



## TEST DATA MANAGEMENT

## Introduction

The software industry is constantly looking for optimization in testing efforts, and one such area is Test Data Management (TDM). It is crucial because testing completeness and coverage mainly depend on test data quality. Efficient testing data management is essential to maximizing return on investment and supplementing testing efforts for the highest success and test coverage. If the data used fails to promote ease of use and adaptability, it adversely impacts the desired outcome and degrades the product quality. To balance the positive results and improve returns, consider the process, potential challenges and possible solutions involved in TDM. The best data is always found in production, as this is where it is used. While using production data, creating a subset of this data is always advisable to reduce the effort involved in test planning, execution, and optimization. Depending on the business, privacy and legal concerns may

surface when using production data. Hence, we recommend not using production data for testing purposes.

In the testing phase, the data is often incomplete; therefore, real-time scenarios are not being tested completely. Testing cycles are shrinking with the advent of Agile and DevOps practices. Creating quality data within that cycle, as well as performing software testing, becomes a challenge.

If you don't have a TDM or are keen to make your test data more robust and production-like, this whitepaper will help you understand and implement TDM. It explains Infosys' TDM strategy and solution encompassing data subsets and synthetic data that reduce cost, time, and effort in the testing cycle while establishing a cost-efficient software deployment process. The practice of not including TDM in the testing life cycle often leads to ignorance about TDM on the part of the software development team.

## IMPORTANCE OF TDM

In today's fast-paced digital world, software testing is crucial to ensuring the quality, performance, and security of any software application. However, using actual production data for testing can expose the system to significant risks, such as data breaches, performance issues and compliance violations. That's where TDM comes in. The quality of the data you feed your tests is more important than the test process itself.

Test data and TDM are more critical. Most companies estimate up to 15% of their testing efforts just to get test data for the backend and frontend systems. There is always a hidden cost associated with data management. All data management tasks included moving data from backend systems to identify test data, data

masking of sensitive data, skipping production defects due to the unavailability of correct test data, manipulation of data for different scenarios, and storage of test data. TDM has, therefore, become a big problem. The functionalities and data usage work perfectly during staging and testing in several scenarios. However, releasing the code into production can lead to intermittent errors due to inefficient TDM.

TDM solutions help organizations accelerate application development speed, code quality, data compliance, and sustainability initiatives by providing timely access to fresh, relevant data downstream for code development, automated tests, troubleshooting and validation



## CHALLENGES WITH TEST DATA

TDM seamlessly manages test data in a software testing project to ensure all test scenarios can access the correct data whenever needed. With TDM, test data is provided in the best format required for test activities and in the right volume to meet all unique testing needs. Different test cases may require data of various types. But there are underlying guarantees that every test data needs to provide before use for test activities, like the quality, availability for use, timely access, realistically real-life use data, and compliance to any regulations that govern test data usage. There are several challenges in TDM:

- Additional time for data setup
- Less realistic data during QA testing
- Additional administrative efforts in TDM

- Sensitivity of private information
- Storage required for test data
- Data masking issues
- Risk of data loss
- Lack of data consistency
- Handling data requests
- Identification of data anomalies.

As businesses increasingly rely on data-driven decision-making, addressing the challenges of managing and accessing test data effectively is crucial. As a result, TDM is a demanding task, considering today's hyper-competitive and complicated digital environment.

## INFOSYS APPROACH FOR EFFICIENT TDM

With ever changing business and technology landscapes, traditional services try to address the paradigm shift in TDM. Still, there is much scope for improvement due to agile and digital transformation, creating a demand for next-gen TDM services to address the gaps. Infosys has helped several clients deploy TDM, resulting in improved data coverage, data setup effort reduction, data provisioning, and a significant reduction in data refresh cycle time and data provisioning effort.

At Infosys, we follow an efficient TDM strategy that helps to avoid fundamental problems like reducing redundancies and seamless management of copies. By eliminating redundancies, effective

TDM lowers storage costs. Additionally, with an efficient TDM, the team can use effective algorithms to sort and identify redundant or duplicate test data and delete unnecessary data, thereby streamlining and managing large volumes of data.

There are various steps involved in TDM: analysis of data requirements, copying production data, data sub-setting, data masking, synthetic test data generation, and data maintenance. Infosys follows standardized processes in each phase, making them much more effective and targets on time, enabling testers to cover more tests in the available time.

