DataSync 360 – Product Requirements Document

Prepared by: Isha Jha

LinkedIn: https://www.linkedin.com/in/isha-jha-85398892/

Date: 2025

Table of Contents

Overview

DataSync 360 is an internal, web-based data observability and integration tool that centralizes monitoring, alerting, and performance analytics across ADF, Snowflake, APIs, and custom ETL scripts. It aligns technical and business KPIs to help teams detect, troubleshoot, and prevent data pipeline issues quickly.

Problem Statement

Existing monitoring tools provide only isolated system visibility and lack the business context needed for holistic understanding. Teams currently depend on scattered logs, emails, and spreadsheets, causing delays, duplication, and reduced confidence in reports.

Objectives & Success Metrics

Objectives:

• Centralize all pipeline visibility across tools.

• Reduce Mean Time To Detect (MTTD) and Mean Time To Resolve (MTTR).
• Display business KPIs—data freshness, SLA breaches—alongside system metrics.
• Enable collaboration through alerting and automated ticketing.

Success Metrics:
- 95% of active pipelines visible in dashboard
- 40% reduction in downtime
- 90% of stakeholders report higher data confidence
- 80% internal adoption in 3 months
- 100% of critical alerts dispatched within 2 minutes.

Target Users

• Primary: Data Engineers – Monitor and debug pipelines.
• Secondary: Data Analysts – Validate data freshness and consistency.
• Tertiary: Product Managers / BI Leads – Track data availability driving business KPIs.

