MigryX × Snowflake

MIGRATE EVERYTHING TO THE AI DATA CLOUD

SAS · TALEND · ALTERYX · DATASTAGE · INFORMATICA · ODI · SSIS · TERADATA → SNOWFLAKE

Land Every Legacy ETL Workload on the AI Data Cloud.

MigryX converts SAS, Talend, Alteryx, IBM DataStage, Informatica, Oracle ODI, SSIS, Teradata, and SQL dialects directly to **Snowpark, Dynamic Tables, Streams & Tasks, Snowpipe, and Cortex AI** — with 95%+ parser accuracy and 99% with AI augmentation.

99%	60%+	10×	8+
AI ACCURACY	COST SAVINGS	FASTER	SOURCE ENGINES

```
silver_claims.py · Snowpark

# Source: SAS DI Studio
# Target: Snowpark + Dynamic Table
from snowflake.snowpark
  import Session, functions as F

def silver_claims(session):
  return (session.table(
    "BRONZE.CLAIMS")
    .filter(F.col("id")
      .is_not_null())
    .dropDuplicates(["id"]))
```

Conversion Accuracy 99.2%

Schedule a Snowflake Demo →

Email us →

THE REALITY CHECK

Manual Snowflake migrations stall. Generic transpilers approximate.

Every legacy ETL platform — SAS, Talend, Alteryx, DataStage, Informatica, ODI, SSIS, Teradata — speaks a different dialect. Manual rewrites take years. Generic scanners miss macros, implicit joins, and column-level lineage. Both leave production at risk.

The True Cost of a Stalled Migration

Legacy ETL licenses, rewrite consultants, and tribal-knowledge risk compound monthly. Most "lift-and-shift" projects miss go-live dates by 12–18 months — and arrive on Snowflake with broken lineage, no STTM, and reverse-engineered logic nobody trusts.

$2M+
ANNUAL LEGACY SPEND

3–5yr
MANUAL REWRITE TIMELINE

⚠ Manual Migration to Snowflake

- ✗ 3–5 years of consultant-led rewrites burning budget
- ✗ SAS macros, DataStage XML, and SSIS .dtsx logic lost in translation
- ✗ No column-level lineage or STTM — governance launched empty
- ✗ Pipelines land as raw SQL — no Snowpark, no Dynamic Tables
- ✗ No Streams, Tasks, or Snowpipe — just code dumps
- ✗ Validation is manual — production incidents discovered post go-live

❄ MigryX → Snowflake

- ✓ Custom AST parsers per source — 95%+ deterministic, 99% with AI
- ✓ Output as Snowpark Python DataFrames with pushdown execution
- ✓ Dynamic Tables with declarative refresh and lag targets
- ✓ Streams + Tasks DAGs replace legacy job schedulers
- ✓ STTM and lineage published to Snowflake object catalog with TAGs
- ✓ Native data matching using Snowflake comparison queries

5-STAGE METHODOLOGY

From legacy artifact to production Snowflake — in five proven steps.

1 — Ingest
SAS scripts, Talend exports, DataStage XML, .dtsx packages, ODI repositories.

2 — Parse + Lineage
Custom parsers build ASTs, expand macros, produce column-level lineage.

3 — Convert
Parser-driven Snowpark, Dynamic Tables, Snowflake SQL, or Tasks DAGs.

4 — Validate
Row-level + aggregate matching using Snowflake-native queries — audit-ready.

5 — Govern
STTM published to Snowflake object catalog with TAGs and Lineage API.

ENGINEERED FOR THE FULL SNOWFLAKE PLATFORM

Not just code conversion.
Production-ready Snowflake architecture.

Every MigryX migration leverages the full Snowflake platform — Snowpark compute, Dynamic Tables, Streams & Tasks, Zero-Copy Cloning, Time Travel, and Cortex AI. Production-ready from day one.

SNOWFLAKE-NATIVE OUTPUT · AUTO-GENERATED

Compute & SQL	Snowpark Python · Snowpark DataFrames · Virtual Warehouses · Snowflake SQL · UDFs · UDTFs · Stored Procedures · 500+ Function Mappings
Pipelines & Orchestration	Dynamic Tables · Streams (CDC) · Tasks (DAGs) · Snowpipe · Openflow · Serverless Tasks
Storage & Tables	Native Tables · Iceberg Tables · Zero-Copy Cloning · Time Travel · MERGE INTO · Search Optimization
Governance	Horizon Catalog · Object Tags · Column-Level Lineage · Data Classification · Masking Policies · Row Access Policies · Lineage API
AI & ML	Cortex AI · Cortex LLM Functions · Snowflake ML · Model Registry · Feature Store · Notebooks · Streamlit in Snowflake
Apps & Sharing	Native Apps · Marketplace · Data Clean Rooms · Secure Data Sharing · Cross-Cloud Replication

Precision Parsing

Custom ASTs per source. SAS macros, DataStage XML, Talend .item, SSIS .dtsx — full fidelity.

