Automating Data Validation in BI System Migration

How Xoriant's QA automation utility enables organizations to realize the best of both worlds: fast migration to a modern BI system and data quality validation

Modern business intelligence tools like Qlik enable more people within an organization to access the same data, analyze it, visualize it and make decisions on it. These modern BI tools offer powerful additional capabilities as well, such as predictive analytics. But today, a modern BI system is not a nice-to-have; it's fast becoming a business imperative as legacy tools are being phased out by their creators.

Indeed, many enterprises are still running legacy BI tools that are no longer supported. In many cases, these enterprises could migrate to the newer versions of the same tools, but when staring down the barrel of an enterprise migration, it often makes the most sense to migrate to the most modern tool available.

Data validation (testing/QA) is an essential component of successful migration

As anyone who's gone through an enterprise data migration before is painfully aware, it's a massive project that gets very complicated very quickly, often with enormous consequences if things go wrong. Two conditions are paramount: migrations have to be fast and the data has to be accurate.

What causes data quality issues? If the underlying data is the same, why would migration from a legacy BI tool to a modern tool like Qlik risk creating data quality issues? The reason is that

> " Most BI tool manipulate the source data, which makes Manual testing a slow and cumbersome approach, so either lot of testers are needed to complete BI testing, or to only test a percentage of the data. Either choice is costly. Thus the need of Automated BI Testing "
>
> **– Data Analyst**

while the underlying data is the same and in the same source database, how it is structured in the BI tool is different. The modern BI tool will employ scripting to restructure the data in order to represent it in a more meaningful, user-friendly way. But that increases the risk of mismatched or missing data.

That's where BI testing comes in. It is the process of validating the data, format and performance of the reports, subject areas and security aspects of the governing project. The goal: build and maintain trust in the data among the people who need to use it. One of the key success factors for a BI project is the level of trust in the data depicted in reports and dashboards. Lack of trust in data causes users not to use the reports and dashboards – and then it's all for naught.

Thorough BI testing is key for ensuring the quality of the reports, which is in turn critical for user adoption. In particular, report data testing is important for assessing the reliability of information visualized on the report. User adoption – and users' trust in the reports – is so important because reports are often used as the basis for management decisions and the user assumes that the information on the report is representative of source data.

Building and maintaining user trust is clearly essential, but it's not the only benefit to BI testing. According to Oracle, "the case for data validation as part of the migration process is unequivocal." And other benefits associated with data validation can be profound:

> **Two conditions are paramount: migrations have to be fast and the data has to be accurate**

- ▶ Improved corporate performance
- ▶ Increased competitive advantage
- ▶ Maximized ROI from enterprise applications
- ▶ Efficient and effective business processes
- ▶ Reduced uncertainty and risk
- ▶ Lower costs of iterations and rewriting code
- ▶ Reduced delays and wasted time
- ▶ Reduced unexpected costs

- ▶ A measurable and accurate view of data
- ▶ Efficient and effective business processes
- ▶ Reduced uncertainty and risk
- ▶ Improved customer service
- ▶ Better accountability and ability to meet compliance targets
- ▶ Additional value for shareholders and stakeholders

But...data validation is hard

BI testing is different from traditional web application testing because the contents of BI reports are automatically generated by the BI tool based on the tool metadata. As a result, testing is often a time-consuming, manual process, which risks the two key conditions of speed and accuracy. In this Solution Guide, you'll learn about a QA automation tool that allows you to realize the best of both worlds: quick migration to a modern BI system and data quality validation.

At a Glance – What Data Organizations Need to Know

▶ Data validation (testing/QA) is an essential component of successful migration to a modern BI tool.

▶ Yet testing is often a time-consuming, manual process – which risks the two key conditions of speed and accuracy.

▶ Xoriant's QA automation utility, a GUI-based tool, automates the validation of Qlik-based reports.

The Challenge

How Do You Efficiently Ensure Report Accuracy?