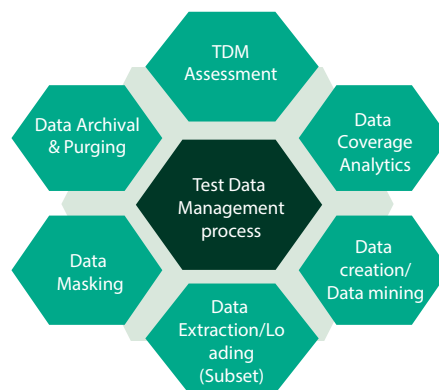


Figure 1 What TDM involves



Infosys ensures data protection, maintains consistency, and creates standardization. The approach is driven by key principles and best practices. Some critical areas considered in our approach include: Test governance and collaboration

- Template based data requests
- Data masking and sub-setting
- Synthetic data generation
- Test data locking system
- Data copy between environments
- Reports for data requests and status
- Intelligent test data generation

Infosys Enterprise Data Privacy Suite (iEDPS) is an enterprise class, cross-platform, data privacy solution. iEDPS helps in preserving data privacy within enterprises by protecting sensitive data. Infosys Enterprise Data Privacy Suite (iEDPS) is a patented enterprise data management solution that leverages intelligent

automation to expedite data protection and modernization. It strikes a balance between utilizing data for growth and safeguarding individual privacy. The suite helps enterprises discover sensitive data, design the right data protection controls, and deliver privacy-compliant data sets to non-production environments. This ensures that organizations can leverage their data assets for business growth while ensuring compliance with data privacy regulations. By providing a comprehensive view of data privacy, iEDPS enables businesses to make informed decisions about data usage and protection. Overall, iEDPS is a robust tool for managing enterprise data privacy and protection.

Data Masking (refer to Figure 2) supports 70+ algorithms. The Infosys TDM tool has numerous capabilities to ease test data management. It has AI-powered TDM for auto identification of data catalog and optimal data volume, can be integrated with CICD and various Rest API interfaces for easy integration with external tools, and provides a service dashboard for anomaly detection.

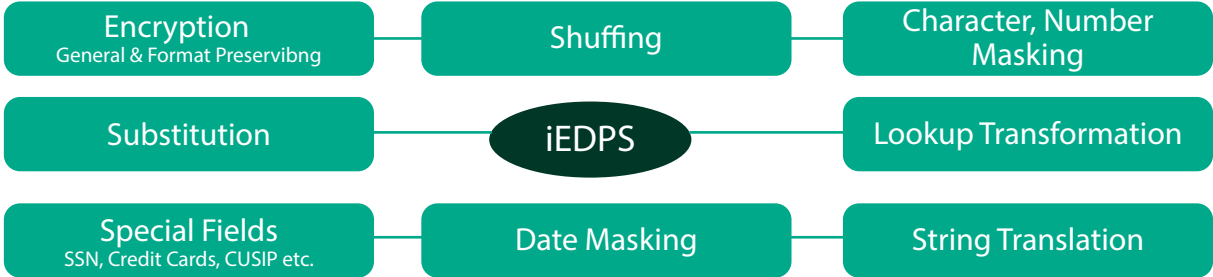


Figure 2 iEDPS components

## INFOSYS CAPABILITIES

01

Test data will be managed so that it remains clean and clear

02

Data protection is given prime importance as client data is protected via masking during test data generation while retaining its original properties

03

We ensure that the testing occurs in a production-like environment, thus decreasing bugs, fixes, and rollbacks that escalate the overall cost

04

We synchronize and handle data in diverse formats and from multiple sources for integrated testing and utilization.

