

Data Modernization: **The NSEIT Advantage**

BROCHURE



Data modernization: **the NSEIT advantage**

Responding to growing competition from digitally native fin-techs and ever-increasing demands from myriad regulators requires financial institutions to reimagine their businesses with technology at its core. Data-driven decision-making underpins this transformation. But extracting and tagging data residing across applications and databases deployed over decades is easier said than done. Financial institutions need an easy, accurate, and quick way to modernize their data.

NSEIT's Data Modernization solution helps our clients extract, tag, ratify, and cloud-enable data in minimal time and with high accuracy. We combine our two decades of experience in helping banks, stock exchanges, and other such financial players leverage technology atop a no-code platform to deliver a fast, seamless and reliable solution.

Data modernization- **the hurdles or just the hurdles**

According to **Aberdeen Strategy & Research**, 'there is a shortage of partners with expertise in data modernization—a highly complex process that is further complicated by legacy systems. This comes in the way of optimizing data infrastructure and leveraging cloud-native and hybrid technologies.'

In short, any data modernization effort must overcome several obstacles such as:

1

Data hosted across multiple, incompatible legacy databases and applications

2

Lack of adequate expertise in handling large, complex data sets

3

Lack of expertise in the deployment and management of automation tools

4

Privacy and security

NSEIT data modernization solution

Our solution combines our deep expertise in the financial sector with a no-code platform to get your data ready for new-age applications and the cloud. The solution has the following six components:



Discovery and assessment phase

- ▼ Assess data complexities of source systems (On Premise and Cloud) using no-code platform DataSwitch.
- ▼ Build code inventory with catalogue for near-to-actual estimates
- ▼ Recommend iterative approach with agile project plan and propose NSEIT's reusable frameworks to save on efforts significantly



Data platform modernization

- ▼ Automated schema redesign to convert structures underlying legacy data models into cloud-friendly modern data structures using no-code platform DataSwitch



Data platform modernization

- ▼ Automated tool to create a data pipeline and their executions using DataSwitch
- ▼ Extract, transform, transfer and load data to the identified cloud platform using SQL based serverless framework SQLake



Testing and reconciliation

- ▼ Reconciliation of the migrated data using automated tools to ensure data migration accuracy
- ▼ Row-by-row, column-by-column comparison, and validation of data by an automated tool



Data engineering pipeline modernization

- ▼ DataSwitch, a No-Code, web-based tool, converts legacy data scripts and ETL tools into modern ones. For e.g. Informatica Mappings gets converted into Python scripts with proper comments, sections and indentation



Technical documentation

- ▼ Documentation of converted code during conversion phase in CSV files
- ▼ Create and share metadata information, entity relationships, functionalities, data lineage along with graphical representation for easy understanding by Business Analysts and Development Team

NSEIT approach for data modernization

1. Check the weather:

- a. Generate code inventory with catalogue using discovery and assessment tool
- b. Identify & prioritize the end user reports after having discussion with stakeholders especially reporting users and visualization team

2. Ride the waves:

- a. Reverse engineer priority reports to identify the database objects i.e., tables/views it fetches data from. Sort the tables in phases if list is too large
- b. Convert the schema/DDIs of these schema/tables/views on priority and deploy them on cloud data warehouse without touching ETL code responsible to load them
- c. Carry out One Time Data Migration for prioritized tables so that they are in sync with On-Premises data warehouse
 - i. It can also be done in phases. For e.g., Executing data migration by dividing, it based on monthly/quarterly data OR using table partitions
- d. Point a copy of priority reports to cloud data warehouse so that it starts reaping the benefits of cloud scalability
- e. Carry out reconciliation of data migration
- f. Develop a data replication service from On-Premises data warehouse to cloud data warehouse so that data gets copied daily
- g. Carry out Steps a-to-f for all prioritized tables in phase wise manner
- h. This approach ensures all priority reports gets migrated on cloud data warehouse without going through ETL code migration to cloud. Reports start reaping benefits out of cloud data warehouse's scalability and flexibility
- i. If necessary, carry out catch up loads post one time data migration and before starting replication service to fill in the missing data to cloud data warehouse