Suppose for example you're migrating from Crystal Reports to Qlik. Now there has to be QA to verify that post-migration data presented in Qlik is accurate and complete. But the report has six million rows and a hundred columns; validating all that data manually is impossible. Testing becomes even more complex when there are different filters; the tester would have to manually test all combinations of filters.

A typical approach to verifying report quality is to verify calculations on the reports themselves to the data in the source report or data warehouse. Thus the BI testing process becomes a manual one, which is slow and error-prone, and expensive. Most BI tools manipulate/transform the source data, which makes it even harder to manually test.
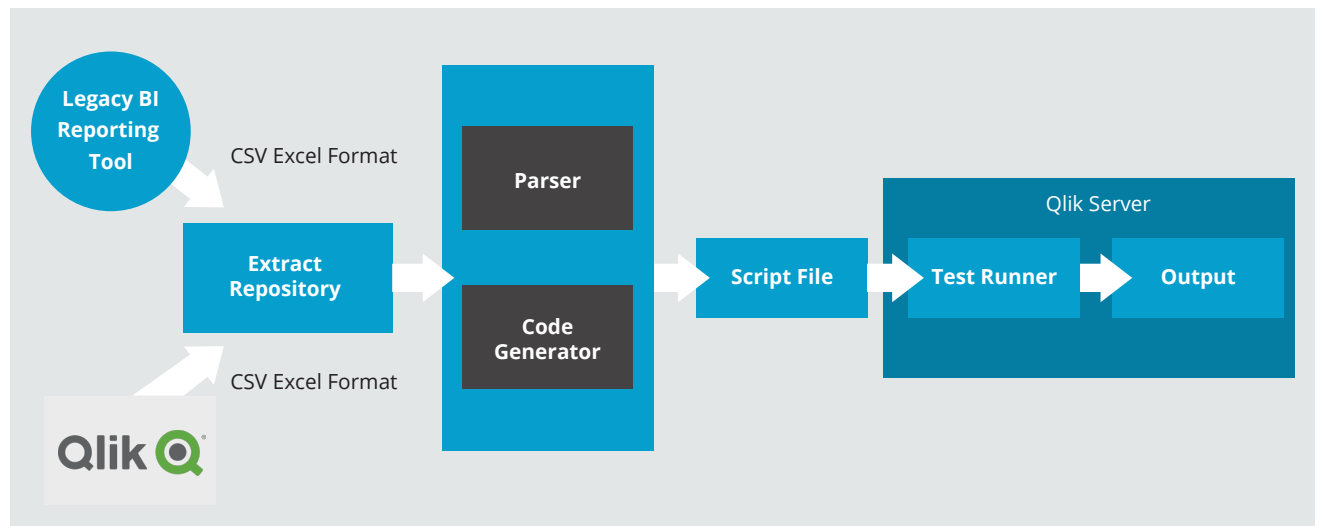
Having a manual QA process to validate reports has multiple downsides:

▶ Reports with large data cannot be accurately validated by manual process, which leaves room for potentially uncovered mismatches.

▶ Huge efforts are required for manually validating the report data.

▶ Manual efforts leave room for human error, which even if it doesn't lead to data quality issues, could lead to a lack of trust in the data.

The Solution

QA Automation Enables Fast Migration and Accurate Reports

Xoriant has developed a QA automation utility to validate Qlik-based reports. Automating testing (QA) allows you to realize the best of both worlds: fast migration to a modern BI system and data quality validation. The tool is designed to be user friendly (it's a GUI-based tool that doesn't require any scripting), so it can be implemented by BI tool users across your organization, even outside IT and the data group.

STEP 1
Data is automatically extracted from the legacy tool and Qlik

Data from the legacy BI tool (it can be any tool) is automatically extracted based on a gold dump (that is, fixed data on which both reports are working – e.g., Oracle) with all combinations of filters automatically applied. The data, in Excel or CSV format, is then stored in the extract repository. Likewise, data is automatically extracted from the migrated Qlik report based on a gold dump, again with all combinations of filters automatically applied, and stored in the extract repository in Excel or CSV format.

