THE NEXT GENERATION OF TEST DATA MANAGEMENT

GenRocket's synthetic data platform creates a new category of Test Data Management (TDM) that we refer to as Synthetic Test Data Automation (TDA). This new and innovative approach automates and accelerates many cumbersome aspects of traditional TDM. It also removes the limitations of other synthetic data platforms that produce a synthetic data replica of a production database to provision test data.

Let's clearly define each TDM category and compare the way each one addresses the most important elements of an enterprise-class test data solution.

Synthetic TDA (GenRocket)

Synthetic TDA is unlike any other form of TDM. It brings the ability to *model and design* any type of test data for any type of test based on pre-defined rules. *Controlled and conditioned* synthetic data is defined by a light-weight instruction set in an executable **Test Data Case.** This instruction set is used to generate a fresh copy of synthetic data in real-time as automated tests are run in the CI/CD pipeline.

TDA allows data to be instantly provisioned by testers using a self-service platform in the volume and variety needed to achieve full test coverage. And TDA is affordably priced according to the number of data environments that are modeled, offering unlimited data generation for each data environment modeled by a *Test Data Project.*

Traditional TDM (e.g., IBM, Broadcom, Informatica)

Traditional TDM is the familiar model for provisioning test data commonly in use today. It has formalized the process of copying and masking a subset of a production database to make it ready for testing. In traditional TDM, test data is often reserved by a tester for a given test and refreshed prior to its next use.

While test data provisioned in this manner is realistic, it's not conditioned to meet the needs of a given test case. Testers must query the test database for the data they need or augment it with manually created data. This can be very costly in terms of time and resources.

Synthetic TDM (e.g., Tonic, Hazy, Mostly AI)

A recent evolution of the traditional TDM paradigm is the emergence of Synthetic TDM. These platforms use machine learning to examine a production database and use synthetic data generation as a data masking technique. Some of these tools produce a statistically equivalent replica of the entire database using synthetic data. Both approaches are alternatives to the traditional TDM data masking process while still eliminating the use of sensitive data (PII or PHI). However, both approaches still have the same limitations in data volume, variety and conditions that are evident in masked production data from traditional TDM systems.

Both traditional TDM and synthetic TDM tools also require dedicated infrastructure for hosting and data storage where GenRocket only requires a small, light-weight Java Runtime and Repository to run and store *Test Data Case* instruction sets.

How GenRocket Compares with Traditional and Synthetic Test Data Management

Let's compare the way GenRocket's Synthetic Test Data Automation Compares with Traditional TDM and Synthetic TDM. The table below contains the most important aspects of TDM and describes how each category of test data management platforms address them.

Capabilities	GenRocket	Traditional TDM	Synthetic TDM
Price	**$25,000 - $100,000** The price range for the GenRocket synthetic TDA platform with unlimited data volume	**$100,000 - $1M** The price range for traditional TDM systems where some license fees increase with data volume	**$50,000 - $300,000** The price range for many of the new synthetic TDM platforms where some license fees increase with data volume
Data Variety	**Unlimited** Generate any variety of new and unique data based on specific rules & conditions regardless of what is in the production database	**Limited** Data variety is limited to what is in your masked production data subset	**Limited** Data variety is limited to what is in your synthesized production data subset
Data Volume	**Unlimited** Generate any volume of data in seconds to minutes, on demand, as needed by each test case	**Limited** Volume is limited to what is available in the production database. Production data has many gaps and does not meet all test case requirements.	**Limited** Volume is limited to what is available in the synthesized production database copy. Production data has many gaps and does not meet all test case requirements.
Data Output Formats	**Unlimited (100+)** GenRocket offers the most test data formats in the industry	**Limited** Limited test data formats are supported. Primary focus of traditional TDM is on inserting data into databases.	**Limited** Limited test data formats are supported. Primary focus of synthetic TDM is on inserting data into databases.
Dynamic Data	**Yes. Data that changes state during workflows** GenRocket data can be dynamic; test data rules are easily created to control the state and condition of the data at any point in the testing process	**No Static Data Only** The data is static; only the data values already contained in the production data base are available for testing.	**No Static Data Only** The data is static; only the data values already contained in the production data base are available for testing

