



www.enov8.com

Po Box 338, St Leonards, NSW 1590, Australia

Tel: +61 2 8916 6391 Email: enquiries@enov8.com

Enov8 Test Data Manager

The company

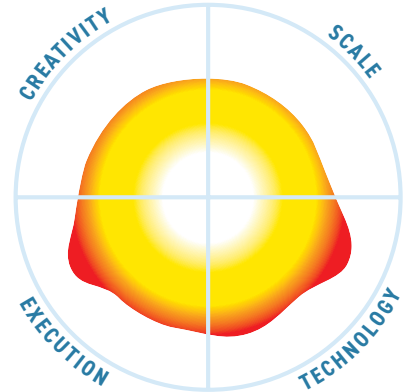
Enov8 is an Australian software vendor offering products that help its customers enhance their DevOps and testing capabilities. Originally spun off in 2013 from an IT services company specialising in this domain, Enov8 now serves large enterprises across various geographies and industries, with a particular focus on financial services.

What is it?

As shown in **Figure 1**, Test Data Manager sits on top of the core Enov8 Application Portfolio Management layer, alongside Release Manager and Environment Manager. As its name implies, it is designed to help you manage your test data, addressing common issues like test data bottlenecks, privacy, and so on. It accomplishes this by using automation to simplify, systematise and accelerate your test data management processes, such as data profiling and data masking. This automation comes in various forms, some of which optionally leverage AI and machine learning in the background.

Test Data Manager offers solutions for two major aspects of test data management. The first of these aspects is data security and regulatory compliance, delivered via automated data profiling, data anonymisation, and compliance validation. The second is test data agility and DataOps acceleration, provided via synthetic data generation, test database orchestration (of either traditional test databases or virtualised database clones), and test data provisioning. Enov8 refers to these two aspects as “Mode-1”

This **Mutable Quadrant** is derived from 13 high level metrics, the more the image covers a section the better. **Execution** metrics relate to the company, **Technology** to the product, **Creativity** to both technical and business innovation and **Scale** covers the potential business and market impact.



and “Mode-2”, respectively, and they can be implemented either individually or (more likely) simultaneously (see **Figure 2**).

In particular, the database virtualisation capabilities offered as part of this process are derived from integration with Enov8’s VirtualizeMe (vME) product, which allows you to rapidly deliver extremely lightweight database clones to your development and testing teams for test data management (by creating virtualised test data sets) and other purposes. This capability is sufficiently advanced that it is worth covering separately, and we have done so [here](#). Test Data Manager functionality can also be delivered through an API, perhaps integrated into an existing DevOps process or CI/CD toolchain, as shown in **Figure 2**. It can operate across domains, including cloud and on-premises environments, in a federated (and load balancing) fashion, via a network of “worker bees” (agents) that report to the core platform. This provides a significant degree of parallelisation and scalability. In terms of data sources, Test Data Manager supports a variety of the most popular relational and NoSQL databases, various kinds

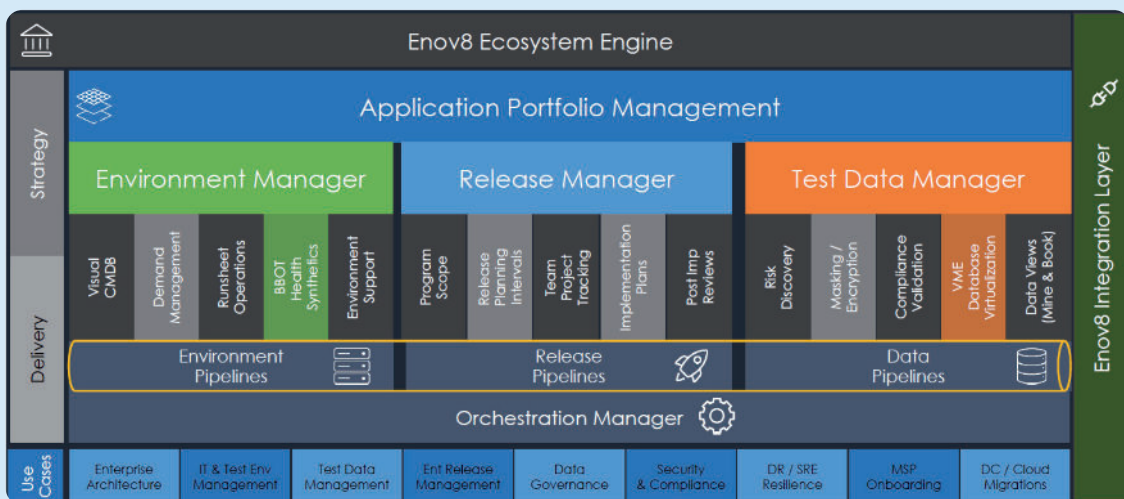


Figure 1 – Enov8 product architecture

of text files (including CSV, JSON, XML, and fixed width), and a smattering of other formats, such as Avro, ORC, and Parquet.