STEP 2
The Qlik script file for validation is automatically generated

The user identifies the dimensions or primary fields by which the rows are identified uniquely and the measure fields whose data is to be compared. The dimension and measure fields are placed in a database source system. The parser combines the data and formats it in a way that the Qlik code can be generated. The Qlik code generator engine initiates and generates the QlikView script, which is automatically stored in a script file.

STEP 3
Data from Qlik is compared to data from the legacy BI tool

Once the script is generated, the test engine automatically pushes the code to a Qlik interpretable QVS file on the Qlik Server and initiates the Qlik data comparison report load. On completion of the Qlik load, the test result is generated and displayed. A summary view provides high-level information about data discrepancies, if any, and a detailed view highlights the missing data and data mismatches in red (see Figure 2). The user has the option to export the test results into Excel.

Comparison summary provides the count of total mismatched data points, the count of mismatched data points by measure and a count of data points which are present in either source or target.

The detailed view highlights the mismatched data points in red. The Table column identifies which source the data row belongs to. (Here, OLTP data means transaction data from SAP Crystal Reports, our legacy BI tool, and OLAP means dimensional data from Qlik.)

The Benefits

QA Automation by Xoriant Saves Time & Money

The raison d'être of the Xoriant QA automation solution is to validate the accuracy and completeness of data reported by Qlik post-migration – for all the reasons already described that data accuracy and completeness are essential and migration to a modern BI tool transformative. But why Xoriant's QA automation solution specifically? There are four key benefits:

▶ **Time savings (and time is money!)**
Using the same parameters as the example above, manually validating the Qlik report data took sixteen hours. Using the Xoriant QA automation tool, it took 15 minutes. That's a savings of 98%.

▶ **Reduced BI migration project cost**
One of the most common reasons organizations delay BI migration projects is cost. Xoriant's automation tool helps mitigate that cost by automating work that would normally require technically proficient (and expensive) testers.

▶ **Maximized user adoption**
The Xoriant QA automation tool can be used by data analysts across the organization, regardless of technical skill. No scripting is required, and even the number of required clicks is minimized. That means it's more likely to be widely adopted – thereby helping you achieve the silo-busting potential of modern BI tools.

▶ **Lightweight and easy to configure**
Xoriant's QA automation solution doesn't require much processing capacity or take up much memory. Plus, it's highly configurable to meet your specific needs.

Bottom Line

Given that data inaccuracies can literally ruin a business, making any changes at all related to the data environment can seem too risky. But as legacy BI tools become unsupported and modern BI tools offer ever-more benefits – chief among them the ability to break down organizational silos and run advanced analytics – migrating to a modern BI tool quickly becomes a business imperative.

There are two conditions that will determine the success of such a migration effort: speed and accuracy. The migration project has to be fast, and the data reported from the new tool has to be accurate. Hence the importance of data validation – the case for which is "unequivocal," as Oracle puts it. With the solution that automates post-migration QA, Xoriant makes that process easier, cheaper and better.

Learn how Xoriant can enable fast migration and accurate reports in your organization:

Connect with us:

www.xoriant.com/technologies/big-data-analytics   info@xoriant.com   +1 408-743-4477

XORIANT

Xoriant is a Silicon Valley based product engineering, software development and technology services firm with offices in the U.S., Europe and Asia. For both technology companies and enterprises, from startups to the Fortune 100, we leverage our expertise in emerging technologies and our highperforming teams to deliver innovative solutions that accelerate time to market and keep our clients competitive.

Across all our technology focus areas – Product Engineering, DevOps, Cloud, Infrastructure & Security, Big Data & Analytics, Data Management & Governance, Digital and IoT – every solution we develop benefits from our product engineering pedigree. For 25 years and counting we have taken great pride in the long-lasting, deep relationships we have with our clients. Learn more at **www.xoriant.com.**