Snowflake-Native

Snowpark, Dynamic Tables, Streams + Tasks, Cortex — not generic Python on a warehouse.

Horizon-Ready

STTM and lineage published into Horizon Catalog with TAGs and policy enforcement.

Air-Gapped Capable

Runs entirely in your network. Zero external calls. Deploy in <1 hour.

SOURCE COVERAGE

Every legacy ETL source.
One Snowflake AI Data Cloud target.

Purpose-built parsers for each source language. SAS macro expansion, DataStage XML, Talend .item files, SSIS .dtsx — full fidelity, deterministic output, no approximation. The same proven methodology applies to every source.

SOURCE ENGINES

- S — SAS Base / EGP / DI →
- T — Talend →
- A — Alteryx →
- D — IBM DataStage →
- I — Informatica →
- S — SSIS (.dtsx) →
- O — Oracle ODI →
- B — Teradata BTEQ →
- P — Oracle PL/SQL →
- F — SAS DataFlux →
- Q — 15+ SQL Dialects →
- M — Mainframe / COBOL →

MigryX Parser Engine

SNOWFLAKE-NATIVE OUTPUTS

- SP — Snowpark Python →
- SQL — Snowflake SQL →
- DT — Dynamic Tables →
- ST — Streams + Tasks →
- SP — Snowpipe →
- HZ — Horizon Catalog →
- IC — Iceberg Tables →
- CX — Cortex AI / LLM →
- NB — Snowflake Notebooks →
- ML — Snowflake ML →
- FS — Feature Store →
- SL — Streamlit in SF →

SOURCE → SNOWFLAKE	PARSER ACCURACY	+ AI AUGMENTED	LINEAGE DEPTH
SAS (Base / Macro / PROC SQL)	95%+	99%	Column-level
Talend (.item / Std Jobs / tMap)	95%+	99%	Column-level
IBM DataStage (DSX / ISX)	95%+	99%	Column-level
Informatica PowerCenter / IICS	95%+	99%	Column-level
SSIS (.dtsx / .ispac)	95%+	99%	Column-level
Teradata BTEQ / FastLoad / TPump	95%+	99%	Column-level

MERLIN AI ENGINE

Custom parsers get you to 95%.
Merlin AI gets you to 99%.

Generic ETL scanners approximate lineage. MigryX parses it exactly — every macro, every column, every dialect — then lands it natively on Snowflake. Merlin AI resolves what deterministic parsers cannot.

Why Merlin lands cleaner code on Snowflake

Merlin understands legacy code semantics — not just syntax. It resolves SAS macro variables, DataStage runtime parameters, and Informatica session expressions that line-by-line transpilers silently mistranslate.

Macro & Parameter Resolution

Resolves SAS macros, DataStage job parameters, Informatica $$ variables, and ODI substitution APIs into deterministic Snowpark.

Snowflake-Aware Conversion

Recognizes streaming vs incremental vs batch logic and routes output to Snowpipe, Streams + Tasks, or Dynamic Tables automatically.

Auto-Generated Documentation

Creates STTM, data contracts, and inline lineage comments — published into Horizon Catalog with TAGs, zero manual effort.

Optimization Recommendations

Surfaces clustering keys, search optimization candidates, and warehouse-sizing hints from the parsed estate.

gold_claims_dt.sql

```
-- INPUT: SAS PROC SQL
-- PROC SQL;
--  CREATE TABLE claims_summary AS
--  SELECT region, SUM(amt) FROM
--  &env_schema..claims GROUP BY 1;

-- MERLIN OUTPUT: Dynamic Table
-- Resolved: &env_schema → 'PROD'
-- Routed: aggregation → Gold layer
CREATE OR REPLACE DYNAMIC TABLE
  GOLD.CLAIMS_SUMMARY
TARGET_LAG = '5 minutes'
WAREHOUSE = XS_WH
AS
SELECT region,
  SUM(amt) AS total_amt
FROM PROD.SILVER.CLAIMS
GROUP BY region;
```

Conversion Accuracy 99.2%

PLATFORM

Three products. One Snowflake migration.

Map your estate, convert your code, govern your outputs — every transformation, end-to-end, on your infrastructure. Compass extracts column-level lineage from your source and publishes it directly into the Snowflake object catalog before you migrate.

DISCOVERY

MigryX Compass

Map your estate before you migrate. Extracts column-level lineage, STTM, and dependency graphs — and publishes them directly into the Snowflake object catalog.

- Inventory + complexity scoring
- MIGRATE / ARCHIVE / DELETE classification
- Dependency graphs & risk scoring
- STTM published with Object TAGs
- Wave-by-wave migration roadmap

LINEAGE

MigryX Atlas

Universal cross-platform lineage spanning SAS, SQL, Python, ETL tools, and 30+ languages — feeding rich, parser-generated metadata into Horizon Catalog.