What does it do?

Data profiling in Test Data Manager (for finding your sensitive data, among other things) consists of creating and running data profiling configurations that specify the selection of data fields, and corresponding data patterns, you want to identify. You can flag fields as sensitive within these configurations, and even specify what kind of sensitivity (such as primary PII, secondary PII, or PHI) each field falls under. In addition, various preset profiling patterns and configurations are available, catering to common, international use cases. Data fields are identified using column name and/or pattern matching, and the results of data profiling are presented within profiling reports available for each profiling job individually and for all of them collectively (see **Figure 3** for an example of the latter).

Profiling configurations can also contain recommended data masking (or other anonymisation) methods for each (sensitive) field. During the anonymisation process, you can either rely on these preconfigured methods or choose your own. Enov8 generally recommends using a hashing algorithm to replace sensitive values with random ones drawn from a pre-generated many-to-one lookup table. This method is flexible, non-reversible, maintains referential integrity and can produce production-like data. A selection of look-up tables for common fields are provided out of the box to facilitate this. The product also uses essentially the same method to create synthetic data sets, except that instead of anonymising a series of existing fields it simply creates a set number of them.

Other anonymisation methods are also available, and include encryption, complex ciphering, random or fixed variance (for numerical fields) and fixed value replacement (which is to say, everything is masked to the same thing). You can also implement your own masking algorithms as Python functions, with templates for the above methods provided. Similarly, you can use Python to extend or modify any of the built-in masking



methods. You can also use parameters to perform calculations and leverage their results during the masking process, such as when calculating a checksum or masking a field that needs a concatenation of other masked fields.

Masking can be executed manually or scheduled as part of the company's Orchestration Manager (and in fact can be positioned as part of an automated pipeline, along with profiling and various other tasks, using the same). The masking process can be paused and resumed intelligently, so that existing progress is not lost, and ultimately produces a report detailing which masking jobs succeeded and which (if any) failed. In the latter case, error details are provided, and you can choose to rerun only the failed jobs once those errors have been fixed. Delta analysis is available for change management purposes where an update to the database schema may have introduced new fields, and masking validation available as part of the report will tell you both how production-like each field's masked data is (using distribution analysis) and what percentage of values have been masked within each field.

Finally, for Test Data Mining, Test Data Manager offers a data library that allows you to both view and reserve data. Various data dashboards are also available (an example of which is shown in **Figure 4**) that can be customised both for your organisation as a whole and by each user individually.

Why should you care?

Test Data Manager provides considerable breadth of test data management functionality. To wit, it offers a complete test data solution, providing sensitive data discovery (via data profiling), data anonymisation (via masking), and database virtualisation (when used alongside vME, which is highly recommended). It also benefits from being part of a product suite alongside Enov8's other offerings, such as Environment Manager, Release Manager, and Orchestration Manager. In short, it allows you to protect your sensitive test data – and comply with any relevant regulations – while accelerating your test data processes and

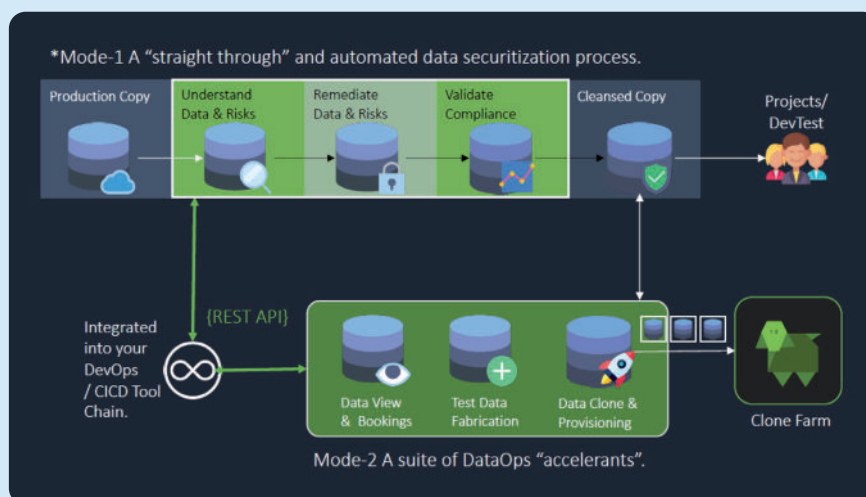


Figure 2 – Mode-1 and Mode-2 Test Data Manager operations

ensuring they do not become a bottleneck for your testing as a whole. Synthetic data generation is also available, although in this case we would say it is best considered as an ancillary capability for complimenting your other test data with new data that may not exist in production.

