DataOps is the art of progressing from data to value in seconds. For us, its all about making data operations as easy and fast as using the email. The main pillars are:

- ✓ **Data accessibility**
- ✓ **Build real time flows**
- ✓ **Ensure compliance and security**

....

Get more from you data with Lenses and DataOps, leverage your data and your skills to iterate quickly and move to production faster.

Focus on:
- ✓ Approaching data ethics
- ✓ Leveraging your whole organisation
- ✓ Data pipelines over infrastructure
- ✓ Data analytics over engineering
- ✓ Configuration over code
- ✓ Production Readiness

ACCESS DATA

Intelligent SQL provides universal data access and self-service data preparation via different clients & channels.

For All Users
- ✓ Data Consumers
- ✓ Data Engineers
- ✓ Data Officers

DATA FLOWS

Free up time from building infrastructure with self service SQL, Kubernetes and Prometheus integration for monitoring flows out of the box.

DATA GOVERNANCE & SECURITY

Operate a data compliant platform with fine grade access control, multi-tenancy features, strong authorisation and authentication integrations, data policies and field level protection.

APACHE KAFKA DISTRIBUTIONS

APACHE KAFKA	Apache Kafka 0.11+
CONFLUENT	CP 3.2.1+
CLOUDERA	Cloudera Manager 5.9.x+
HORTONWORKS	HortonWorks HDP 2.6.1+
AMAZON	MSK - Managed Streaming for Kafka
AZURE	HDI 4.0+

Supports all versions of Apache Kafka up to latest 2.1.0 Nov 2018 release

MANAGE, SECURE, AND GOVERN YOUR DATA

SECURITY	LDAP, Kerberos, TLS Certificates, Role based access
COMPLIANCE	Certify, Govern and Protect your Data assets
LINEAGE	Audits, Processor & Application Topology, Schemas
ETL READY	Kafka Connectors to simplify your streaming ETL
OPEN SOURCE	Prometheus, Grafana integration
CLOUD READY	Azure, Amazon, Google and others
MONITOR	150+ Key Indicator metrics of your infrastructure
ALERT	Setup and Receive alerts
CLIENTS	Python, JDBC, Go, ReduxJS, Java

Lenses ® Features & Description

Manage Data	
SQL Engine	Process your data to get timely insights with zero code - docs.lenses.io/lenses-sql
Supported Formats	Primitive Types, JSON, XML, CSV, Avro, Protobuf, Thrift, Custom Serializers
Complex Types	Complex type and CQRS support including: arrays, unions, structs, maps
Data Preparation	Filtering, transformations, calculations and aggregations
Quality Assurance	Insert data into topics with payload suggestions for better QA experience
Delete Data	Delete data from compacted topics
Replay offsets	Edit consumer commit offsets to replay messages

Access Data	
Web User Experience	Easily access and query data in motion • SQL editor with IntelliSense • Time travelling • Quick Filter • Navigate Partitions & Offsets • Access Metadata (ie. key, partitions, timestamp, offsets, headers) • Download query result data sets
APIs	Rest / WebSocket endpoints to integrate with data via Lenses. You can fire queries and subscribe to data for your custom data integrations with security and applied
CLI	**C**ommand **L**ine **I**nterface to automate running queries and build your CI/CD stream
JDBC Driver	JDBC driver plugin to access data in motion for your custom apps or BI tools
Native Clients	Python client for Jupyter notebooks. GO client and Redux middleware for building real-time dashboards for data in motion
Continuous Queries	Scalable continuous queries over Kubernetes with stream processing capabilities. See more on Lenses SQL Engine.
Query Management	View and Manage live queries from users via the admin CLI.
Access Metadata	Metastore engine: Lenses identifies the structure of your data for the key and the value and represents them in a schema, for any data format, which can also be updated. Key/Value types: Lenses autodetect the data format of the key and value data format. Users can also override types for each topic.
Avro Schemas	AVRO messages, accommodate for schema evolution, and enforce a data contract on your messages, being a fast and useful data serialization. Lenses integrates with widely used Schema Registries to manage schemas with RBAC.

