

## Effiziente Erzeugung von massgeschneiderten, komplexen und konsistenten synthetischen Daten

Der Einsatz von Produktionsdaten oder anonymisierten Produktionsdaten zu Testzwecken ist oft nicht möglich, nicht erlaubt oder führt zu unerwünschten Nebenwirkungen.

Synthetische Daten sind die Alternative. Allerdings war bis jetzt die Erstellung von synthetischen Daten zu aufwändig, qualitativ ungenügend und oft einfach nicht praktikabel.

iSynth ermöglicht die effiziente und kostengünstige Erstellung von synthetischen Daten.

### Key Features

**Unterstützung aller Teststufen:** iSynth kann alle Teststufen mit synthetischen Daten beliefern. Vom Unit Test bis hin zum voll integrierten Systemtest eines grossen Systemverbands. Die synthetischen Daten von iSynth können als Primärdaten oder als Ergänzung von bereits bestehenden Daten genutzt werden. iSynth eignet sich sowohl für die Generierung kompakter wie auch grosser Testdatensets.

**Flexibel und erweiterbar:** iSynth kann leicht auf die spezifischen Bedürfnisse von Projekten und Kundenumfeldern angepasst und erweitert werden. iSynth-Nutzer können spezifische Funktionen als Erweiterungen in Python realisieren.

**Teamarbeit:** iSynth unterstützt die Arbeit im Team, sowohl bei der Definition, der Generierung wie auch der Nutzung der Daten.

**Massgeschneiderte Testdaten:** Mit iSynth werden Testdaten für Testfälle direkt massgeschneidert. Die aufwändige Suche nach den passenden Testdaten in einem Datenbestand entfällt. Die für die Testfälle benötigte Datenvarianz kann über Regelwerke oder mittels simplen Excel-Tabellen gesteuert werden.

**Datenverbrauch managen:** Über das in iSynth integrierte Data Checkout Tool können einzelne Datensätze als verbraucht markiert und der Datenverbrauch verwaltet werden.

**Metadaten:** iSynth erlaubt die Definition beliebiger Metadaten als Ergänzung der eigentlichen Daten. Z.B. kann die Testfallidentifikation hinterlegt werden.

**Wiederverwendbarkeit:** Ein Grundkonzept von iSynth ist die Definition von wiederverwendbaren Elementen, sowohl auf Ebene der Modellelemente (Bausteine), der Kopiervorlagen wie auch der Deployment Funktionen.

### Herausforderungen

**Regulatorische Vorgaben:** Die Vorgaben zum Umgang mit Personen-, Gesundheits- und Bankdaten werden immer restriktiver. Im Zeitalter von Big Data Analytics erfüllen Lösungen auf der Basis von Datenanonymisierung die Anforderungen im Sinne des Daten- und Persönlichkeitsschutzes nur ungenügend.

**Agile Projekte und DevOps Testing:** Für aussagekräftige Tests braucht es neben entsprechenden Entwicklungsmethoden und –werkzeugen, Toolchains, Service-Simulatoren, Testwerkzeugen vor allem auch konsistente, qualitativ gute Daten in der richtigen Menge.

**Externes Sourcing:** Oft werden Software-Komponenten durch externe Partner entwickelt. Ob diese Partner im Inland, nearshore oder offshore tätig sind, spielt keine Rolle. Der Entwicklungspartner benötigt Daten, welche die Kundendaten des Auftraggebers akkurat repräsentieren und ohne Sicherheitsbedenken verwendet werden können.

**Komplexität von Daten und Applikationen:** In den IT Systemen werden immer komplexere, stärker vernetzte Daten bearbeitet. Gleichzeitig werden die IT Landschaften von grossen monolithischen Applikationen zu feingranularen Komponenten (Stichwort „Micro Services“) umgebaut. Als Konsequenz sind aus fachlicher Sicht die Daten stärker verteilt, müssen aber für Tests über alle betroffenen Komponenten konsistent sein.

**Effizienz im Testing:** Diese ist sowohl beim manuellen wie auch automatisierten Testing sehr stark von den Testdaten abhängig. Die Suche nach geeigneten Testdaten, der Umgang mit wechselnden Daten sowie der irreversible Verbrauch von Testdaten führen zu Ineffizienzen im Testing.

### Lösungsüberblick

Mit iSynth können komplexe, konsistente, qualitativ hochwertige synthetische Daten rasch und kostengünstig erstellt werden.

Mit iSynth wird der gesamte Lebenszyklus von synthetischen Daten unterstützt. Der objektorientierte, modellbasierte Ansatz von iSynth ermöglicht es, in einem strukturierten aber flexiblen Prozess, in mehreren Schritten die benötigten synthetischen Daten zu erzeugen.



Die zu synthetisierenden Bausteintypen werden aus einer möglichst fachlichen Perspektive in einer iSynth spezifischen, objektorientierten Notation im Schritt 1 definiert. Die in Schritt 2 zu erstellenden Kopiervorlagen nutzen die Bausteintypen und definieren, mit welcher Varianz Daten im Schritt 3 synthetisiert werden sollen. Technische Attribute und Zielsystem spezifische Formate werden erst im Schritt 4 berücksichtigt.

# Technologie und Architektur

**Open Source:** iSynth basiert auf Open Source Technologien und Frameworks wie Python, Django, Vue.js.

**GUI:** iSynth wird über ein **Web-GUI** bedient. Alternativ können gewisse iSynth-Dateien auch mit Text-Editoren oder Python IDEs bearbeitet werden.

