

## Efficient generation of customized, complex and consistent synthetic data

It is often not possible or permissible to use production data, anonymized or otherwise, for test purposes and doing so can lead to unwanted side effects.

Synthetic data is the alternative. Until now, however, creating synthetic data has been too complex, of insufficient quality and, often, simply not practicable.

iSynth enables you to create synthetic data efficiently and cost-effectively.

### Key Features

**Support of all test levels:** iSynth can supply synthetic data for all test levels. From unit tests to fully integrated system tests on large system landscapes. iSynth's synthetic data can be used as primary data or to complement existing data and is suitable for generating both concise and large test data sets.

**Flexible and extensible:** iSynth is easy to adapt and extend according to specific requirements in terms of projects and customer environments. iSynth users can implement specific functions themselves as extensions in Python.

**Teamwork:** iSynth supports collaboration within the team during the data definition, generation and usage phases.

**Customized test data:** iSynth customizes test data directly for specific test cases, making time-consuming searches for suitable test data a thing of the past. The data variance required for the test cases can be managed using rules or simple Excel sheets.

**Manage data consumption:** iSynth's integrated Data Checkout Tool can be used to mark individual data records as depleted and thus manage data consumption.

**Metadata:** iSynth allows you to define any metadata you wish to complement the generated data (e.g. you can add a test case ID).

**Reusability:** One of iSynth's basic concepts is the definition of reusable components, both at the level of the model elements (building blocks) and copy templates, as well as deployment functions.

### Challenges

**Regulatory requirements:** Requirements for handling private, health and bank data are becoming increasingly restrictive. In the age of big data analytics, solutions based on data anonymization will most likely fail to meet the requirements in terms of data and confidentiality protection.

**Agile projects and DevOps testing:** In addition to appropriate development methods and tools, frameworks, service simulators and test utilities, it is essential to have the right amount of consistent, high-quality test data in order to perform meaningful tests.

**External sourcing:** Software components are often developed by external partners. Whether these partners are active onshore, nearshore or offshore is irrelevant. The development partner needs data that accurately represents the client's sensitive data and to be able to use it without security concerns.

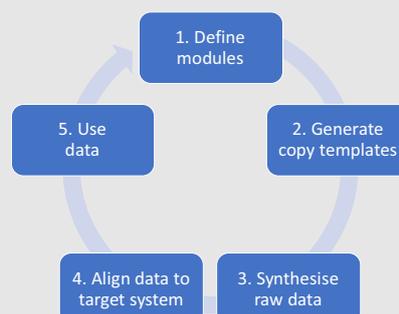
**Complexity of data and applications:** IT systems are being used to process increasingly complex and cross-linked data. At the same time, IT landscapes are being remodeled from vast, monolithic applications to fine-grained components (keyword "Micro Services"). As a result, the data is better distributed from a business point of view, but needs to be consistent across all components concerned for testing purposes.

**Efficiency in testing:** Efficiency is highly dependent on the test data, for both manual and automated testing. The search for suitable test data, the handling of varying data and the irreversible consumption of test data produce inefficiencies in the testing process.

### Solution overview

iSynth can be used create complex, consistent, high-quality synthetic data quickly and cost-effectively.

iSynth supports the entire life cycle of synthetic data and its object-oriented, model-based approach enables the required synthetic data to be generated in multiple steps in a structured yet flexible process.



The module types to be synthesized are defined from a business perspective using iSynth-specific, object-oriented notation during Step 1. Step 2 involves generating the copy templates using the module types and defining the level of variance at which data is to be synthesized during step 3. Technical attributes and target system-specific formats are only taken into account in step 4.

## Technology and architecture

**Open Source:** iSynth is based on Open Source technologies and frameworks, such as Python, Django, Vue.js.

**GUI:** iSynth is operated via a **Web GUI**. Alternatively, certain iSynth files can also be modified using text editors or Python IDEs.

**Database Management Systems:** iSynth requires a DBMS supported by Django to create the synthetic data. **PostgreSQL, Oracle** and **SQLite** have proven suitable for this task. Any DBMS can be used as a target database system.

