

## SAS2PY Platform Overview

**SAS2PY** is the industry's first fully on-premise, parser-driven **IBM DataStage** migration and transformation platform. Purpose-built for organizations modernizing legacy DataStage environments—including parallel jobs, sequences, shared containers, and ETL workflows—SAS2PY transforms these pipelines into modern, cloud-native formats optimized for **Snowflake, Databricks, BigQuery, Redshift, Azure Fabric, and DBT**.

With unmatched accuracy, security, and speed, SAS2PY helps organizations reduce risk, control cost, and accelerate their modernization timeline.

### What Makes SAS2PY Different

- **Parser-Driven DataStage Conversion**  
No AI hallucination—every DataStage component is parsed, transformed, and fully auditable.
- **Secure On-Premise Deployment**  
Run entirely within your infrastructure (on-prem or VPC). No code or data ever leaves your environment.
- **Full Metadata & Transformation Preservation**  
DataStage sequences, reusable jobs, parameters, and schema mappings are intelligently retained and mapped.
- **Built-In Execution Layer**  
Execute migrated workflows natively in Snowflake, Databricks, and other platforms using visual controls.
- **(Optional) Merlin AI Developer Assistant**  
Real-time AI guidance for debugging and optimizing DataStage-to-modern transformations.
- **Comprehensive Validation & Matching**  
Validate schema, data, and logic with row-level and partitioned comparisons across platforms.

## End-to-End DataStage Migration Process

### 1. Analyze

- Automatically inventories DataStage jobs, stages, containers, and dependencies
- Maps parameters, variable files, schemas, and data sources
- Prioritizes migration efforts by complexity and impact

### 2. Convert

- Converts DataStage logic into Python, SQL, and modern pipeline formats (e.g., DBT, PySpark)
- Retains business logic and mappings with minimal refactoring
- Produces clean, production-ready code for cloud platforms

### 3. Validate

- Confirms schema alignment and transformation accuracy
- Performs output comparisons between DataStage and target systems
- Supports full regression and functional testing

### 4. Execute

- Enables direct execution in Databricks, Snowflake, and other environments
- Includes job versioning, scheduling, and packaging features
- Supports DAG-based orchestration

### 5. Data Matching

- Validates row counts, aggregates, and calculated KPIs
- Performs partition-based checks by date, region, or customer segments
- Aligns source and target schemas automatically

### 6. Document

- Generates audit-ready migration reports
- Provides end-to-end data lineage, code annotations, and transformation traceability
- Ensures rollback visibility for compliance and governance



### Automated Schema Mapping

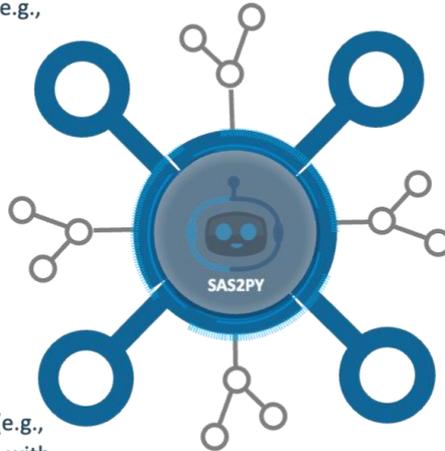
Automatically maps source schemas (e.g., SAS, Oracle, Teradata) to the target system for fast and efficient setup.

### Metadata Comparison

Compares metadata (e.g., table structures, indexes) between the source and target systems to ensure complete structural alignment.

### Data Type Validation

Ensures accurate translation of column types (e.g., numeric, string, date) into formats compatible with the target system, eliminating type mismatches.