Using distribution analysis to determine how closely your masked data resembles your production data is a rarely seen, but welcome, feature. It allows you to create test data sets that are highly secure but retain the statistical properties of the original data, enabling you to test as if you were using the original data but without the accompanying privacy concerns. The selection of masking and matching methods provided by Test Data Manager are relatively minimal, but its primary method of masking – using many-to-one lookup tables – compensates for this through sheer flexibility, and you can add to the available masking methods yourself if required.

It is also worth noting that Test Data Manager is quite easy to use through its web interface, to the extent that, at least in

principle, it can be leveraged by both technical and non-technical users. At minimum, this makes your testers' lives easier. More significantly, this can allow parts of your organisation outside of your dedicated testers to assume some responsibility for test data, helping you to spread the load and utilise tribal knowledge. In theory, you could even federate out your test data management entirely, with your dedicated testers primarily acting in a test data governance capacity.

The bottom line

Enov8 Test Data Manager is a robust solution for data profiling and data masking, with additional capabilities for test data provisioning and synthetic data generation (among other things). When combined with vME, the database virtualisation offering also from Enov8, it offers a holistic and comprehensive solution for test data management, ensuring test data security, compliance, and efficiency, especially within DevOps environments.

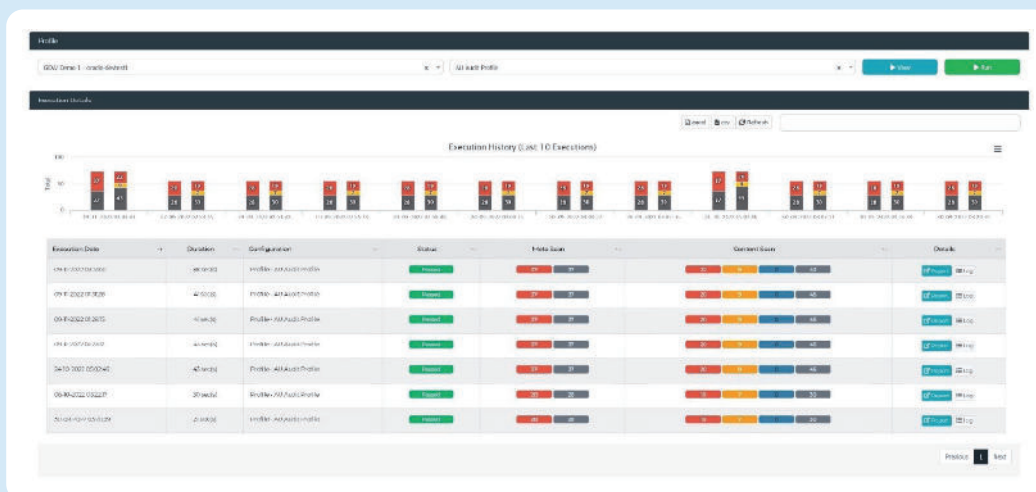


Figure 3 – Test Data Manager profiling execution report

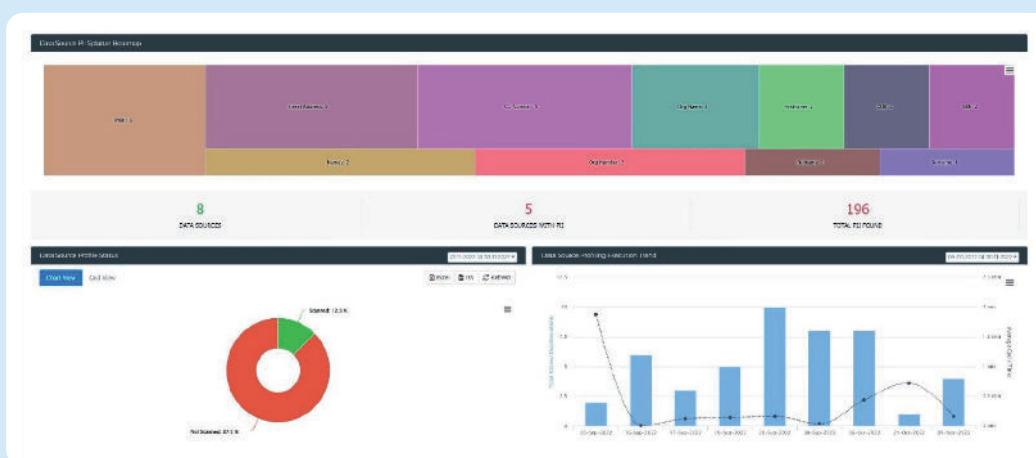


Figure 4 – Test Data Manager executive insights dashboard