**Interfaces to target systems:** The synthetic data can be transferred flexibly to the target system via any interface. Files (CSV, XML, JSON etc.), SQL Loader Scripts, XML messages via JMS, Web service calls etc. are commonly used for this. Target systems can also actively retrieve the data from the iSynth database via SQL queries. iSynth also supports bidirectional interfaces to target systems. Bidirectional interfaces are required when the target system assigns keys that are required later on in the iSynth process.

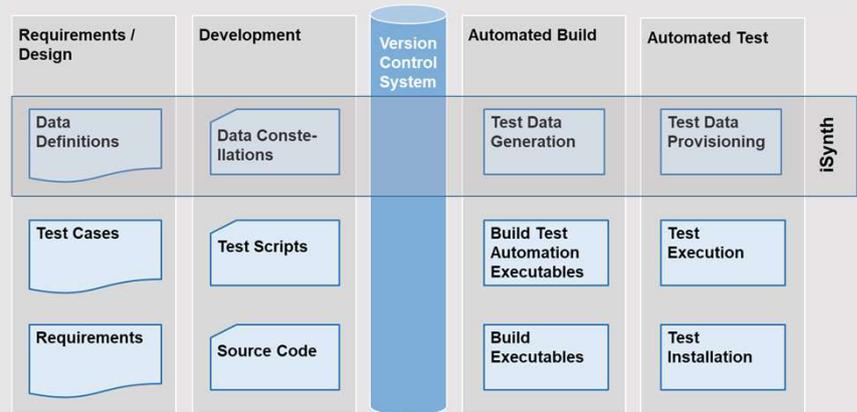
**Integration of test automation and service virtualization:** This is handled in the same way as integrating any other target system. LeanFT and LISA were successfully interfaced to iSynth via JSON.

**Runtime environment:** iSynth runs on Windows, macOS, Linux and any operating system that supports Python 3.6. iSynth is also available as a Docker Image and is particularly easy to use in this way.

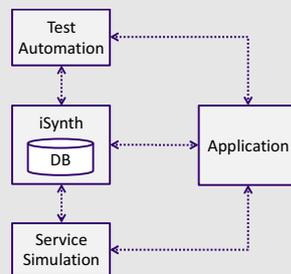
**Version Control System:** The project-specific artefacts are available as files and are managed via a standard version control system such as GIT. The data in the iSynth databases does not have to be versioned, since it can be created again at any time due to the other versioned artefacts.

## Integration in DevOps

It is also essential that the right data is available in order to perform meaningful, automated tests in the DevOps environment. Based on textual data definitions that are under version control, iSynth can automatically generate, distribute and enable the use of the required data in the DevOps cycle.



iSynth delivers consistent data to the tested application, as well as to the test automation tools and service simulation/virtualization.



Ideally, iSynth is used to define data objects that describe test cases, test steps, and the data required to perform the test steps. This way, the tests to be performed can be controlled using the defined and synthesized data in iSynth. This will then make the data in the test automation scripts obsolete, thus simplifying the creation and maintenance of these scripts and saving costs.

## Benefits

### Efficiency in testing

- Aligned design of test cases and test data
- No more searching for test data
- No coordination problems with test data
- Optimized volume of test data
- Fast definition of test data via reusable modules

### Effectiveness in testing

- Optimization of test cases through defined variance
- Generation of data that does not (yet) exist in production

### Automation in Test and DevOps processes

- Provision of test automation with test and control data, so that the sequence of automated tests is managed in a data-driven way
- Automatic synthesization of data within DevOps processes

### No privacy issues

- Synthetic data is inherently secure

### Data for further usage purposes

- Training, demo system, honeypot

## Contact

Have we piqued your interest? We would be pleased to advise you personally about the various possibilities of iSynth and discuss the application scenarios in your environment.

Your iSynth contacts:

Josef Bösze, Partner, [josef.boesze@itopia.ch](mailto:josef.boesze@itopia.ch)

Ralph Schibli, Managing Partner, [ralph.schibli@itopia.ch](mailto:ralph.schibli@itopia.ch)



corporate information technology

itopia ag - corporate information technology  
Technoparkstrasse 1, CH-8005 Zurich  
[www.itopia.ch](http://www.itopia.ch)

© itopia ag, September 2018