3. Boil the ocean:

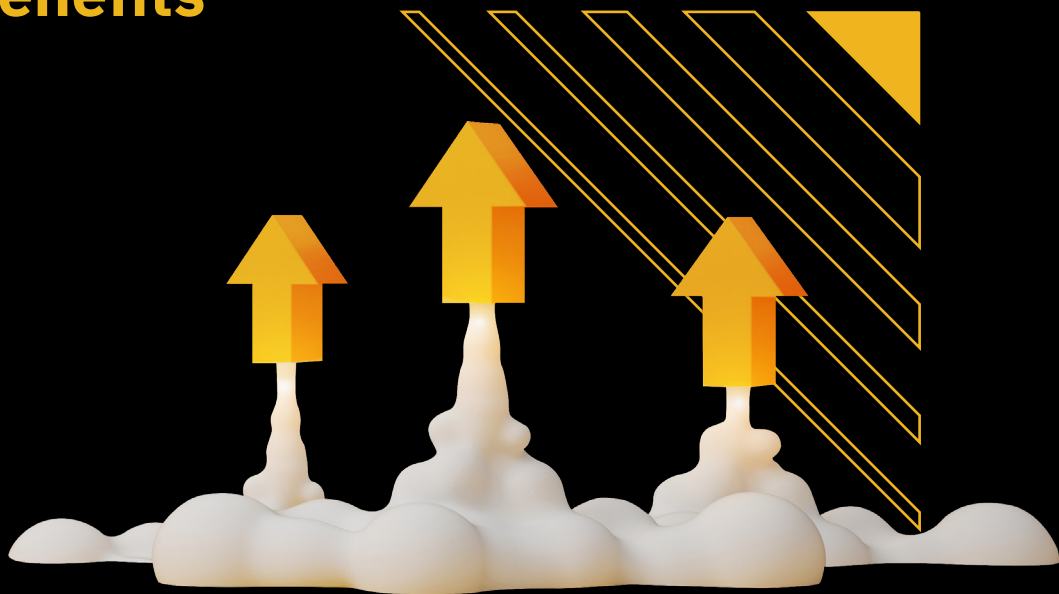
- a. Identify the ETL routines responsible for loading the priority tables in phase wise manner. Distribute them in phases, if necessary.
- b. For each phase, convert and test ETL routines into Data Engineering pipelines with below approach
- c. Orchestrate Data Engineering pipelines on public cloud and start executing them so that they start populating the data into respective target tables on cloud data warehouse
- d. However, before above step, do ensure to stop/decommission the replication service of respective tables so that it will not replicate On-Premises data onto cloud data warehouse anymore
- e. Gradually after converting all ETL routines into respective data engineering pipelines, entire replication service will stop
- f. This ensures that that all ETL code gets converted into modern tech stacks without having any impact on reporting/visualization there by improving ETL routine performance as well

g. This phase takes most of the time in project as converting and testing ETL code would consume bandwidth of most associates and hence kept as last phase of data modernization journey

Document metadata for all code that gets converted and prepare a report out of it to be presented to business users in a graphical manner to provide transparency in data engineering pipelines

Document metadata for all code that gets converted and prepare a report out of it to be presented to business users in a graphical manner to provide transparency in data engineering pipelines

The **benefits**



Our expertise in transforming the data operation of financial institutions results in:

- Discovery and Assessment Phase tools provide estimates that are close to actual thereby significantly mitigating the risk of project over run due to last minute surprises in scope change
- Data Platform Modernization (DDL conversion) tool save conversion efforts significantly thereby making the cloud-based data warehouse ready for data migration phase
- Data Platform Modernization Migration Frameworks provide ready-to-be-used code components with all best practices saving ~65% of efforts
- Testing and Reconciliation capability of the tool saves testing efforts significantly and ensures the quality of migrated data required to take sign off from all stakeholders
- Data Engineering Pipeline Modernization tool converts the legacy code into Python/Spark/SQL which can be consumed by cloud-based services readily.
- Technical Documentation is generally the neglected yet crucial aspect of modernization journey which is well taken care by tool with data lineage and graphical representations. This helps leadership to drive the data modernization programs to its highest maturity with self-service tools as future.

The **NSEIT** edge

A subsidiary of the National Stock Exchange [NSE], the largest in India, NSEIT is a finance vertical-focused IT solutions provider with a proven track record across digital transformation initiatives.

Over 200 experts in data modernization and data analytics

- ▶ During Assessment phase, **no-code automation tool** connected to all On-Premises DBs and generated a data catalogue from their data dictionary under **15 days**. This helped in planning whole project with stringent timelines
- ▶ Converted > 3000 table DDLs from **Greenplum into Snowflake using no-code automation tool** in just **7 days**
- ▶ Using **NSEIT's Data Migration Framework**, migrated **~0.6 Petabytes** of data from On-Premises Data Warehouse to public cloud in response to a 3X rise in trading volume—**within 24 days**—for the leading stock exchange

Converted & tested **~2000 Informatica Mappings** and **~400 Workflows** into **Snowflake SQLs** and **Apache Airflow DAGs** respectively using **no-code automation tool** under **5 months**

End-to-End data lineage created under **2 months** along with **customized B-tree based graphical representation** at table as well as column level along with **ETL Code and tagging it to PDM (Physical Data Model)** as well.

Quality **certifications**



About **NSEIT**

NSEIT Limited is a digital native technology company that engineers world-class solutions to help our global customers accelerate their digital transformation journeys. Our key service pillars are Application Modernization, Business Transformation, Data Analytics, Infrastructure &

Cloud Services, and Cybersecurity, through which we create intuitive digital experiences and tangible business impact. For over two decades, our innate drive for excellence has made us the partner of choice for global organizations. At NSEIT, we fuel digital progress.

For more information, visit us at nseit.com

Follow us at:



© NSEIT Limited. All rights reserved.

All trademarks, logos, and brand names are the property of their respective owners. All company, product, and service names used are for identification purposes only. Use of these names, trademarks and brands does not imply endorsement.