

## In dieser Ausgabe:

- ⇒ Warum einen Testdatengenerator verwenden?
- ⇒ Testdatengeneratoren für funktionale und nicht-funktionale Tests
- ⇒ Synthetische vs. anonymisierte Testdaten
- ⇒ Links, Literatur und Zitate

## Warum einen Testdatengenerator verwenden?

In der Entwicklung von komplexen Softwaresystemen ist die Testautomatisierung eine unabdingbare Tätigkeit, um die Qualität von Software zu steigern bzw. zu erhalten. Ein Großteil der heute verwendeten Software arbeitet im Hintergrund mit Datenbanken oder anderen Datenformaten, wie beispielsweise einer XML-Struktur. Dies macht es schwierig, eine angemessene Testabdeckung zu erreichen. Die geringere Testabdeckung vergrößert jedoch die Gefahr, kritische Fehler zu übersehen.

Eine vollständige Abdeckung ist in der Praxis kaum erreichbar, weil die Anzahl der möglichen Testfälle schnell anwächst (kombinatorische Explosion). Die Anwendung von Äquivalenzklassen und Grenzwerten kann die Anzahl der benötigten Tests deutlich reduzieren. Immer wieder steht der Tester vor der Situation, die Äquivalenzklassen-Methode nicht oder nur teilweise anwenden zu können, da keine Klassen und Grenzwerte existieren. Beispiele wären Konfigurationsmöglichkeiten wie Telefonnummern, Adressen, Länder usw. auf einer Benutzeroberfläche. Ein anderes Beispiel ist, wenn mehrere System-Varianten wie z. B. 5 Browser, 5 Zielsysteme (Windows, Linux, Mac sowie zwei Mobile Betriebssysteme), 2 Datenbanken (Oracle und DB2), 8 Sprachen und 2 Server in Kombinationen getestet werden. Wenn alle Kombinationen gültig sind, werden  $5 \times 5 \times 2 \times 8 \times 2 = 800$  Kombinationen für eine vollständige Überdeckung benötigt. Das sind nicht 800 verschiedene logische Testfälle, sondern 800 Testdurchläufe mit unterschiedlichen Testdatenkombinationen, d. h. das 800-fache Ausführen der Systemtestfälle.

Testverfahren wie Entscheidungstabellen-Reduktion und Klassifikationsbäume verringern hier die Anzahl der Testfälle und Testdurchläufe. Was passiert aber, wenn bei der Anwendung aller Testverfahren und Reduktionen die Testfallanzahl und die Varianten immer noch sehr umfangreich sind? Spätestens dann sollte der Testmanager an eine automatische Datengenerierung mit einem Testdatengenerator denken.

Da das Thema Testdatengeneratoren sehr umfangreich ist, konzentriert sich dieser Knowledge-Letter im Folgenden nur auf ausgewählte Ansätze der Testdatengenerierung.

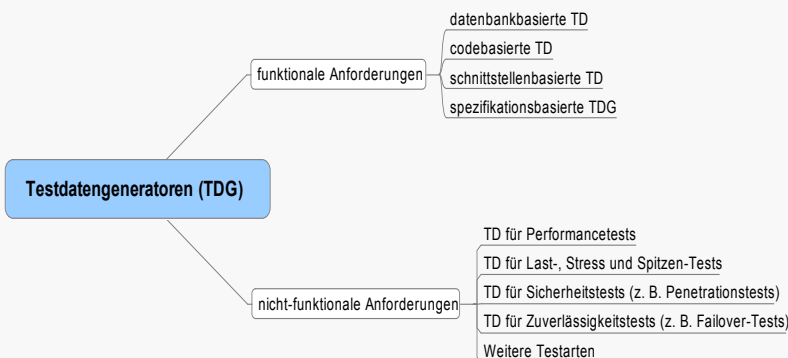


Abbildung 1 – Klassifikation der Testdatengeneratoren [Quelle: SWQL]



**Qualität vor Quantität**

oder

**Qualität mit Quantität?**

Was nützen uns 1000 Testfälle, wenn Sie alle dieselbe Aussagekraft haben und keinen inhaltlichen Mehrwert bringen?

Insofern ist ja die Aussage „Qualität vor Quantität“ besonders in der professionellen Qualitätssicherung zu befürworten.

Aber Ausnahmen bestätigen eben auch die Regel „Qualität mit Quantität“:

Denn wohlüberlegte Kombinationen von Testdaten haben in der Regel ja einen Sinn und decken bestimmte Varianten eines Testfalls ab, die in der Praxis vorkommen.

