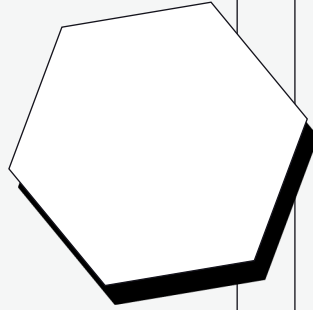


CASE STUDY

Leading HealthTech Software Provider

Automated Test Data Generation and Masking Automation for a Pathology Software Provider



SYNTHESIZED

Overview

A well-known HealthTech software provider chose Synthesized to unlock mission-critical performance testing and compliant support debugging using production-relevant but fully anonymized data. The team turned test data from a blocker into a safe, on-demand asset that supports both engineering and offshore support operations.

Technical Challenge

The company's legacy Laboratory Information System (LIS) monolith suffered from performance issues in production because there was no way to test against large, representative datasets matching real customer environments of 30–200 GB. At the same time, a high volume of support tickets was driven by issues that the offshore support team could not reproduce because they were blocked from accessing live customer data for regulatory and security reasons. This was particularly critical because the LIS orchestrates end-to-end sample-to-diagnosis workflows, where any downtime or slowdown directly affects patient care timelines and lab revenue. Support engineers were forced to handcraft data scenarios to debug issues, a process that was slow, inconsistent, and still failed to expose deeper performance bottlenecks.

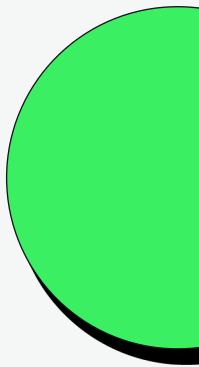


TDM Requirements

The HealthTech provider defined clear requirements for any new test data management approach.

They needed to:

- ① Work with very large relational databases (30–200 GB) without degrading performance or sacrificing referential integrity across complex LIS schemas.
- ① Automatically detect and anonymize PII across nested JSON fields, documents, and binary blobs, ensuring zero patient or customer data exposure in non-production.
- ① Generate production-realistic, human-readable test data for performance testing, regression, and live debugging, while preserving statistical properties of real workloads.
- ① Provision compliant datasets to the offshore support team so they could reproduce and fix high-priority issues without accessing raw production data.
- ① Select a mature, responsive vendor that could handle their scale and timelines, with strong support and pricing flexibility, after evaluating alternative test data management tools.



The Solution

The HealthTech company implemented Synthesized Platform as an AI native test data automation layer on top of their production databases. Using Synthesized, the team:

- ✓ Ingested and anonymized full-scale customer databases in the 30–200 GB range, using automated sensitive data discovery to find PII even in JSON fields, documents, and other complex structures.
- ✓ Applied privacy by design transformations that removed all PII while preserving relational integrity and business logic, ensuring the anonymized datasets behaved like the original LIS environments.
- ✓ Generated synthetic, human-readable values (for example, realistic names and addresses) to keep test data understandable for engineers and support staff while remaining fully compliant.
- ✓ Stood up representative performance test environments and ran load and regression tests at a realistic scale, surfacing previously hidden performance bottlenecks and enabling focused tech debt reduction.
- ✓ Provisioned safe, anonymized copies of customer datasets for the offshore support team, who could now live debug critical tickets against realistic data without breaching data residency or privacy controls.

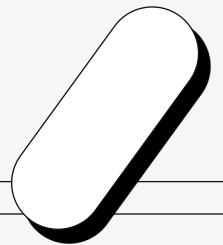
This approach aligned directly with Synthesized's core strengths: AI-powered schema analysis, automated masking, and production-quality data generation at scale, delivered through a unified, automation-ready platform.

Proof and Impact

By adopting Synthesized, the HealthTech provider significantly reduced customer frustration associated with recurring performance issues and slow ticket resolution on a mission-critical LIS where “seconds count” for pathologists. The engineering team identified and addressed deep performance bottlenecks that had not been visible with smaller or handcrafted test datasets, improving system responsiveness under real-world loads.

The offshore support team moved from guesswork and manual data creation to precise, data-driven debugging on anonymized copies of real environments, accelerating fixes for high-priority incidents. While specific metrics were not available, internal feedback from support and professional services teams confirmed a marked drop in frustration and faster resolution of critical issues.

For this HealthTech company, Synthesized proved to be a trusted foundation for future expansion as they explored scaling test datasets beyond their largest customers to de-risk growth and future product changes.



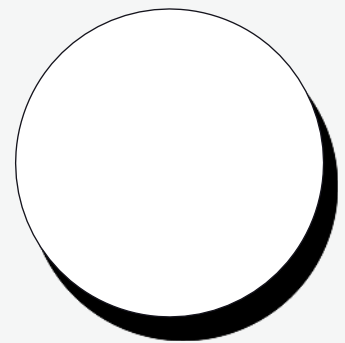
Why Synthesized Was Chosen

DIMENSION	CUSTOMER NEED	WHY SYNTHESIZED WON
Large-scale performance testing	To test against 30–200 GB datasets that mirror real customer characteristics	AI-native test data automation designed for production-realistic data generation at scale with guaranteed referential integrity.
Compliance and privacy	Ensure zero exposure of patient and customer PII in non-production, including complex JSON and document fields.	Automated PII detection and masking with privacy-by-design synthetic data that preserves statistical properties without retaining identifiers.
Supports debugging	Enable offshore support to reproduce and debug live issues without access to production data.	Safe anonymized copies that behave like real customer environments, enabling live debugging while staying within residency and security constraints.
Product maturity	A robust, enterprise-grade TDM platform that was evaluated against competitors such as Syntho.	Unified, AI-native platform with proven automation, schema intelligence, and enterprise readiness across regulated industries.
Vendor partnership	Responsive support and flexible pricing to match the customer's engineering growth and evolving requirements.	Close collaboration, rapid support, and a roadmap aligned with continuous test data automation.

Conclusion

By deploying Synthesized Platform as an AI-native test data automation layer, the company automated PII detection and anonymization at scale, preserved LIS business logic, and generated human-readable, production-realistic data for performance testing and compliant offshore debugging.

The result was a shift from test data as a blocker to a safe, on-demand asset that improved system performance, accelerated issue resolution, and provided a scalable foundation for future growth.



About Synthesized

Synthesized has redefined software quality by pioneering AI-native test data automation. Recognized as a Gartner Cool Vendor and finalist for multiple AI innovation awards, Synthesized serves global enterprises across financial services, insurance, healthcare, telecommunications, government, and technology.

Co-headquartered in London and New York, Synthesized's "Data as Code" approach transforms test data management from an operational liability into a strategic enabler of continuous AI-driven software testing and delivery excellence.