05

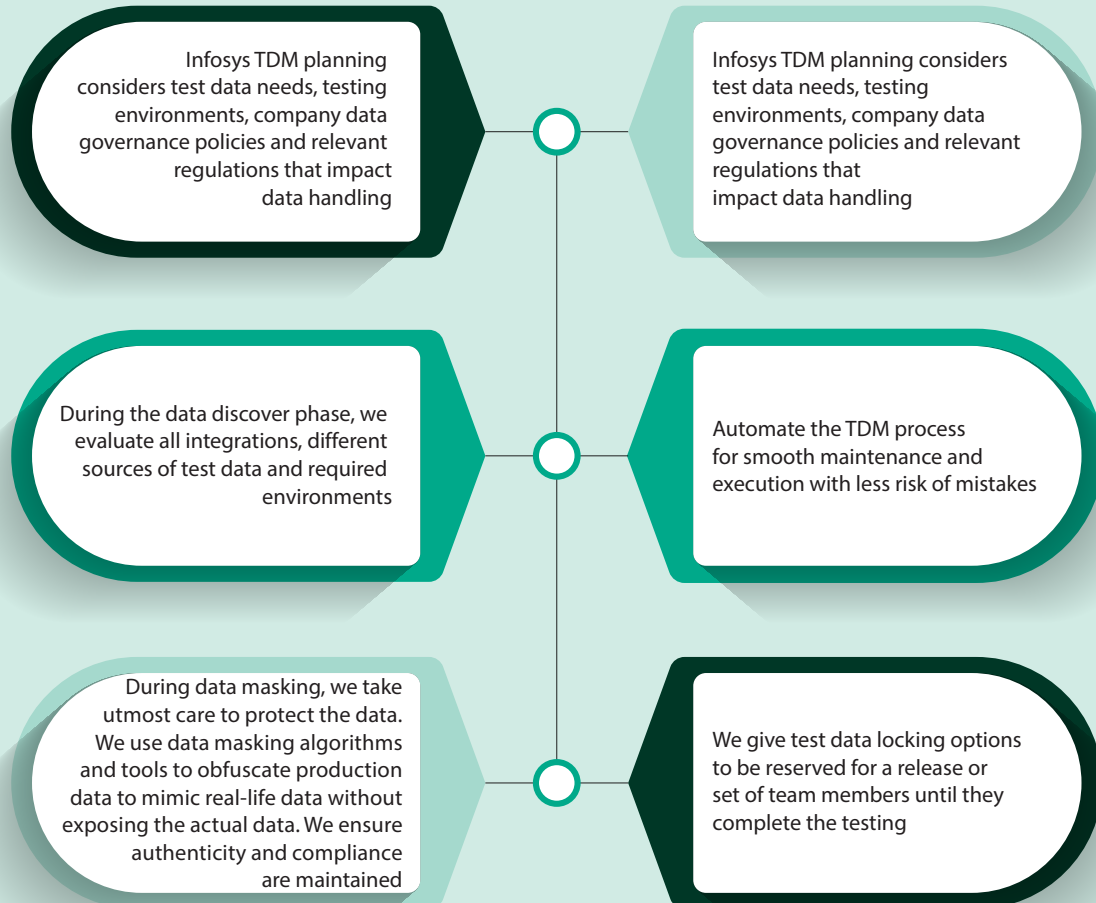
Infosys Learning Platform Lex has TDM training and articles on multiple domains and databases.

06

Infosys TDM Suite is a web-based tool for centralized test data management. It provides a single, easy-to-use interface for testing and data provisioning teams. The tool comes with test data generation, extraction, refresh, masking, sub-setting capabilities, and the data request-based workflow for end-to-end TDM process management.



## BEST PRACTICES



## CONCLUSION

TDM enables organizations to meet the test data needs of testing teams by assuring that the test data is of the right quality and quantity with the appropriate format and proper environment. This ensures that the provided data is referentially intact, of the correct size, not too large or small, and aptly fits the testing requirement. The data can be provisioned by synthetic data creation, production extraction or sourcing from lookup tables.

Infosys has served multiple clients in establishing TDM with client-licensed tools and is experienced with open-source tools. Infosys has developed an enterprise-level license tool, in-house tools, and accelerators, which provide more capabilities with the highest security to data, resulting in meeting user expectations and product releases with the highest quality.

# About the Authors



**Sai Krishna Mohan Chitikam**  
Principal Consultant, Engineering Services, Infosys

---



**Dr Sumit Goyal**  
Delivery Manager, Engineering Services, Infosys

---



**Krishna Markande**  
Associate Vice President, Senior Principal Technology Architect,  
Engineering Services, Infosys

---



For more information, contact [askus@infosys.com](mailto:askus@infosys.com)



© 2024 Infosys Limited, Bengaluru, India. All Rights Reserved. Infosys believes the information in this document is accurate as of its publication date; such information is subject to change without notice. Infosys acknowledges the proprietary rights of other companies to the trademarks, product names and such other intellectual property rights mentioned in this document. Except as expressly permitted, neither this documentation nor any part of it may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, printing, photocopying, recording or otherwise, without the prior permission of Infosys Limited and/ or any named intellectual property rights holders under this document.