

Business Technology

Architektur & Management Magazin

Expertenwissen für IT-Architekten, Projektleiter und Berater



Peter Schaar:
„Datenschutz setzt
die Spielregeln.“

SICHERHEIT

**IT-Sicherheit
integral betrachtet**

Data Loss Protection

**Enterprise Wide
Integration Security**

Alles sicher oder was?

Die BSI-Sicherheitsanforderungen an
Cloud-Anbieter

Rechte und Rollen

Zentrales Rollen- und Rechtecockpit
für SOA-Applikationen

Business Continuity Management

Sicherung kritischer
Wertschöpfungsprozesse

Schatzkammer Datenbank

Sicherheitsstrategien für Datenbanken

Testdaten – Generierung oder Verfremdung?

Herausforderungen beim Umgang mit Testdaten

Ein Softwaretest ist nur so gut wie die Daten, mit dem die Systeme getestet werden. Die gängige Praxis, Abzüge von Produktionsdaten zu verwenden, kann datenschutzrechtliche Folgen bergen, die vielen Unternehmen und Dienstleistern anscheinend nicht hinreichend bekannt sind. Auch haben Unternehmen zwar oft Datenschutz-, IT-Compliance- und IT-Governance-Regeln, doch wie wendet man diese im Softwaretest an?

AUTOREN: MATTHIAS RASKING UND STEPHAN OSWALD

Der Umgang mit Testdaten ist weiterhin eines der größten Probleme im Softwaretest. Sei es die schwierige Identifikation von Testdaten für Testszenarien, die Herausforderung, Testdaten über einen längeren Testzeitraum hinweg nachzuhalten oder die mangelnde Verfügbarkeit von Testdaten aus Schnittstellensystemen – ca. 30 Prozent aller abgewiesenen Fehler in Projekten mit einem niedrigen Reifegrad der IT-Prozesse sind auf mangelhafte Testdaten zurückzuführen. Mit definierten Prozessen für Testdatenverwendung lässt sich diese Fehlerquelle auf unter 5 Prozent reduzieren. Diese Verbesserung verlangt allerdings ein umfassendes Verständnis der Anforderungen an Testdaten sowie einen projektweiten Fokus auf die Erstellung und Pflege von Testdaten.

Die Bereitstellung von Testdaten wird in Releases zu oft auf wenige Schritte beschränkt, anstatt den gesamten Lebenszyklus von Testdaten vollständig abzubilden (Abb. 1).

Hierbei wird davon ausgegangen, dass Testfälle grundsätzlich Testdaten aus dem Produktionsbetrieb be-

nötigen (entweder aus einem Datenabzug oder migrierte Daten bei neuen Anwendungen). Außerdem setzt dieser Prozess zwar eine umfangreiche Kenntnis der Tester in Bezug auf die relevanten Daten für die Testfälle voraus, vernachlässigt aber wesentliche Punkte in Bezug auf Nachvollziehbarkeit, Wiederverwendbarkeit und letztlich auch Prozesseffizienz. Oft wird eine gewisse „Realitätsnähe“ der Daten für Testzwecke gefordert, was aber meistens bedeutet, dass die benötigten Testszenarien nicht ausreichend detailliert wurden, um auf andere Bereitstellungsmethoden zurückzugreifen. Der pragmatische Ansatz, Testdaten während des Tests herauszusuchen, bedingt sehr erfahrene Tester und erschwert die Planbarkeit der Tests. Der Aufwand für einen Produktionsabzug ist meistens immens (es gibt Ausnahmen, bei denen mit viel Aufwand auch eine wöchentliche Bereitstellung ermöglicht wird). Dazu kommen Ladezeiten, während denen die Testumgebungen nicht zur Verfügung stehen. Daher aktualisieren viele Firmen ihren Testdatenbestand nur ein bis zwei mal im Jahr, was wiederum zu Proble-

men mit der Aktualität der Testdaten (zum Beispiel durch zwischenzeitlich eingespielte Änderungen) führen kann.

