



The integration of Fixed Income (FI) data from the new system with the existing Equities (EQ) data, and switching off the legacy FI data feed, had significant challenges.

Section 1.02 The Data Testing Challenges and Solution

For the multimillion dollar project, the key question was “How can we test hundreds of new data-feeds and thousands of existing ETL Processes and at the same time monitor progress in the most effective way? That was the key question to be answered

(a) A Need Formal Methodology for ETL and Data Testing

In a data-centric project, data quality management is as important as project management. Data, not process, is the focus of such projects. Testing has to prove that the quality profile of the production data is maintained across all components of the project. This means that a dedicated framework for testing data-centric project is a must. The iQA Solution provides multi-step action plan. The first step is to operationalize *iQA Data Quality Assurance and Testing Framework*.

Just quality assurance software is not enough. You need proper methodology and a tool that supports it. S. Gawande

The iQA Data Testing Framework has

- ✓ Clearly defined activities
- ✓ Roles and responsibilities
- ✓ Workflow and communication protocol that ties it all together
- ✓ Clear guidelines for each activity.
- ✓ Categorization of data into multiple subject areas based on functional areas, so that respective SMEs can be assigned.

Nomura Securities implemented the Testing Framework. Next, iQA was configured to support the Testing Framework to achieve goals of communication, accountability, global collaboration and visibility to management.

Key Technical Challenges in Data testing without iQA

“The beauty of iQA is in its capability to rapidly test for maximum number of test cases and data volumes.” – Nomura Securities Data Analyst

- **Manual comparisons** of huge amount of data
- Running **data quality rules across two databases** to compare source and target data
- **Absence of rule checklist** for Business Analysts and users for sign-off on ETL processes
- Create rules once; **reuse** across DEV, QA, UAT and PROD environments
- **Visibility** to management
- **Global collaboration** of team
- **Automation and scheduling** of rules in QA and UAT regions

(b) Data and ETL testing with iQA

Once the actual testing began, iQA quickly proved to be up to the job. Over a period of nine months, testers, developers and SMEs implemented roughly five thousand data testing rules of different types, including:

- Source data values compared to target data values
- Lists of value comparisons based on set theory
- Reconciliation of data both within a system and external to the system
- Expected data from transformational business rules
- Predictive testing

On each release cycle, many data issues were discovered. iQA routed those data issues and reports to developers, SMEs, and even source system users.

Initial runs quickly discovered thousands of critical and non-critical data quality defects. As the testing cycles proceeded, these issues were resolved and users could sign off on the ETL processes.

As new code was developed, previous rules were recombined to provide regression tests. Also, the same rules were reused in different environments such as UAT and integration.

iQA's web-based interface was effective for global collaboration. Access to inexpensive resources meant low cost. The simplicity and centralized control made it easy to onboard new resources and outsource testing.

iQA's automated data comparison and testing permitted testing on complete datasets instead of restricted sample test data.

Integration of iQA with HP Quality Center, which had been acquired by Nomura Securities, allowed use of this infrastructure and its accompanying methodology.

The Rules Knowledge Repository stored all the rules discovered for data testing so that they are accessible for future use. There was no need to maintain documents for testing as both descriptive metadata and actual testing results were stored in the repository.

The testing results Dashboard provided visibility to status of testing at any time. Rapid feedback in the form of reports of data with material errors together with drill-down abilities enabled quick decisions and responses.

For business users it was easy to sign off on the ETL process based on success of predetermined lists of rules.



"iQA gave us the capabilities to reach both end of the pipe and validate the data flow" – Nomura Securities Offshore Data Quality Rules Analyst

It is a typical characteristic of data-centric projects that information requirements are *not clear* at the beginning. Many of them are discovered as more data is reviewed. This gradual crystallization of information requirements is aided by iQA, with its Rules Knowledge Repository. Another characteristic of data-centric projects is that data tests must be carried into production. There is no guarantee that production data will remain stable into the distant future. Again, iQA has this capability.

Section 1.03 How the bank benefited from iQA implementation for Data Testing!

The following lessons were learned from the success of the project:

- ✦ iQA's Data Testing framework provided a **unified strategy** and **visibility** for the developers, testers, SMEs, business users, and management.
- ✦ iQA supported **test-driven development** since in data-centric projects not all business and data transformation rules are predefined - but are discovered during the development.
- ✦ Up to **20% fewer resources** were required for testing over a period of nine months.
- ✦ The ability to utilize **offshore** resources provided up to **33% in direct cost** saving for testing resources.
- ✦ iQA enabled **continuous, consistent** and **automated** testing of **all the data** passing through the ETL processes.
- ✦ The ability to get **automated** and **quick feedback** to SMEs, and the capability to **drill down** on the problem dataset led to higher efficiency.

The downstream data was used by trading, and any bad data could have had serious financial impact. Because of this, the key success factor for the success of the project was obtaining the sign-off on data quality. And with iQA, the Bank was successful in proving this success within given parameters of time, money and resources.