Data Analysis & Preparation	
Know your data	Enjoy visibility for discovering, exploring and querying your Data
Access Data	Easy access to data irrespective of the underlying data format
Structure your data	Manipulate the layout of your data to be efficient and easy to work with.
Cleanse Data	Refine your data by leaving out corrupted records or enhance it in order to make it correct by the time the business logic is applied on it.

Transform Data	Apply the necessary transformations and aggregations and enrich data on the fly
Move Data	Move data between applications and source and target data systems
Extract Insights	Perform fast SQL operations for instant response to BI dashboards, reports, and machine learning applications. Lenses SQL integrates with native clients like JDBC driver, ReduxJS middleware for custom frontend apps, Python for integration with Jupyter or other notebooks.
Analytics	Continuous λ(sql), lambdas expressed in SQL as continuous queries, analysing, enriching and transforming data with a stream first approach that can even support live dashboards.
Distributed Joins	Express massively parallel distributed joins, without having to worry about partitioning schemantics
Self-service Infrastructure	Move your SQL logic into production in seconds, and enjoy monitoring alerting and control over your data pipelines
Anonymization	Anonymize and Obfuscate data while it transitions across your staging area

Data Transformation & Stream Processing	
SQL Processors	SQL processors are continuously unbounded SQL queries, specifically for streaming data. Each query is providing its internal topology which is visualised in Lenses UI as interactive nodes, where we have live input and output data.
SQL Processors Scalability	SQL processors provide 3 execution modes: • In process: which is good for development purposes • Kubernetes: where each runner is managed by a Kubernetes pod • Connect: where each runner is orchestrated by Kafka connect workers
Native scale via Connect Workers	Distribute your stream-processing jobs to Kafka-Connect clusters. Lenses SQL stream processors provide a native integration with Connect workers to scale up or down. The number of processor runners can simply be specified via Lenses UI.
Native scale via Kubernetes	Run and manage your data stream processing application within Kubernetes. Lenses SQL stream processors provide a native integration with Kubernetes to scale up or down. The number of processor runners can simply be specified via Lenses UI. Lenses manages the pod lifecycle and exposes the logs for debugging purposes.
SQL Processors Capabilities	• Self-service data preparation • Topic to Topic transformations • Filter and Enrich data • Join the stream (with repartition support) • Aggregations
Processing with Custom Apps	Lenses provides support for custom streaming applications. You can integrate the custom application nodes in the global topology view and leverage internal flow views and monitoring metrics for each app. Native support for Spark, Kafka Streams and Akka Streams
Monitoring	For each processing entity (SQL or Custom) Lenses collects performance metrics for which threshold alerts can be applied. For example data in/out, rates, data size.
Alerting	Threshold alerts can be specified for the input data for each SQL processor

Data Pipelines / Streaming ETL ready	
Global Topology	Lenses works out each step in your data processing pipelines (topics, connectors, SQL processors. micro-services) and renders them in a friendly and interactive graph.

Lenses Source Connectors	Bring data onto Kafka from multiple sources: Twitter, BlockChain, MQTT, JMS, RDBMS, FTP, File location, Yahoo, Bloomberg, Blockchain, Cassandra, CoaP, ReThink and more.
Lenses Sink Connectors	Popular distributed NoSQL and caching systems: Cassandra, DocumentDB, Elastic Search, SolR, HDFS, RDBMS, HazelCast, InfluxDB, MongoDB, HBase, Kudu, Redis, Pulsar, VoltDB, Rethink and others
SQL for Ingestion	Lenses connectors (see https://github.com/Landoop/stream-reactor) support SQL configuration to give extra capabilities and integrate with target systems
Custom Connectors	Lenses is compatible with all Kafka connectors as it implements the connect API. Lenses makes connectors available for spin up instances as well as available in the global topology.
Change Data Capture	Extract Change Data Capture (CDC) events from MainFrame or RDBMS via bin-logs.
Custom Stream Processing	Hook custom applications to your flows. Native support for Spark, Kafka Streams and Akka Streams and compatible with other types. (see Data Transformation)
Manage Lifecycle	Admin actions for Connectors & Processors, to edit, scale, start/stop.
Import / Exports	Lenses CLI supports import/export of different elements as configuration, to create repeatable flows and promote across environments.
Multi-tenancy	Lenses provides multitenancy support via blacklisting/whitelisting which can protect the relevant flows via topic-centric configuration. Restricts the access to involved connectors, processors and consumers.
Alerting	Apply data consumption threshold alerts for each entity involved: consumers, sink connectors, SQL processors

