A stylized human figure, composed of glowing orange and yellow lines, stands on a circular platform that also glows with a bright orange light. The background is a blurred green and blue, suggesting a digital or data environment.

# **Data-Centric QA for Customer 360: Enabling Accurate, Actionable Insights**

A Whitepaper



## Retailers' data conundrum to creating candid Customer Experiences

Retailers are constantly improving customer experiences and service levels pinning their hope on Data. Being able to know a customer 360 degrees throughout their dynamic lifecycle is like holding a handbook of winning strategies. Build a profitable customer base, model customer treatments through sentiments, personalize campaigns and offers, market basket analysis, make customer engagement a configurable entity...all driven through an aggregate data center. The scheme is to join the discrete data dots of customer's interactions across retail channels and form a Facebook-like profile of customers to tackle quite a lot of customer intelligence case, predictive analytics, recommendations and so on.

While the benefits of a single customer data are golden, conceiving a big data warehouse model that delivers QUALITY and CONTEXT to DATA in ACCURATE FORMATS from the available disparate data sources and structures is no paper and pencil effort. ETL Tools and operations is a utopian process to extract data, cleanse, standardize and model the data; To mentor efficient ETL processes, Data-centric Testing for data quality, accuracy and completeness at every stage takes precedence to achieve a successful Customer 360 business case.

The end users of these reports view them as a representation of the end-to-end business journey but are unconversant with the data sources or the layers of rules that are applied during transformation. In this status quo, creating confidence on the data quality is tricky; the impacts are battered on the 'decision-making' dashboards.



“A recent Data Quality survey reports 92% of organizations suspecting data quality issues and on average suspect that 26% of their total data may be inaccurate”)

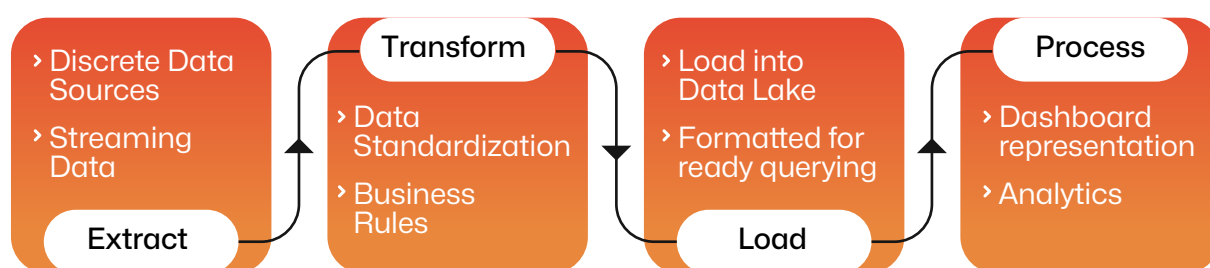
To build resilience to this ecosystem, taking a step back from the Reports/Visualization phase to trace the entire data revolt (data sourced, transformed and loaded: ETL Dev) will gain transparency to the problem. The number and variety of data sources involved, data specifications, staging, update frequencies and complexity of processing rules should go through a ‘Quality Drill’ to label ‘integrity’ to the BI output.

A formal Data-centric QA Strategy in a real time Big Data-ETL routine will mitigate the quality problems, boosting authenticity of the customer intelligence information for Retail business users.

## Meet the Testing challenges

The new normal of the heavy-duty (Big) Data Transformation in Retail

With a basic understanding of ETL process in a general setting (irrespective of domain), the risks of Data Quality at different stages better exposed.



Test strategy doesn't come easy in this set up. A few of the pressing concerns to craft a solid and sustainable test process could be as follows:



- › Transactional data of customers comes in large scales and forms in streams of micro seconds - insisting powerful databases and hence robust Test Data Management approach (standard approaches possibly will result in gaps unless customized)
- › Ambiguous definitions of Data Models
- › Access limitations to Testers on ETL Jobs, Business Workflows and Dashboards (keeping the end to end picture in dark)
- › Missing Test Scenarios due to complex data workflows which when identified at the reporting end might result in gaps unless costomized
- › Testing in less time with more coverage





# Testing Scope in a typical Retail ETL ecosystem

An illustration of the Retail ETL ecosystem and the judicious role of QA

## Source to Target Testing

In the process of Data transformation into a Customer 360-like system (an aggregate database), source to target compatibility tests play a significant part-- a direct consequence of data loss risks and therefore generating a misleading profile of customer

## Data Transformation Testing

Identify the data fields that are resistant to the positive transformation logic (driven by Business Intelligence needs)

## Data Quality Validations

Verify data profiles for completeness, accuracy and integrity. Validate

- Meta data rules
- Compliance
- Data responses to rejections or replacements

## Integration Testing

Validate the ETL design in upstream and downstream flows



### **Performance and scalability**

Ensure that the system is resilient to the data and concurrent operational loads. Data is incremental which drives the need to define performance expectations and impacts for streaming/batch loads

### **User Acceptance Testing**






To verify if the data model and workflow expectations are attended

### **BI Testing**

Verify data at field level through the sources and sequence of ETL stages on a querying interface.