**Datenbanksysteme:** iSynth benötigt für die Erstellung der synthetischen Daten eine von Django unterstützte Datenbank. **PostgreSQL, Oracle** und **SQLite** haben sich bewährt. Als Zieldatenbanksysteme können beliebige DBMS erschlossen werden.

**Schnittstellen zu Zielsystemen:** die synthetischen Daten können sehr flexibel über beliebige Schnittstellen zu den Zielsystemen transportiert werden. Häufig verwendet werden Files (CSV, XML, JSON u.a.m.), SQL Loader Scripts, XML Meldungen über JMS, Webservice Calls u.a.m. Zielsysteme können sich die Daten aber auch aktiv aus der iSynth-Datenbank mit SQL Zugriffen holen. iSynth unterstützt auch bidirektionale Schnittstellen zu Zielsystemen. Bidirektionale Schnittstellen werden verwendet, wenn das Zielsystem Schlüssel vergibt, die im weiteren Prozess von iSynth benötigt werden.

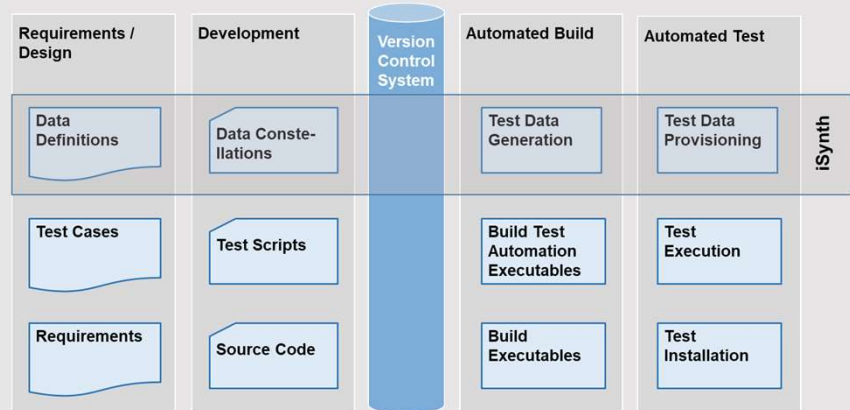
**Anbindung von Testautomatisierung und Service-Virtualisierung:** Diese erfolgt analog zur Anbindung der Zielsysteme. LeanFT und LISA wurden über JSON erfolgreich an iSynth angebunden.

**Runtime Umgebung:** iSynth läuft auf **Windows, macOS, Linux** und auf jedem Betriebssystem, welches Python 3.6 unterstützt. iSynth liegt auch als **Docker** Image vor und kann so besonders einfach genutzt werden.

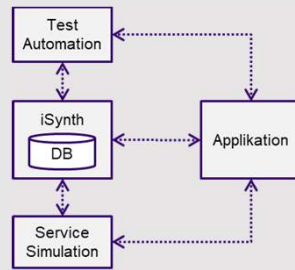
**Version Control System:** Die projektspezifischen Artefakte liegen als Dateien vor und unterliegen der Kontrolle eines Standard Version Control Systems wie z.B. GIT. Die Daten in den iSynth-Datenbanken müssen nicht versioniert werden, da sie aufgrund der anderen, versionierten Artefakte jederzeit wieder erstellt werden können.

# Integration in DevOps

Für aussagekräftige, automatisierte Tests im DevOps Umfeld müssen auch die richtigen Daten bereitgestellt werden. Basierend auf textuellen Datendefinitionen, die unter Versionskontrolle sind, können mit iSynth die benötigten Daten im DevOps Zyklus automatisch generiert, verteilt und genutzt werden.



Durch iSynth werden sowohl die zu testende Applikation als auch die Testautomatisierungswerkzeuge und die Service-Simulation/Virtualisierung mit konsistenten Daten beliefert.



Idealerweise werden in iSynth Datenobjekte definiert, welche Testfälle, Testschritte und die für die Ausführung der Testschritte benötigten Daten beschreiben. So kann die Steuerung der auszuführenden Tests vollständig durch die in iSynth definierten und synthetisierten Daten erfolgen. Die Testautomatisierungsscripts enthalten so keine Daten mehr. So wird die Erstellung und die Wartung dieser Scripts einfacher und günstiger.

## Ihr Nutzen

### Effizienz im Testing

- Abgestimmtes Design von Testcases und Testdaten
- Keine Suche nach Testdaten
- Keine Abstimmungsprobleme bezüglich Testdaten
- Optimale Menge von Testdaten
- Rasche Definition von Testdaten über wiederverwendbare Bausteine

### Effektivität im Testing

- Optimierung der Testfälle durch definierte Varianz
- Erzeugung von Daten, die es in der Produktion (noch) nicht gibt

### Automatisierung in Test- und DevOps Prozessen

- Versorgung der Testautomatisierung mit Test- und Steuerungsdaten, so dass die Abfolge der automatisierten Tests datengesteuert erfolgt
- Automatische Synthetisierung von Daten im Rahmen von DevOps Prozessen

### Keine Datenschutzprobleme

- Synthetische Daten sind inhärent sicher

### Daten für weitere Einsatzzwecke

- Schulung, Demosystem, Honeypot

## Kontakt

Haben wir Ihr Interesse geweckt? Gerne informieren wir Sie persönlich über die vielfältigen Möglichkeiten von iSynth und diskutieren die Einsatzszenarien in Ihrem Umfeld.

Ihre Ansprechpartner für iSynth:

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



corporate information technology

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

© itopia ag, März 2018