Security	
Role Based Access	Lenses supports role-based authorisation to protect your cluster and your data, especially in a multi-tenant environment. Also provides a no-data role for operation and support teams
Basic Authentication	Token-based authentication scenarios
LDAP	Integrate via memberOf or pluggable LDAP, or AD (Active Directory) and provide the people in your organization role-based access to Kafka streaming data
Kerberos	JaaS, Utilise Kerberos Authentication, Authorization, Single sign-on
TLS Certificates	Authenticate and Authorise Clients and Application via signed TLS certificates
ACLs	Manage Access Control Lists (Kafka ACLs) via Lenses interface / CLI
Quotas	Manage Kafka quotas for clients via Lenses Interface / CLI
Multi-tenancy	Blacklist and Whitelist capabilities for user groups. It applies topic-centric security which can restrict access to various topics and its involved pipelines

Governance & Lineage	
Audit Logs	Immutable audit logs to track all CRUD actions in a who did what and when. Lenses also audits actions happen outside Lenses, by periodically monitoring your infrastructure.
Audit Data access	Compliance is hard when it comes to data access so Lenses is logging all queries and data access for its users.

Data Schemas	Lenses keeps its own metastore for data types and schemas and also integrates with widely used Schema Registries to poll the schemas, including schema management with RBAC. Visualizes history of Avro schemas and ensures evolution. See Browse Data.
Topologies	Application / Topology level view of data-pipelines, to highlight relations of data with private or GDPR regulated data.
Query Topologies	Visualise the internal data processing flow structure for continuous queries and custom applications that processing streaming data.

Kafka Monitoring	
Lenses Services overview	Monitor the health status of the services involved in your infrastructure exposing basic metrics via JMX (Brokers, Schema Registry, Connects, Zookeeper)
Data Metrics	Topic Data metrics for infrastructure, to identify the most used topics, how topic traffic looks like and how data have been distributed across partitions. Ensure sync replicas.
Consumer Lag	Calculated and exposed consumer lag for each consumer group and its instances.
Metrics API	Lenses exposes calculated metrics via an endpoint for further integration
Prometheus Exporters	Templated agents for exporting critical key-metrics of the infrastructure, and reporting to Prometheus.
Grafana Dashboards	Collection of Grafana templates, curated over many years with suggestions and additions from our clients in production environments. Monitor Cluster overview, Kafka Metrics, Infrastructure, Consumer/Producer rates, Client Apps.
Monitor Client Apps	Our metrics exporter works with your custom JVM applications and we give visual metrics for each application
Alert Manager	Integrates with Alert Manager for alert routing management
Infrastructure Alerts	Built-in and configurable infrastructure and domain alerts
Threshold Alerts	Create custom alerts for the different consumer groups.

Find out more

➡ The Lenses.io Documentation page
➡ Interesting Resources

About Us

A team of passionate distributed system specialists based in the City of London with high-frequency trading, investment banking, betting, media and retail experience are the creator of **Lenses** ®

We aim to define Data Ops and assist companies open up their business data to all relevant users seamlessly by giving them the power of data operations. DataOps is currently transforming data management. Building a data-driven culture mandates that all data personas work with data, enabling participation from the entire business.

Modern organisations are trying hard to expose their data via platform teams. As a result, many end up with customised DIY solutions with in-house engineering teams spending most of their time building infrastructure and tooling.

We believe that data should be in the hands of its users, as simple as their email, to enable innovation and minimise time to value and used ethically.

Wanna give it a try ?

Lenses.io