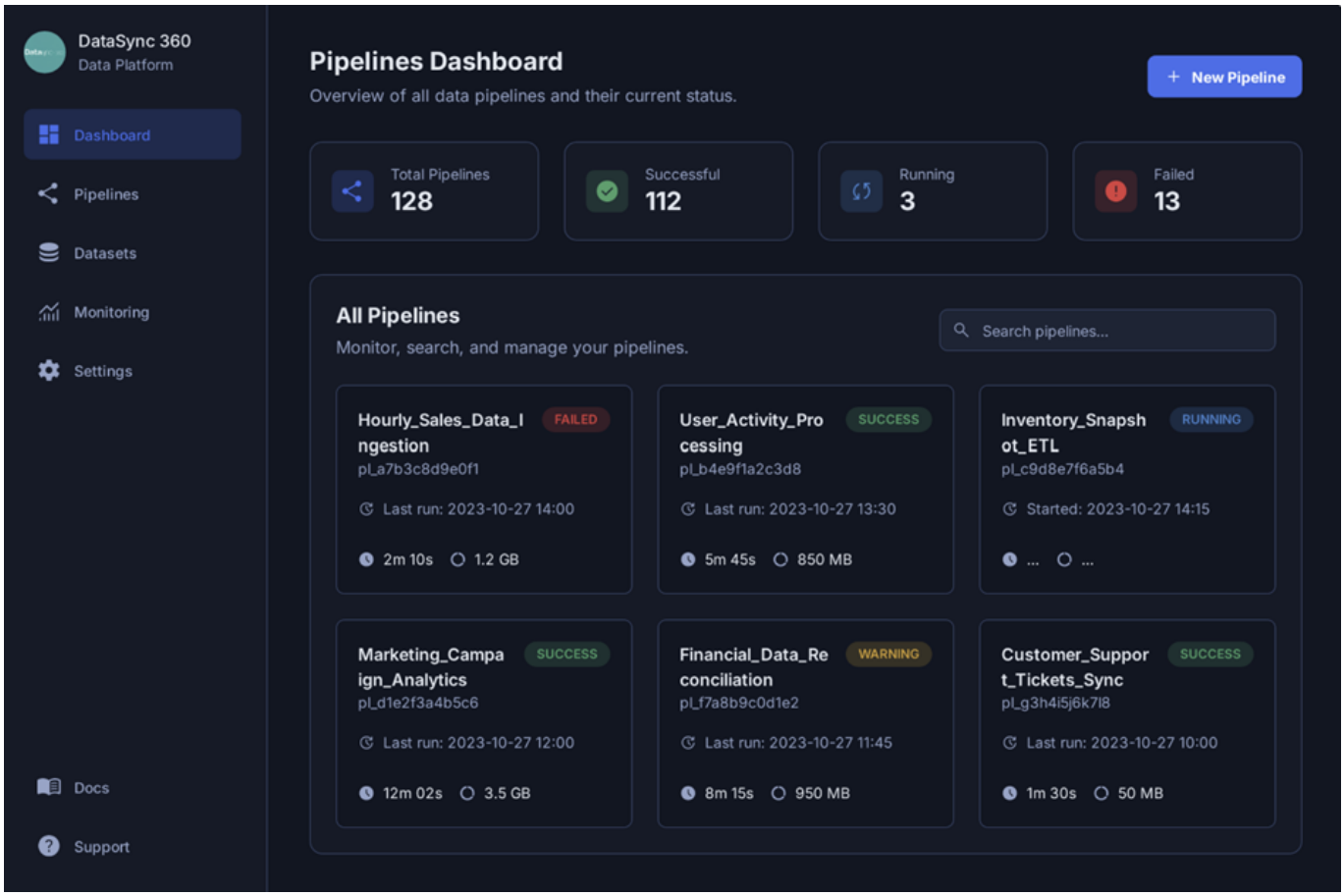
Key Features

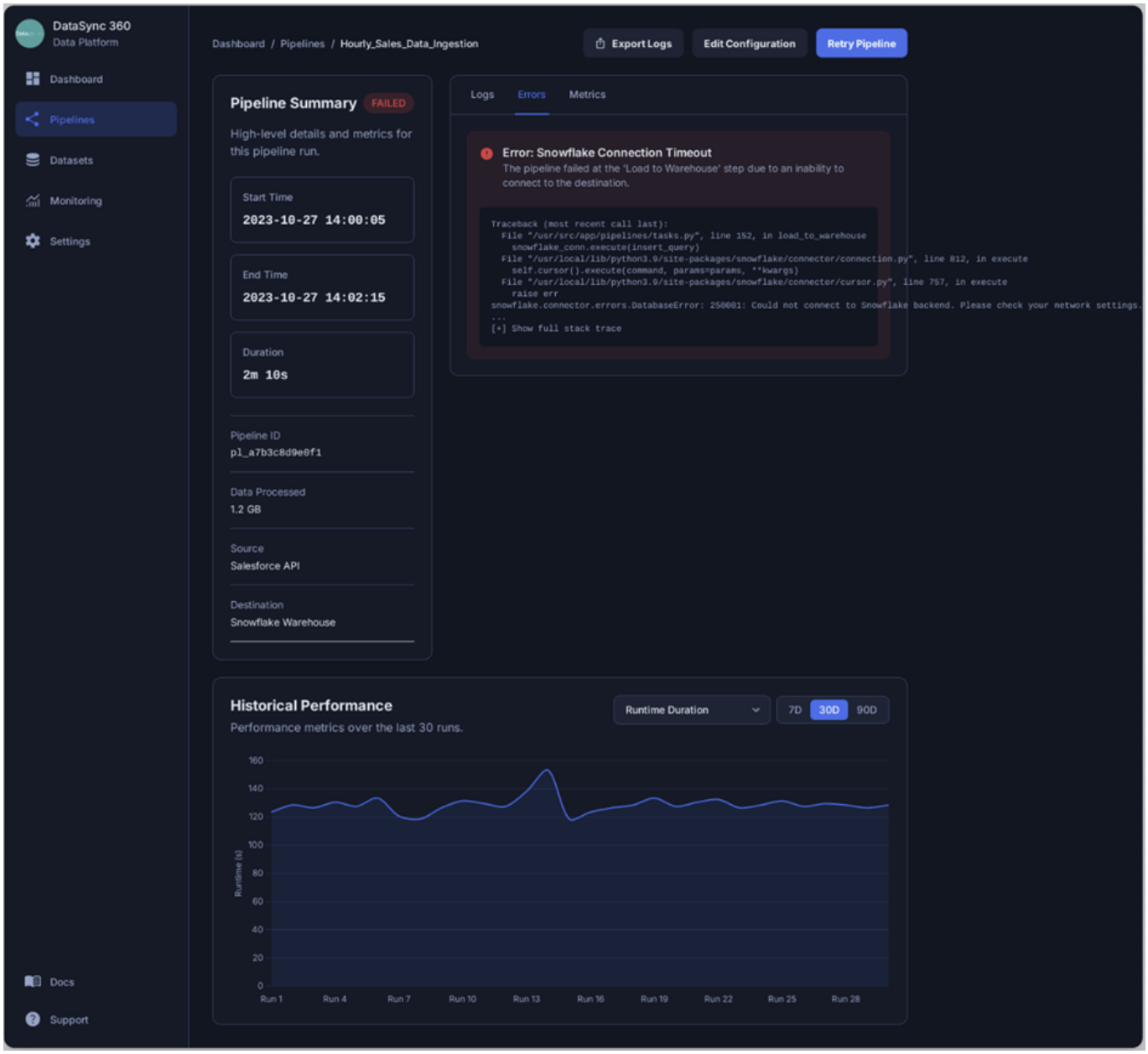
1. Pipeline Dashboard: Live status cards with filters and quick actions.
2. Pipeline Details: Logs, stack traces, and performance trends over time.
3. Monitoring & Alerts: Integrations with Slack, Teams, and Jira.
4. Historical Analytics: Runtime trends and data throughput.
5. Role-Based Access Control: Tailored dashboards by role.
6. Dark Mode: Intuitive UI optimized for rapid detection.