### DATA SOURCES / DISCRETE DBs

-  Real time streaming data
-  Structured and Unstructured Data Files
-  Emails / Social Posts
-  Website / Web chats / Logs / Mobile App
-  Transaction Logs

### DATA STAGING

Data transformation for Customer 360



- Legacy Data (historic data for a desired period)
- Incremental Data (defined by desired frequency)

### SINGLE SOURCE OF DATA



Customer 360 / Aggregated view of a customer

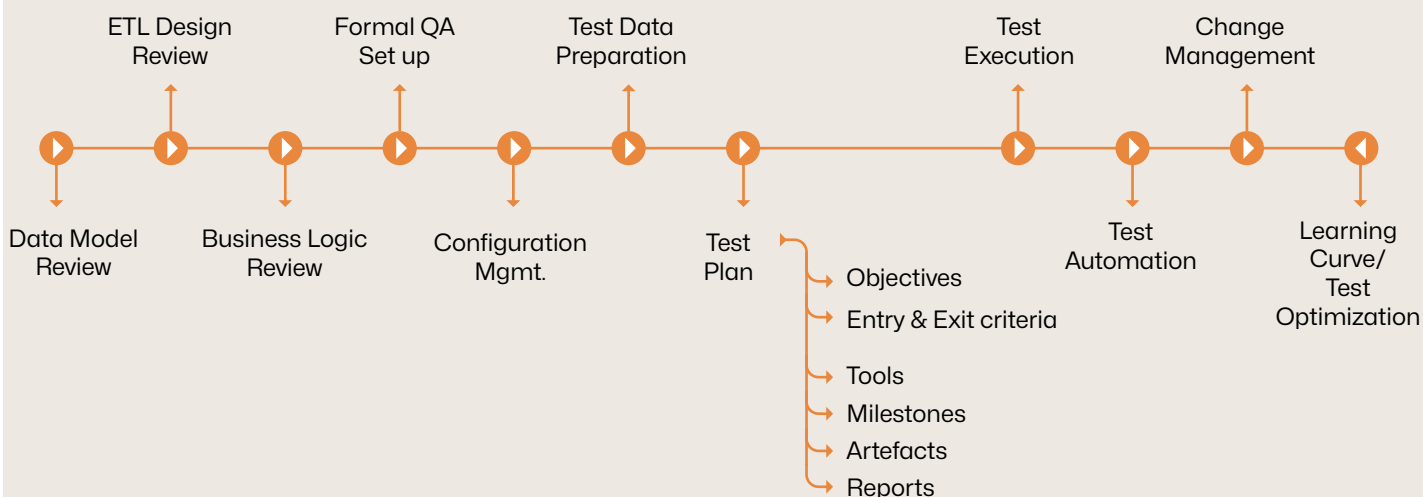


- Test Data Validations
- Integration Testing
- Workflow Tests

- Source to Target Testing
- Data Mapping Testing
- Data Completeness Testing
- Data Quality Validations
- Data Transformation Testing
- Metadata Testing
- Data Archival and Retrieval Testing

- Production Validations
- Customer Profile Validations
- Initial and Incremental ETL Testing
- Performance and Scalability Testing
- UAT
- Regression Testing
- BI/Dashboard Testing

### TEST PROCES





# ETL Testing Cube of Strategic Dimensions

## Knowledge of Domain

Understand business workflow for entry and exit areas, data transactional points to validate

## ETL Mapping

Maintaining ETL Mapping of Source to Target DB to reference and trace test case requirements with ease

## Data Validation Coverage

Map Test Scenarios to BI use cases and ensure 100% coverage (for example, the fields populated in the BI engine, frequency of this information loaded into the database etc.)