Capabilities	GenRocket	Traditional TDM	Synthetic TDM
Data Storage Cost	**Low Storage Cost** By design, test data is delivered in real time for each test case and does not need to be stored. This "data on demand" approach can lead to huge costs savings for data storage.	**High Storage Cost** Test data is maintained in many databases in the lower environment; for larger organizations there is a substantial data storage cost	**High Storage Cost** Test data is maintained in many databases in the lower environment; for larger organizations there is a substantial data storage cost
Data Provision Time	**Seconds / Minutes** GenRocket integrates the volume and variety of test data needed directly into the test case. For a typical functional test case, data is delivered in 100 milliseconds. For other tests, data is delivered in seconds to minutes.	**Hours / Days** Traditional TDM does not integrate the volume and variety of test data directly into the test case; data is delivered to a database forcing developers and testers to hunt for and modify the test data they need – a slow process.	**Hours** Synthetic TDM does not integrate the volume and variety of test data directly into the test case; data is delivered to a database forcing developers and testers to hunt for and modify the test data they need – a slow process.
Data Security	**No Production Data is Accessed or Stored** GenRocket never copies or stores production data. Metadata is used to model a production database as a Test Data Project. Test Data Cases are designed to generate test data for test cases in the lower environment	**Production Data is Accessed, Stored & Masked** Sensitive production data must be copied and stored in the TDM system. Then it is profiled, masked, and transferred to a database in the lower environment.	**Production Data is Accessed, Stored & Synthesized** Sensitive production data must be copied and stored in the TDM system. Then it is scanned, analyzed, synthesized, and transferred to a database in the lower environment.
Deployment Complexity	**Medium** GenRocket implemention is more automated and nimble than traditional TDM systems	**Very High** Traditional TDM systems are highly complex and cumbersome systems that are known to take up to 18 months to deploy.	**Medium** Synthetic TDM systems are more automated and less cumbersome than traditional TDM systems
Data Profiling	**Not Required** Designed and generated synthetic data is 100% secure, by definition, and does not require profiling to detect PII / PHI.	**Required** Data must be profiled to identify sensitive data and appropriate subsets	**Required** Data must be profiled to identify sensitive data and appropriate subsets
Data Masking	**N/A – Synthetic Data is 100% secure** Because synthetic data is not real data, there is no need to mask it.	**Masking Required** Sensitive production data must be carefully masked prior to its use for testing.	**N/A – Synthetic Data is 100% secure** Because synthetic data is not real data, there is no need to mask it.

Capabilities	GenRocket	Traditional TDM	Synthetic TDM
Data Reservation	Not needed, by design With GenRocket, fresh data is generated for each test run, so each developer or tester gets what they need and there is no need to reserve it.	Yes Because testers often share the same test data, it must be reserved to ensure its integrity for different tester's tests.	Yes Because testers often share the same test data, it must be reserved to ensure its integrity for different tester's tests.
Data Refresh	Always fresh, by design With GenRocket, fresh data is generated for each test run, so there is no need to refresh it.	Required Because the testing process changes data values in a shared test database, it must be refreshed frequently to ensure data validity.	Required Because the testing process changes data values in a shared test database, it must be refreshed frequently to ensure data validity.
Direct CI/CD Test Case Integration	Yes Developers and testers can quickly find categorized Test Data Cases in a self service Portal and integrate them into their test cases. GenRocket *Test Data Cases* are called by each test case, delivering specific volume and variety of data to each test as part of an automated CI/CD pipeline.	No Traditional TDM does not integrate the volume and variety of test data directly into the test case; data is delivered to a database forcing developers and testers to hunt for and manually modify the test data they need – a slow process. Data is not easily integrated into a CI/CD pipeline	No Synthetic TDM does not integrate the volume and variety of test data directly into the test case; data is delivered to a database forcing developers and testers to hunt for and manually modify the test data they need – a slow process. Data is not easily integrated into a CI/CD pipeline

LEARN MORE

REQUEST A DEMO