



# Eliminating Data Related Defects

## How to Speed Up the Elimination of Data-Related Defects in Your SDLC Workflow

Every enterprise has business and technology initiatives tied to its ongoing digital transformation efforts. This requires the ability to quickly innovate, develop, and iterate on applications and features to gain a competitive edge and align with changing customer requirements. With so much depending on the successful delivery of applications, developers and quality assurance teams set high standards for the number and quality of releases. One of the key bottlenecks to achieving these goals is time and resources spent to address data related defects which are generally found during later QA/test stages of SDLC workflow.

Data related defects are the defects introduced in SDLC workflows due to testing done with data which is not fully representative of enterprise production data. It leads to software systems break down due to unanticipated, incoming data exercising the software in unexpected ways. It causes software release delays or production downtime scenarios due to insufficient use-case test coverage with ever-growing production data and evolving applications. Consequently, AppDev and QA/Test teams are always looking for new ways to improve test data management within the entire SDLC workflow that not only helps meet software quality success criteria earlier but also frees up time to develop newer features and produce more feature releases.

---

*Consequently, AppDev and QA/Test teams are always looking for new ways to improve test data management within the entire SDLC workflow that not only helps meet software quality success criteria earlier but also frees up time to develop newer features and produce more feature releases.*

---

---

*Enterprises often have stale, outdated data in their application development (AppDev) and quality assurance (QA) environments, which becomes a barrier to achieving software quality goals.*

---

## Challenges With Data-Related Defects

For most enterprises, reproducing the production data environment copy or synthetic data environment repetitively is time-consuming, non-scalable and infrastructure intensive for each step of SDLC workflow. Enterprises often have stale, outdated data in their application development (AppDev) and quality assurance (QA) environments, which becomes a barrier to achieving software quality goals. If QA has the wrong data or just can't get production-quality data quickly enough, unnecessary work is created and project releases get delayed. And often, these tasks have tight interdependencies which delay other project tasks. When QA and developers wait for data, they spend a lot more time context switching, making very expensive incremental progress and solving unnecessary errors. Without the right data, false positives and false negatives are found. Errors are detected later in the pipeline or even in production as production-quality data is not easily available in the earlier SDLC phases. Finding data-related defects later in the SDLC workflow lowers software quality, extends the release cycle and increases the cost to fix these defects.

Enterprises try to get production-quality data in the SDLC workflows earlier to help detect and prevent data-related defects faster. Enterprises manage these challenges by adopting the following approaches:

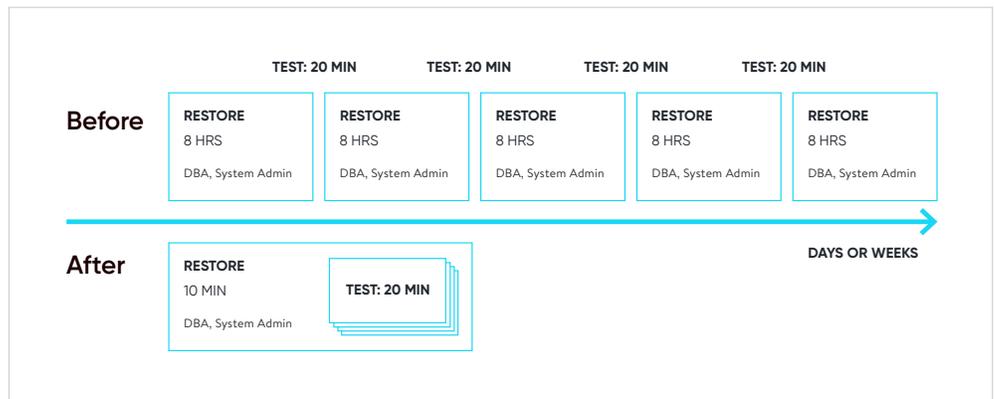
- **Provide full production copies of different data sources earlier in SDLC workflows.** Full-size production copies for test environments require longer provisioning times and consume additional storage. Both of these problems are amplified when multiple data sources are involved in testing. To bypass these objections under aggressive timelines, developers and QA teams often stick with data which is not refreshed often or has a quarterly refresh cycle. Further, the associated deployment time and cost of using old and existing data is not negligible. But this pushes defect discovery and fixing to later stages of the SDLC, which is more expensive and causes significant delays within the software release cycles.
- **Share production-quality data environments among team members and across the organization.** The inability to provision dedicated QA environments to individual team members because of cost, time and resources often leads teams to share datasets. The concurrent access to a common data environment results in conflicts and unnecessary scheduling and planning for data accessibility when there is more than one stakeholder who contends for the same resources all at the same time. The result is often chaotic test environments with longer development cycles, test cycles, and lower quality releases.
- **Subset data to accelerate data delivery.** Data subsetting partially addresses the challenges of copying and moving full, production-sized datasets to lower tier environments earlier in the SDLC workflow. But subsets fail to adequately embody the breadth of real-world data conditions where the applications are extensively used. Relying on subsets prevents QA teams from identifying edge cases and outliers, such as defects that manifest themselves only when multiple parameters reach extreme levels. Subsetting also delays the defect discovery in the SDLC workflow. During the later stages of the SDLC and closer to the production release, AppDev and QA teams run software test cycles against full production data copy, leading to longer cycle times to fix the bugs and delaying software releases.