- Column-level traceability across stacks
- Cross-platform STTM & impact analysis
- GDPR / SOX / BCBS 239 audit reports
- Visual lineage explorer
- Snowflake Lineage API integration

AI ENGINE

Merlin AI

Domain-trained AI that understands legacy semantics, routes output to Snowpark, Dynamic Tables, or Streams + Tasks, and pushes accuracy to 99%.

- Macro & parameter resolution
- Snowflake-aware code routing
- Clustering & SO key hints
- Auto-documentation + data contracts
- Conversational Q&A on your estate

99%	10×	60%+	8+
AI-AUGMENTED ACCURACY	FASTER THAN MANUAL	COST SAVINGS	SOURCE ENGINES

On-Premise & Air-Gapped

Your source code never leaves your network. Fully isolated, zero external dependencies. Deploy in under an hour.

Docker & Kubernetes

Containerized on Docker, Kubernetes, or OpenShift. Scales with your workload — no infrastructure surprises.

Enterprise Security

SOC 2 certified. Role-based access control, complete audit trails, and compliance reporting built in.

TRUSTED BY SNOWFLAKE ADOPTERS ACROSS

Fortune 500 · Financial Services · Healthcare · Government · Insurance · Retail · Telecom · Pharmaceutical

WHY THE AI DATA CLOUD, WHY NOW

Your competitors are already on the AI Data Cloud. Are you still rewriting SAS by hand?

Every month on legacy ETL compounds the debt — in licensing, talent scarcity, compliance risk, and missed AI/ML opportunity. The Snowflake AI Data Cloud is the destination. MigryX is how you get there.

⚡ **The talent cliff is real:** The pool of engineers who know SAS, BTEQ, DataStage, and legacy ETL shrinks every year. The longer you wait, the more expensive your migration becomes — and the harder it is to find anyone who can do it. Modern data engineers want Snowpark and Cortex.

💰

Eliminate Legacy Licensing

SAS, Informatica, and DataStage licenses consume millions annually. Cutover to Snowflake delivers more capability for a fraction of the cost — typically 60%+ savings.

🤖

Unlock Cortex AI on Your Data

SAS analytical models land as Snowflake ML with Model Registry, Feature Store, and Cortex LLM functions ready to use — agentic AI from day one.

🚀

Streaming + Batch in One Place

Dynamic Tables unify streaming and batch with declarative refresh. Streams + Tasks handle CDC. Snowpipe ingests continuously — no Kafka, no Airflow.

🛡️

Governance Built-In

Horizon Catalog provides Object TAGs, masking policies, and lineage from the moment your migrated tables land. Audit-ready GDPR, SOX, BCBS 239 reporting.

📈

Preserve Decades of Logic

Decades of SAS macros and Informatica mappings encode institutional knowledge. MigryX preserves every transformation with column-level lineage — nothing guessed.

⚡

Zero-Copy Cloning & Time Travel

Spin up dev/test environments instantly with zero-copy cloning. Time Travel for free recovery. Capabilities legacy ETL stacks simply cannot match.

FOR DECISION MAKERS

See your legacy code land on Snowflake — live, in 30 minutes.

No slides. No generic demos. We migrate a sample of your actual SAS, Talend, DataStage, Informatica, ODI, or SSIS code into Snowpark and Dynamic Tables with full Horizon Catalog lineage so you can evaluate accuracy, structure, and governance yourself — zero commitment.

60%+	10×	99%	<1hr
COST SAVINGS	FASTER	AI ACCURACY	TO DEPLOY

Schedule Your Snowflake Demo →

Or send a sample — we'll convert to Snowpark/Dynamic Tables and return Horizon Catalog lineage free.

hello@migryx.com · (617) 512-9530

MigryX for Snowflake

Stop rewriting legacy.
Start running on Snowflake.

Join the enterprises that have already landed millions of lines of SAS, Talend, DataStage, Informatica, and SSIS code on the Snowflake AI Data Cloud — Snowpark-native, Horizon-governed, Cortex-ready — in weeks, not years.

> *"MigryX cut our SAS-to-Snowflake timeline from 3 years to 4 months. The Horizon Catalog lineage alone justified the entire investment."*

VP **VP of Data Engineering**
Fortune 500 Financial Services

> *"We submitted DataStage jobs Tuesday. Thursday we had Snowpark and Dynamic Tables running with full STTM. Nothing else comes close."*

CD **Chief Data Officer**
Global Healthcare Enterprise

❄ **Free Proof of Concept — On Snowflake**

Send a sample of SAS, Talend, Alteryx, DataStage, Informatica, ODI, SSIS, or BTEQ. We'll convert it, deploy to a Snowflake account, and return Horizon Catalog lineage — **free, no commitment.** **Start your POC →**

60%+	10×	99%
COST SAVINGS	FASTER	AI ACCURACY

Schedule Your Snowflake Demo →

snowflake.migryx.com | hello@migryx.com | (617) 512-9530