Problematisch ist meist nur die Wirtschaftlichkeit der Testdurchführung von sehr vielen Varianten. Wenn es nichts kostet, würde man ja gerne alle Möglichkeiten durchspielen.

Mit Testdatengeneratoren hat man in manchen Einsatzbereichen die Möglichkeit, auf einfache Art eine Vielzahl von Varianten zu generieren. Durch unterschiedliche Arten von Testdatengeneratoren wie spezifikations-, schnittstellen-, datenbank- oder codebasiert können praktisch alle Testebenen durch diese Tools sinnvoll unterstützt werden.

Durch die automatisierte Generierung von vielen Testfall-Varianten bleiben die Kosten meist relativ gering und es hilft dies dem Tester, eine deutlich bessere Testabdeckung zu erreichen und verschiedenste Situationen in den Testabläufen einfach durchzutesten.

**Dipl.-Ing. Johannes Bergsmann**

Staatl. befugter und beedeter Ingenieurkonsultent für Informatik

Der Quality-Knowledgeletter ist ein periodisches Informationsmedium von Software Quality Lab und dessen Partnern mit den Schwerpunkten IT-Qualitätsmanagement, Projekt- und Prozess-Management.

Inhalt: fachliche Beiträge und Schwerpunktthemen, Vorstellung neuer Produkte und Leistungen, neue wissenschaftliche Erkenntnisse und andere Fachbeiträge aus unseren Themenbereichen.

Aktuelle Fach- und Forschungsbeiträge sind willkommen. Einsendungen an [info@software-quality-lab.at](mailto:info@software-quality-lab.at).

Weitere Infos zu diesem und anderen Themen finden Sie auf <http://www.software-quality-lab.at>.

# Testautomatisierung mit Testdatengeneratoren

DI Daniel Roschanek - Software Quality Lab

Zur umfangreichen Testabdeckung werden neben Testautomatisierungswerkzeugen sogenannte Testdatengeneratoren verwendet. Die Testdatengeneratoren für funktionale Tests basieren entweder auf einer Datenbank, dem Quellcode, einer Schnittstelle oder auf einer Spezifikation. Bei nicht-funktionalen Tests werden die generierten Daten für Performance-, Last-, Stresstests usw. eingesetzt. Dieser Artikel gibt einen kurzen Überblick über die Arten von Testdatengeneratoren.

## Begriffsdefinitionen

### Testdaten

Daten, die vor der Ausführung eines Tests existieren und die die Ausführung der Komponente oder des Systems im Test beeinflussen.

### Anonymisierte Testdaten

Anonymisierten Daten sind personenbezogene Daten, aus denen kein unmittelbarer oder mittelbarer Rückschluss auf die mit den Daten verbundene Person möglich ist.

### Synthetische Testdaten

Ein synthetischer Datensatz ist ein künstlich erstellter Datensatz, der nicht auf realen Daten basiert.

### Testdatengenerator

Ein Testunterstützungswerkzeug, mit dem die für einen Test benötigten Daten generiert, bereitgestellt, verändert oder aus einer Wissensdatenbank selektiert werden können.

## Klassifikation der Testdatengeneratoren

In allen Bereichen des automatisierten datengetriebenen Testens können Testdatengeneratoren eingebunden werden. Wie in Abbildung 1 zu sehen ist, werden die Testdatengeneratoren nach der Verwendungsart der Massendaten für funktionale und nicht-funktionale Anforderungen klassifiziert.

Für **Funktionstests** werden Testdaten aus unterschiedlichen Quellen (Datenbank, Software, Schnittstelle, Spezifikation) generiert oder abgeleitet. Das Vorgehen zur Testdatengenerierung auf der Basis dieser vier Ansichten wird im nächsten Kapitel genauer beschrieben.

Für **nicht-funktionale Tests** werden Testdaten für Performance-, Last-, Stresstests usw. zur Verfügung gestellt. Diese Testarten können effizient meist nur mit Hilfe eines Werkzeuges mit bereitgestellten Massendaten durchgeführt werden.

## Testen von funktionalen Anforderungen

Generatoren für die Testfallspezifikation von funktionalen Anforderungen können in zwei Kategorien eingeteilt werden: **Testfallgeneratoren** und **Testdatengeneratoren**.

Testfallgeneratoren erstellen automatisch logische Testfälle und Testdatengeneratoren generieren konkrete Daten aus den logischen Testfällen.