Design Overview

Wireframes (conceptual):

The following conceptual wireframes illustrate the core interfaces used in DataSync 360.





Technical Architecture & Dependencies

Core Integrations:

- Azure Data Factory APIs (metadata, run status)
- Snowflake and Redshift connectors
- Azure AD for SSO & RBAC
- Slack, Teams, Jira for collaboration

Data Storage & Processing:

- Azure SQL / Cosmos DB for metadata and metrics
- Logs stored in Azure Blob Storage
- Azure Functions for orchestration and aggregation

Risks & Mitigations

API rate limits – Medium – Apply throttling and batching.
Delayed alerts – High – Implement near-real-time polling.
Unauthorized access – High – Enforce RBAC, encryption, and auditing.
UI performance issues – Medium – Use pagination and lazy loading.
False alerts – Medium – Optimize alert thresholds.

Roadmap

Phase | Duration | Deliverables
MVP – 6 weeks – Core dashboard and ADF integration.
Monitoring – 4 weeks – Logs, retries, alerts.
Integrations – 5 weeks – Slack/Jira/Snowflake + RBAC.
Improvements – 3 weeks – Reports, reliability scores, admin tools.

Positioning & Differentiators

Positioning Statement:
DataSync 360 unifies technical and business KPIs across diverse data sources, enabling proactive reliability management.

Differentiators:
• Built for Azure-native infrastructure.
• Correlates system metrics with business outcomes.
• Delivers customization at a lower cost than enterprise tools.

User Stories & Acceptance Criteria

• As a Data Engineer, I can view all pipeline statuses in one dashboard for fast resolution.
Acceptance: Real-time, searchable dashboard view.
• As a Data Analyst, I receive alerts for pipeline failures to act promptly.
Acceptance: Notifications delivered to configured channels within 2 minutes.
• As a Product Manager, I track performance to forecast reliability.
Acceptance: Visual trend charts showing runtime and success rates.

Operational Plan & KPIs

Ongoing Tasks:

- Weekly reliability reviews.
- Monthly alert tuning.
- Bi-weekly pipeline audits.

KPIs:

- MTTD/MTTR trends
- False alert rate
- User adoption percentage

Appendix – Glossary & References

Glossary:

ADF – Azure Data Factory

ETL – Extract, Transform, Load

SLA – Service Level Agreement

RBAC – Role-Based Access Control

References:

- Microsoft Azure Data Factory Docs
- IBM Data Observability Whitepaper
- Bigeye & Monte Carlo Product Pages