## Production Reconciliation

Custom test strategy due to compromises made during data updates

## Quality at Speed

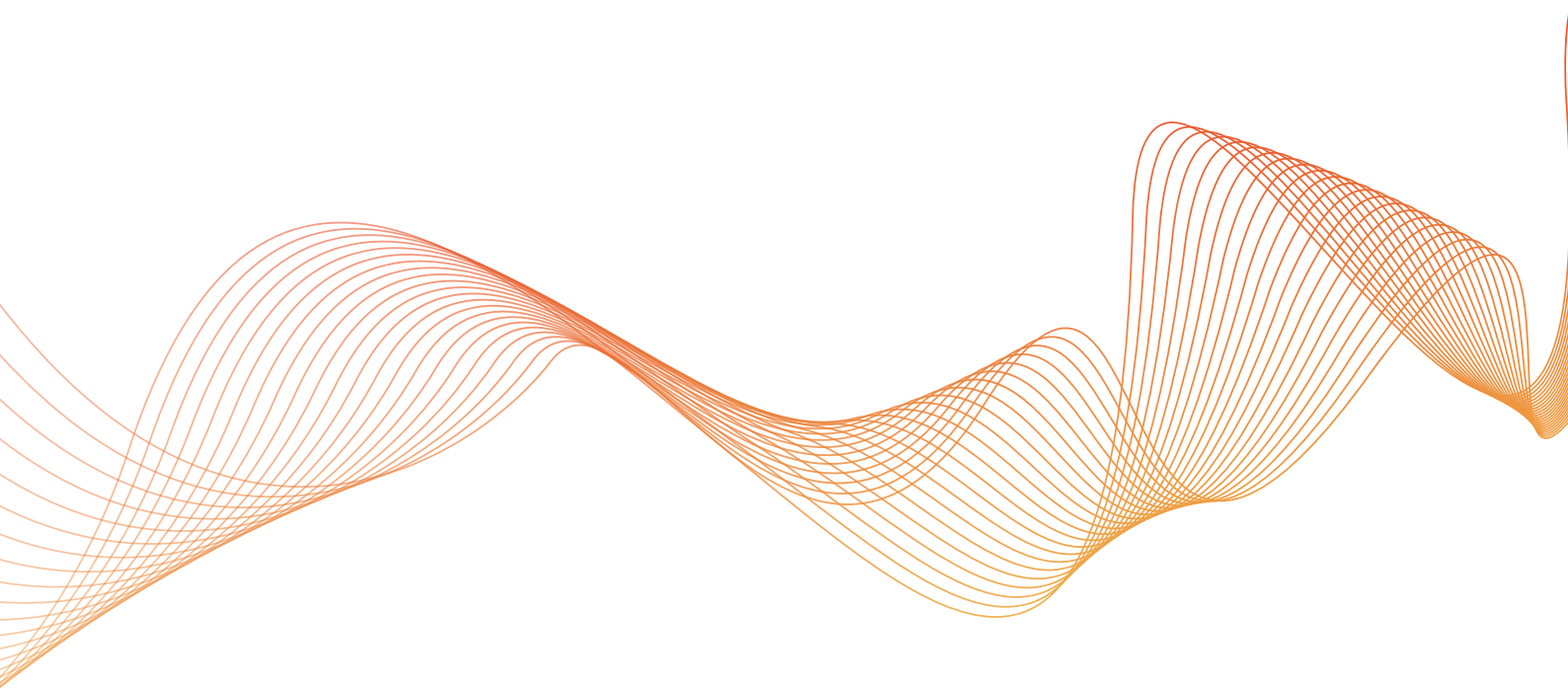
Operate between Static evaluations, Manual Inspections and Dynamic Tests Continuous Testing and Automation



## Automated journey in ETL Testing

ETL testing strategies can be protracted to blend into the trending dev and deployment methods like DevOps and CI/CD. Development cycles are high demanding on the sprint timelines and test coverage efficiency. Testing in this environment should chase along the efficiency in interactive cycles.

Test Automation is an extended scope to rationally create reusable functions and frameworks that drive the optimization of STLC timelines (Quicker time to market). It would require an Automation Expert to thoroughly analyze the reusable scenarios in the ETL Process, define the entry, exit and dependent scenarios and plug in a custom Test Framework (Continuous Integration) at the right instances.





## About Indium

Indium is an AI-driven digital engineering company that helps enterprises build, scale, and innovate with cutting-edge technology. We specialize in custom solutions, ensuring every engagement is tailored to business needs with a relentless customer-first approach. Our expertise spans Generative AI, Product Engineering, Intelligent Automation, Data & AI, Quality Engineering, and Gaming, delivering high-impact solutions that drive real business impact.

With 5,000+ associates globally, we partner with Fortune 500, Global 2000, and leading technology firms across Financial Services, Healthcare, Manufacturing, Retail, and Technology—driving impact in North America, India, the UK, Singapore, Australia, and Japan to keep businesses ahead in an AI-first world.

### USA

Cupertino | Princeton  
Toll-free: +1-888-207-5969

### INDIA

Chennai | Bengaluru | Mumbai  
Hyderabad | Pune  
Toll-free: 1800-123-1191

### UK

London  
Ph: +44 1420 300014

### SINGAPORE

Singapore  
Ph: +65 6812 7888

[www.indium.tech](http://www.indium.tech)



For Sales Inquiries  
[sales@indium.tech](mailto:sales@indium.tech)



For General Inquiries  
[info@indium.tech](mailto:info@indium.tech)