## Strengths of Using Data Platform

Instead of taking shortcuts, enterprises are adopting a data platform approach that not only addresses immediate test data management challenges, but scales with increasing application releases, and growing developer and QA teams. A data platform can support in the following ways:

- **Accelerate the delivery of lightweight production-quality and synthetic data copies.**

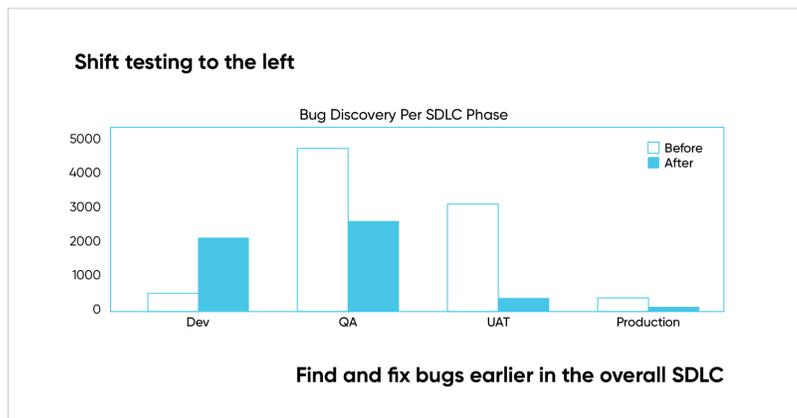
Developers and QA teams can easily get access to lightweight production-quality and synthetic data as part of their SDLC, allowing for high-quality testing cycles earlier in the workflow. By shifting left on testing, teams can identify defects when they are less expensive to fix, eliminating future rework that often contributes additional cost to the software development budget.



- **Provide personalized production-quality and synthetic data environments with data controls.**

Every developer and tester can have his or her own data environment with the necessary data controls to manipulate environments according to his or her respective requirements. For example, they can bookmark a data environment and then create a branch to test a special use case with it as they are working on other defects within the same environment.

- **Create a positive impact on engineering efficiency.** As data-related defects are eliminated earlier in the SDLC, developers can spend more time on new feature development instead of defect resolutions late in SDLC workflow. Likewise, QA teams can add more test cases and have more exhaustive test coverage in each software cycle.