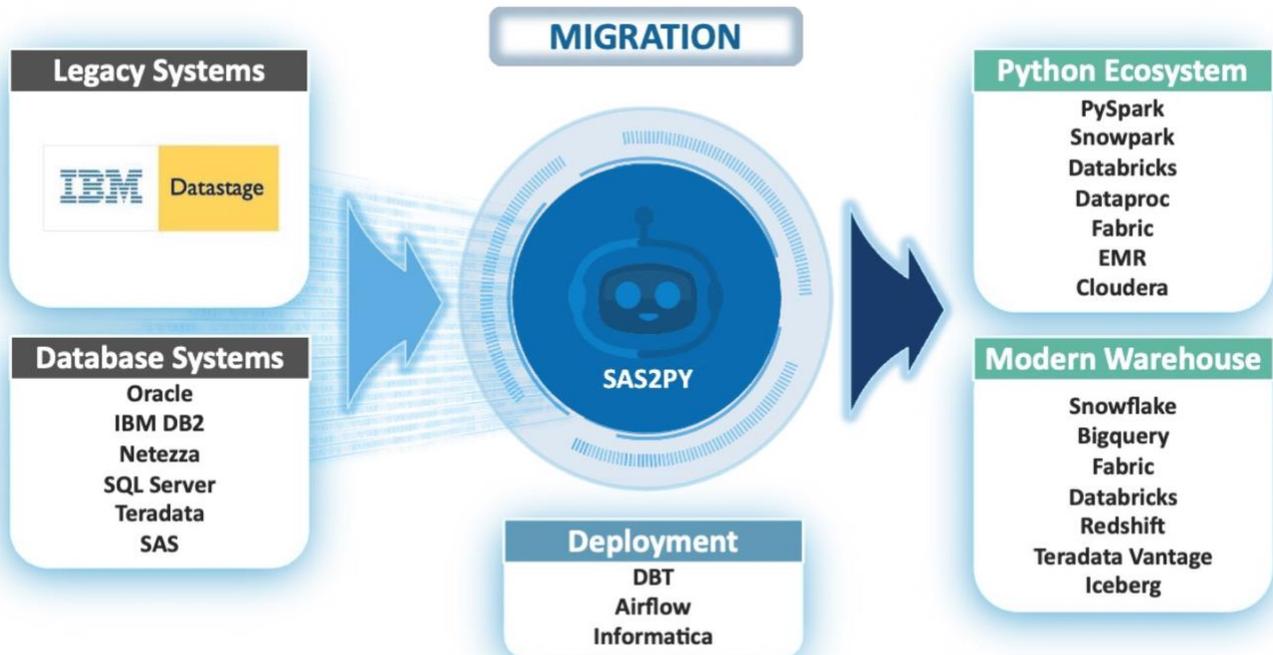


### Metrics Comparison

Validates key metrics (counts, sums, averages, etc.) between source and target systems to confirm data accuracy.

### Partitioned Validation

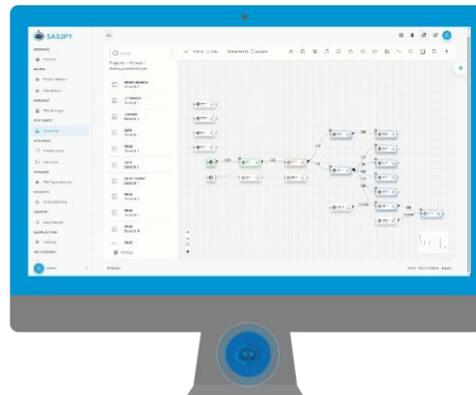
Performs aggregate checks by partitions (e.g., date, region) to ensure data consistency within subsets.



## Code/Data Lineage

**End-to-End Code Lineage:** SAS2PY automatically maps dependencies across all workflows, ensuring **100% traceability** before, during, and after migration.

**Clickable Lineage Exploration:** Users can **click on any dataset, transformation, or script** to immediately view the original code alongside its migrated equivalent.



**Legacy Code Preservation & Full Auditability:** Every **original script** remains **fully viewable, traceable, and auditable**, ensuring **long-term historical reference and compliance validation**.

## Merlin AI – The Developer Assistant (Optional)

**Merlin AI** is a GenAI developer assistant for developers and data engineers working with migrated code. Built directly into the SAS2PY platform, Merlin supports:

- Real-time debugging and transformation assistance
- Auto-rewrite of SQL and Python blocks for optimization
- Error tracebacks, logic flow explanations, and quick fixes
- GenAI-enabled context-aware chat for specific code inquiries
- On-premise optional for security-sensitive environments

## Enterprise-Grade Data Matching & Validation

SAS2PY includes a powerful data validation engine to ensure consistency and trust throughout your migration:

- **Automated Schema Mapping:** Aligns source and target columns and formats
- **Data Type Normalization:** Ensures compatibility across platforms
- **Metadata Comparison:** Verifies structures, indexes, and dependencies
- **Metric Validation:** Confirms row counts, aggregates, and KPI accuracy
- **Partition-Level Checks:** Detects mismatches by date, region, or business slice



## Deployment Options & Infrastructure

- Deploy within your infrastructure or in the cloud – you decide
- Docker-based delivery; compatible with AWS, Azure, GCP, on-prem VMs
- Standard Instances: AWS m5.2xlarge, Azure D8s v3, GCP n2-standard-8
- No data or logic ever leaves your environment

## Licensing & Engagement Options

- **Volume Pricing** Tiered by LoC (lines of code) and platforms migrated
- **Proof of Concept (PoC)** 4-week paid PoC
- **Partner-Friendly** We work with many SI's, technology partners and cloud migration partners

## Take the Next Step in Digital Transformation

**SAS2PY** is the world's most secure, flexible, and automated legacy modernization platform built to meet the needs of large enterprises across regulated, data-intensive industries. Whether modernizing IBM Datastage code, rebuilding ETL pipelines, or migrating entire platforms to the cloud, SAS2PY accelerates your path forward while reducing risk and cost.

**Schedule a Demo:** [www.sas2py.com](http://www.sas2py.com) | [sales@sas2py.com](mailto:sales@sas2py.com)



**Migrate Smarter. Analyze Faster. Transform with SAS2PY.**

[www.sas2py.com](http://www.sas2py.com)

[sales@sas2py.com](mailto:sales@sas2py.com)