

Nexla Platform Overview

INTEGRATE. PREPARE. DELIVER. MONITOR.

Build & run high performance, automated data workflows for your business

Organizations understand the important role data plays as a strategic asset for various business goals, such as improving customer experiences, running efficient business operations, or improving products and services.

However, accessing, managing, and getting data that is ready-to-use for these initiatives has become very complex. In order to make more data usable, data engineering teams are generally made responsible for building data pipelines to reliably and efficiently deliver data. But the process of getting and delivering data comes with numerous difficulties for data engineering teams:

1. Data engineers are spending a lot of time hand-coding repetitive data ingestion tasks.
2. Since data workflows and platforms continuously change, data engineers build and maintain, and then rebuild again complex infrastructure using old time-consuming ways.

3. Building real-time data pipelines for low latency systems becomes even harder to build and maintain.
4. Huge efforts are made by engineering teams to manually tune pipelines, optimize performance, and architecture to meet business demands.

As a result, engineering teams remain under immense pressure and fall behind on their projects. Moreover, they are unable to focus on strategic projects because they can't get out of the firefighting mode.

On the other hand, operations and analytics team become dissatisfied with data engineering teams as they struggle to get their hands on data that's crucial for the business operations and generating insights for the business.

How can Nexla help?

With the Nexla platform, data engineers as well as data scientists, analysts, and operations users have access to an end-to-end data engineering solution for ingesting, preparing, delivering, and monitoring data.

Universal Connectors, **Continuous Metadata Intelligence**, and **Nexsets** offer a unique approach that's innovative and provides a scalable solution for modern challenges in the organizations.



Figure 1: Nexla fits in between data infrastructure and data consumption layer. Universal Connector Architecture means that you will never be limited by connectors for your existing and future data engineering projects. Continuous Metadata Intelligence feeds the insights coming from observed and inferred metadata to create logical data products called Nexsets. Nexsets are automatically created abstractions that make it simple for users of any skill level to work with data. The Nexsets are used to transform, enrich, validate, and monitor data coming from and going to anywhere. Nexsets materialize the data to wherever it's needed whether its an app, data warehouse, or a stream. The result is a collaborative environment that empowers your data engineering, analytics, and operations teams to solve the toughest data challenges.

Universal Connector Design

You don't hand code connectors with Nexla. You just configure them. Even though we provide hundreds of connectors out of the box, we know that connector needs grow exponentially in a data-driven organization. That is the reason we invented a new connector approach. You can now configure a brand new, high-performance, bidirectional connector for any system in less than 60 minutes.

Continuous Metadata Intelligence

Nexla's continuous metadata intelligence observes data at a record level, infers metadata, and combines that with system metadata to generate a deep understanding. It automatically creates logical data products, Nexsets, that organizes schema, samples, descriptions, annotations, ratings, and other metadata.

Automated Data Products: Nexsets

Nexsets are logical data units that abstract the underlying heterogeneity of data and provides a consistent interface to data regardless of source, format, and velocity. You can transform and send Nexsets to data users or your teammates across organization or to a consuming application. Just like collaborating in a Google Document, sharing and collaborating with data is very simple in Nexla without compromising data governance or quality.

No/Low-Code Data Engineering & Operations

Nexla provides a collaborative, no/low-code approach to designing data flows that empowers users of any skill-level to work with data. They can integrate, transform and design data flows to get ready-to-use data from anywhere or share it with internal or external collaborators.

Adaptive Data Provisioning

Nexla's adaptive data provisioning make it simple to create ELT, ETL, API integration, streaming integration, API proxy, and real-time data as a service integration patterns. Nexla's universal connector design and Nexsets as an abstraction layer make this multi-modal integration possible.

Developer Friendly

In Nexla, you can declaratively create data flows and programmatically control Nexla with a complete set of APIs. Developers can customize Nexla or build their own application on the platform. Our team is always here to give you complete support for your projects.

Modern Data Flow and Data Processing

Nexla runs Kafka and Spark under the hood and it also allows bringing in of your own processing engine. The system can automatically select the right processing engine depending on the use case. Modern containerized micro-services based architecture means the system can auto-provision and auto-scale in response to changing processing requirements.

Run Anywhere

Nexla is available as a SaaS offering but can also be installed as private or hybrid installation on any infrastructure – AWS, Azure, or GCP.

Monitoring, Alerting, and Intelligent Error Handling

Nexla has extensive mechanisms that catch any instance where pipelines fail. For example, a required field might be missing in data, RPC timeout from a service might occur, data might be incorrectly formatted. Intelligent error handling addresses most of these errors automatically with easy to setup alerting capabilities.

LEADERS TRUST NEXLA FOR MISSION-CRITICAL DATA



CAPABILITIES AT A GLANCE

Data Integration

- Files: FTP, SFTP, S3, GCS, Dropbox, Box
- Data Formats: JSON, JSONL, XML, CSV, TSV, PSV, XLS, Log files (any Grok pattern), Image recognition (OCR), Avro, Parquet, ORC, EDI, fixed width, custom separator text, protobuf, gzip, zip
- Automated Database Connectors: Oracle, MySQL, Postgres, SQL Server
- Optimized Warehouse Connectors: Firebolt, Redshift
- Data Science: Pandas
- Streaming: Kafka, Webhook, Email
- 100+ Automated API connectors
- Advanced REST API authentication: OAuth, JWT, HMAC
- API chaining, API pagination
- SOAP API
- Custom API templates
- Advanced scheduling
- Streaming and real-time data processing
- Real-time API Proxy
- Data API (API access to data)
- Role-based data source authentication
- Custom sources and destinations

Data Catalog

- Automated dataset cataloging
- Automated data characterization
- Schema management
- Dataset annotation
- Column level lineage
- Integration with your in-house catalog tool

Data Preparation

- No-code excel-like functions:
- String manipulation, Math, Arrays, Date-time
 - Lookups, IP-Geo, Log Parsing
 - AI Functions: Sentiment analysis, one-hot encoding, entity analysis
 - Conditional logic, hashing, Security, and more
- Data filtering
- Validation rules
- Custom functions (Python, JS)
- Extensible function library
- Static and dynamic lookups
- Metadata transformations

Collaboration

- Create and manage teams
- Share, branch, and reuse workflows
- Share and reuse custom transforms
- Advanced permissions
- Schema templates
- Shared data prep functions

Data Engineering

- Create declarative pipelines
- Full API Access
- Nexla SDK
- Customize and plug in your own data connections, custom logic, and orchestration
- Full feature Command Line Interface for operations teams

Data Pipeline Management

- High speed pipeline: 2X, 4X, 8X
- Choice of SaaS hosted, Private VPC Deployment in AWS, GCP, Azure, On-premise deployment
- Integration with data governance & data catalog tools
- Multiple federated data backplanes

Data Operations & Governance

- Kafka-based stream data processing
- Custom real-time processing engine
- Spark-based batch data processing
- Error notifications in app and email
- Audit log
- Error quarantine
- Custom notifications
- Record-level lineage
- SLA guarantees
- Autoscaling containers, Dynamic orchestration

Data Security

- Credentials stored in vault
- AES 256 encryption for data in motion & at rest
- VPN Tunneling
- Google SSO, Active Directory, SAML, OIDC

Users and Support

- Unlimited Seats
- Team and organization management
- Web, Email, Slack Support
- Dedicated customer success
- Phone and video support
- Onsite training and onboarding