ANSÄTZE ZUR VERBESSERUNG DER SITUATION

Eine Möglichkeit zur Verbesserung besteht in der Nutzung von synthetischen Testdaten, also Datensätzen, die manuell oder mit Werkzeugen generiert wurden. Einige Firmen verwenden bereits im Einsatz befindliche Werkzeuge (die eigentlich zur Testautomatisierung oder für Performance Tests geeignet sind), um über die Oberfläche Testdaten anzulegen. Für einfache Datenkonstellationen und die Vervielfältigung von Daten mag dieser Ansatz zwar zweckdienlich sein, allerdings muss dazu die Anwendung stabil sein, und komplexe Datensätze oder historisierte Daten können nicht erzeugt werden. Hinzu kommt noch ein Thema, das im Zuge der Novellierung des Bundesdatenschutzgesetzes, des Telekommunikationsgesetzes und weiterer Gesetze und Regularien nicht nur in Deutschland mehr und mehr Aufmerksamkeit erhält: Datenschutz und Datensicherheit.

In Hinsicht auf die abzuleitenden datenschutzrechtlichen Hürden, die das Softwaretesten berühren, müssen alle Testprozesse und ausgeführten Testprojekte überprüft werden. Es ist zu prüfen, ob die Dokumentation von Testfällen, die personenbezogene oder sensitive Daten enthalten, aufbewahrt und zurückgeführt werden müssen oder vernichtet werden können. Dazu muss natürlich mit der Revision und Rechtsabteilung erst einmal definiert werden, welche Daten sensitiv sind und wie mit diesen umgegangen werden darf (zeitliche Beschränkung der Tests, Logging und Einschränkung des Zugriffs, Anonymisierung etc). Es muss geprüft werden, ob die Prozesse zur Übergabe von Daten an interne oder externe Testorganisationen noch den gesetzlichen Bestimmungen standhalten. Darüber hinaus ist zu prüfen, wie auf Produktionsdaten verzichtet oder der Prozess so gesteuert werden kann, dass eine Anonymisierung der Daten (sofern möglich und aus einer Kosten-Nutzen-Betrachtung heraus sinnvoll) dort stattfindet, wo Testmanager und externe Dienstleister sowie deren Erfüllungsgehilfen keinen Zugriff haben. Tests mit Daten aus Produktivsystemen sind auf ein Minimum zu reduzieren und es ist ein methodischer und toolgestützter Rahmen aufzubauen, der die Simulation realer Testszenarien mit rein synthetisch erzeugten und/

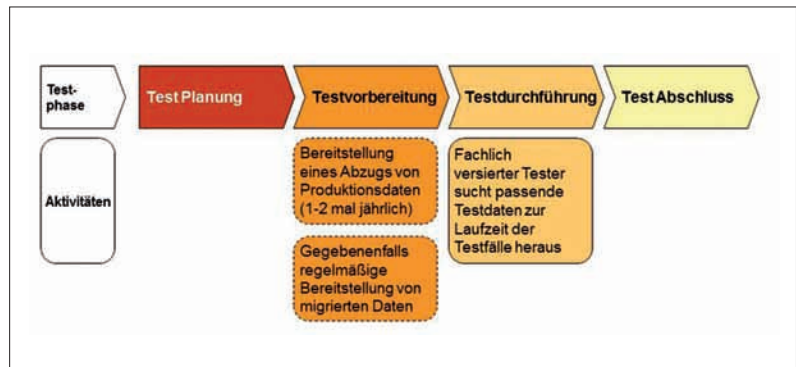


Abb. 1: Herkömmlicher Prozess zur Verwendung von Testdaten (vereinfacht)

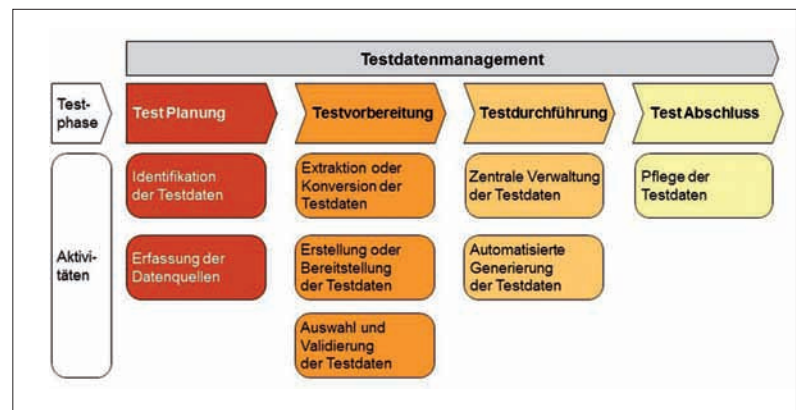


Abb. 2: Robuster Testdatenmanagementprozess

oder anonymisierten Testdaten ermöglicht, und bei dem die Testergebnisse trotzdem noch nachvollziehbar und aussagekräftig sind.