Eine der größten Herausforderungen bei der Testdatengenerierung für Funktionstests ist es, synthetisch realistische Daten passend in Struktur, Kontext und Quantität generieren zu können. Beispielsweise sollte ein Name als 'Hans Müller' nicht nur als eine Zeichenfolge 'R2 D2' ausgeprägt sein. Je

realistischer die Datenwerte sind, umso effizienter können diese für Testzwecke eingesetzt und resultierende Fehlermaskierungen vermieden werden [Scholze 2005].

## Datenbankbasierte Testdatengeneratoren

Mit dieser Klasse von Testdatengeneratoren werden Testdaten aufgrund eines vorhandenen Datenbankschemata oder Datenbankbeständen erzeugt.

Die komplexen Datenbankstrukturen werden aus der Datenbank extrahiert (gute Tools stellen dazu einen Datenbank-Import-/Export-Assistenten zur Verfügung).

Der Testdatengenerator erstellt aus dem eingelesebenen Aufbau und den Eigenschaften der Tabellen (z.B. Indices, Datentypen, Längen der Felder, etc.) und deren Beziehungen untereinander möglichst passende (synthetische) Testdaten, welche dann zur Ausführung von automatisierten Tests verwendet werden können.

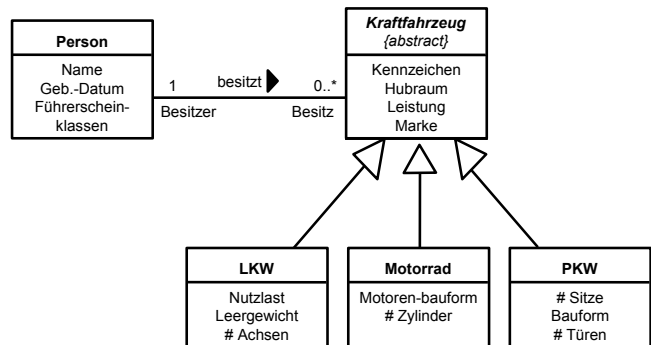


Abbildung 2 - Datenmodell als Basis für Testdatengeneratoren [Grafik: SWQL]

Problematisch bei der datenbankbasierten Testdatengenerierung ist, dass die aus den Tabellen-Strukturen generierten synthetischen Testdaten nicht unbedingt auch passende Testdaten für die zu testende Applikation sein müssen (z.B. kann ein Textfeld in der Datenbank die Länge 255 haben und in der Benutzeroberfläche der Applikation jedoch nur 40 Zeichen lang sein. Weiters müssen auch entsprechende Schlüssel und Datensätze für die Abbildung von Beziehungen zwischen Instanzen (Datensätzen) der Entitäten eingefügt werden. Des Weiteren müssen die Konsistenz- und Integritätsbedingungen der Testdaten sichergestellt werden, was nicht immer alleine aus der Datenbankstruktur ableitbar ist.

Datenbankbasierte Testdatengeneratoren können eine gute Ergänzung zu anderen Testdatengeneratoren sein bzw. sind sehr hilfreich, wenn es darum geht, Veränderungen an Datenbanken durch Testfälle zu erkennen. Als primäre Basis für Testfallgenerierungen sind sie jedoch nicht ausreichend.

## Vollständiger Knowledge Letter Zugang

Wir freuen uns, dass Sie an diesem Thema Interesse haben und den Knowledge Letter von Software Quality Lab bis hierher gelesen haben.



**Dieser Knowledge Letter ist eine Vorschau (gekürzte Version des gesamten Artikels).**

Wenn Sie den ungekürzten Knowledge Letter lesen möchten, registrieren Sie sich bitte unter <http://www.software-quality-lab.com/download/knowledge-letter/anfrage-knowledge-letter/>

Sie erhalten nach der Registrierung vollen Zugang zu allen bisherigen Knowledge Letters von Software Quality Lab und erhalten automatisch künftige Knowledge Letter per E-Mail.

### Software Quality Days — Die größte Konferenz zum Thema „Software Qualität“ in Europa!



Besuchen Sie die Top-Konferenz mit allen Infos rund um Software Qualität.

Beste Qualität der Vorträge und Tutorials sowie eine Mischung aus praktischen und wissenschaftlichen Beiträgen machen die Software Quality Days zum Top-Event.

In den 3 praktischen Tracks werden anwendungsorientierte Vorträge präsentiert. Der wissenschaftliche Track zeigt Beiträge mit hohem Innovationsgrad und praktischer Anwendbarkeit, basierend auf Forschungsergebnissen. Im Solution Provider Forum präsentieren Aussteller ihre neuesten Tools mit Praxis-Beispielen.

Nähere Infos unter

[www.software-quality-days.com](http://www.software-quality-days.com)