## How Delphix Can Help

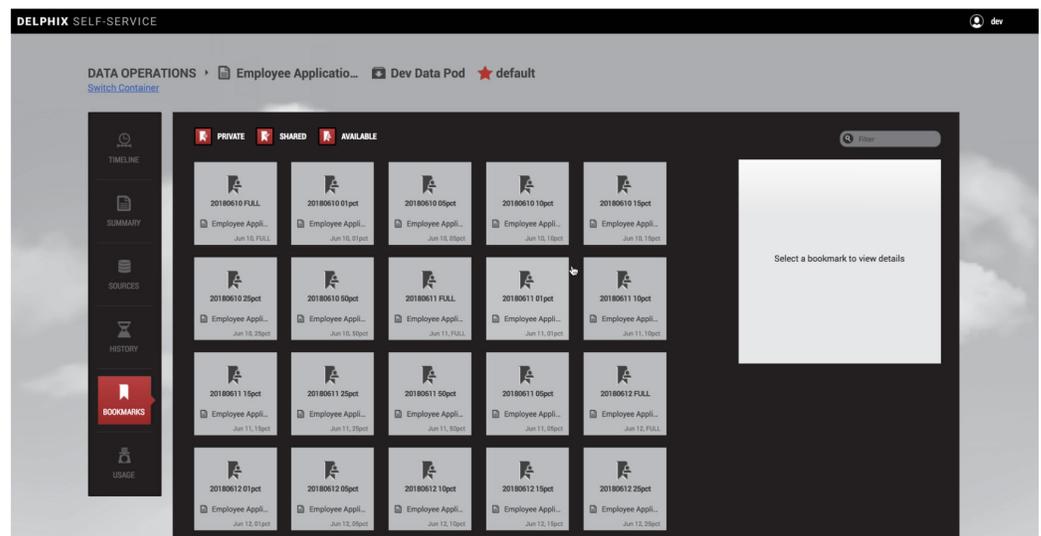
Delphix as a single data delivery platform for production-quality and synthetic data provides access to the right test data at the right time to the right stakeholder. Each team member has his or her own personalized environment with the following data controls that help identify and eliminate data-related defects earlier in the SDLC:

### INSTANTLY REFRESH PERSONAL DATA ENVIRONMENTS

Developers and QA teams can refresh their personal test data environments from the latest version of production data in minutes instead of waiting for days or weeks. Instant availability of production-quality data allows developers to uncover and resolve defects in real-time without going through an IT ticketing process to request data environments and waiting for them to become available.

### BOOKMARK AND SHARE TEST DATA WITH TEAMMATES

Bookmarks of data environments associated with defects can be instantly shared across teams to improve collaboration and reduce the time to resolve defects.



### RESET TEST DATA BACK TO A PRIOR STATE WITH A SINGLE CLICK

Synchronize test data from multiple, heterogeneous sources to the exact same point in time for integration testing of distributed applications.

## Fortune 500 Financial Services Enterprise

Here's an example of a Fortune 500 financial services enterprise that was challenged with a number of data-related defects in SDLC. Learn how they addressed their issues and the results they achieved leveraging Delphix Dynamic Data platform.

One Fortune 500 financial services institution used Delphix for the development and testing of their online platform that provides market insights to clients and enables them to make smarter financial decisions. The investigation was triggered by a massive platform growth: Over a span of just a few years, the company experienced massive growth, where financial data had doubled, usage had tripled, and development efforts had quadrupled.

A large percentage of defects were discovered in late-stage user-acceptance testing, which created a high risk of impacting the customer experience. Database teams were struggling to keep pace with the increasing number of data sources and associated exploding storage costs. Teams also missed several releases because of slow environment provisioning.

To solve these challenges, the company's IT organization implemented the Delphix Dynamic Data Platform for both their Oracle and MS SQL production data sources. After deployment, the results were immediately visible.

Rather than waiting a full day for a DBA team to restore an environment following a 20-minute test run, dev and QA engineers leveraged Delphix's self-service capability to initiate a 10-minute reset process that brought the environment back to a bookmarked state.

Less waiting times enabled QA teams to execute more test cycles earlier in the SDLC using a "shift left" approach to testing. Ultimately, this led QA teams to discover and resolve errors when they were easier and less expensive to fix.

The firm estimated they reduced overall defect rates by 60 percent and improved productivity across 800+ developers and testers by 25 percent.

### ABOUT DELPHIX

Delphix's mission is to free companies from data friction and accelerate innovation. Fortune 100 companies use the Delphix Dynamic Data Platform to connect, virtualize, secure and manage data in the cloud and in on-premise environments. The Delphix Dynamic Data platform serves as the foundation for DataOps across hundreds of the world's leading enterprises.

© 2019 Delphix Corp. All rights reserved. | [delphix.com](https://delphix.com) | [info@delphix.com](mailto:info@delphix.com) | +1.650.494.1645  
04-2018