EIN END-TO-END-PROZESS FÜR TESTDATEN-MANAGEMENT

Ein Prozess für den Umgang mit Testdaten sollte also möglichst alle Testphasen abdecken, von der Planung bis zu Durchführung und Abschluss der Tests. Ein wohl definierter Prozess beginnt mit der Identifikation von Datenquellen, Datentypen und den eigentlichen Anforderungen an Testdaten. Eine besondere Rolle spielen hierbei Abhängigkeiten zwischen Daten untereinander (zum Beispiel bei Schnittstellen), aber auch Abhängigkeiten zu Datenquellen, die im Laufe der Testphase aktualisiert werden. In der Testvorbereitungsphase ist die Erfassung und Erstellung der Testdaten wichtig. Daten können mit Werkzeugen ausgewählt, extrahiert und validiert werden. In diesem Schritt sollte auch ein Verfremdungsansatz verwendet werden, um bei extrahierten Daten sensible Felder zu maskieren.

Werkzeuge können dabei eine entscheidende Rolle spielen. Einige Firmen, vor allem im Versicherungs-



Es gibt viele Szenarien, um Testdaten bereitzustellen – diese müssen aus verschiedenen Gesichtspunkten wie Datensicherheit, Datenschutz und Kosten-Nutzen-Verhältnis betrachtet werden.

Stephan Oswald: Als Beratungshaus und Technologiefirma dürfen Sie Ihre Kunden rechtlich nicht beraten. Was sagen Sie Kunden, die nichtanonymisierte Produktionsdaten zum Test verwenden wollen?

Matthias Rasking: Zunächst einmal ist es wichtig, dass man die Datenquelle als solche überhaupt kennt. Wir führen seit mehreren Jahren innerhalb von Accenture Schulungsprogramme durch, um auf die Aspekte Datenschutz und Datensicherheit in den verschiedenen Ländern hinzuweisen. Für diese Themen gibt es jährliche Auffrischungsschulungen, die verpflichtend sind. Da Accenture in über 120 Ländern Kunden berät und IT-Projekte umsetzt, ist das lokale Verständnis in Bezug auf Testdaten besonders wichtig. Unsere Testmanager sind geschult, gezielt Fragen über Testdaten zu stellen, um eine erste Risikoanalyse (angelehnt an CobiT und COSO) der zu bearbeitenden Daten durchführen zu können. Anhand dieser Analyse ist dann der Angebotsverantwortliche in der Lage, unsere Kunden auf wichtige Aspekte im Umgang mit Testdaten hinzuweisen, um den Kunden zusammen mit deren Datenschutzbeauftragten einen Lösungsweg aufzeigen können. Es gibt viele Szenarien, um Testdaten bereitzustellen – diese müssen unter verschiedenen Gesichtspunkten wie Datensicherheit, Datenschutz und Kosten-Nutzen-Verhältnis betrachtet werden. Hier kann Accenture zwar Beispiele für mögliche Lösungswege von anderen Projekten einbringen, die Entscheidung für ein Szenario liegt aber immer beim Kunden.

Oswald: Wenn Sie „Offshore“-Software testen, was ist für Ihre Kunden hinsichtlich personenbezogener Daten zu beachten?

Rasking: Grundsätzlich gehen wir davon aus, dass Daten für Testzwecke entweder anonymisiert sind oder wir synthetische Daten verwenden können. Der Zugriff auf personenbezogene Daten ist für externe Mitarbeiter ja generell nur in Verbindung mit einem separat aufgesetzten Verfahren möglich. Beim Einsatz von Mitarbeitern in „Offshore“-Lokationen treten wiederum besondere Probleme auf. Wichtig ist in diesem Kontext, dass offshore keine allgemeingültige Bezeichnung ist, sondern zwischen den entsprechenden Voraussetzungen in EU- und Nicht-EU-Ländern (also auch Marokko, Ägypten, Brasilien etc.) unterschieden wird. Für unsere Kunden haben wir innovative Modelle entwickelt, um bereits in frühen Phasen mit synthetischen Daten zu arbeiten, die unsere Testspezialisten selbst erzeugen, um dadurch eine Unabhängigkeit von personenbezogenen Daten zu erlangen. Zusammen mit Werkzeugen, die die Erstellung und Verwaltung von

Stubs und Simulatoren für Schnittstellen vereinfachen, können so Werkzeuge zur nachvollziehbaren und automatisierbaren Erzeugung von Testdaten sehr helfen.

Oswald: Was raten Sie Ihren Kunden hinsichtlich der Übergabe von Testdaten bei Cloud-basiertem Testing?

Rasking: Das Tückische an der Nutzung von Testdaten in Cloud-basierten Szenarien ist, dass der Testspezialist sich oftmals nicht mehr sicher sein kann, wo die Daten eigentlich gespeichert werden. Schauen Sie sich mal das Testmanagementtool eines Kunden an: Wenn ich in einem „Defect“ eine Fehlerbeschreibung anhängen, die gegebenenfalls personenbezogene Daten enthält – wo werden diese Daten dann gehalten? Viele große Unternehmen setzen bereits Software-as-a-Service-Lösungen ein, ohne sich im Klaren zu sein, wo genau ihre Daten nun sind, daher wird es immer wichtiger, mit synthetischen Daten zu arbeiten. Und das, ohne zu wissen und nachvollziehbar zu dokumentieren, welche Daten überhaupt in den Systemen und Werkzeugen gehalten werden.

Oswald: Würden Sie Ihren Kunden raten, sich nach neuen Methoden umzusehen, damit künftig keine Produktionsdaten mehr im Softwaretest eingesetzt werden müssen?

Rasking: Auf jeden Fall sollten sie diese Möglichkeit in Betracht ziehen, selbst wenn man für bestimmte Tests weiterhin auf produktive Daten angewiesen sein wird. Viele Unternehmen scheuen die Anfangsinvestition, das komplette Testfallportfolio so zu überarbeiten, dass ein Großteil mit synthetischen Daten getestet werden kann. Es wird dabei nur der Datenschutzaspekt betrachtet, nicht aber die enorme Effizienzsteigerung und bessere Planbarkeit von Tests mit synthetischen Daten. Tests werden häufig durch Mitarbeiter der Fachbereiche durchgeführt, die die wichtigsten Szenarien kennen und die in der Produktion relevanten Testdatensätze identifizieren können. Das führt oft dazu, dass Testfälle nur rudimentär beschrieben sind. Dadurch entfällt die Möglichkeit, Testspezialisten für diese Tests einzusetzen, die eventuell aus den Anforderungen effizientere Kombinationsmöglichkeiten und effektivere Testabdeckungen ableiten könnten. Ohne ein systematisch abgeleitetes Testfallportfolio verliert man aber den Überblick der Testabdeckung und damit der Produktqualität. Und zu einer systematischen Ableitung von Testfällen gehört nun mal die Definition der wichtigsten Parameter und damit der benötigten Testdaten. Mit synthetisch erzeugten Testdaten können vielfältige Kombinationen simuliert werden, die die Testabdeckung erhöhen und die Effektivität der Tests steigen lässt.

umfeld, verwenden zwar häufig selbstgeschriebene Programme, um bei einem Produktionsabzug Datenschutzrichtlinien umzusetzen, aber mittlerweile sind die Werkzeuge in diesem Bereich so flexibel geworden, dass sie auch komplexe Applikationslandschaften (zum Beispiel heterogene SAP- und Nicht-SAP-Systeme) abdecken können. Zudem sollte ein gestuftes System zur Testdatenbereitstellung eingesetzt werden:

- Für die frühen Teststufen können Testdatengeneratoren eingesetzt werden. Für diese Aktivitäten benötigt man üblicherweise eine geringe Einarbeitungszeit und kann auch wiederverwendbare Aufträge zur Testdatenerzeugung verwalten. Dazu müssen aber die Testfälle entsprechend granular beschrieben werden, damit Testdaten bereits frühzeitig erzeugt werden können.
- Um gesamte Testumgebungen mit verfremdeten Testdaten zu bestücken, können ebenfalls Werkzeuge zum Einsatz kommen. Diese Aktivitäten benötigen jedoch eine höhere Anfangsinvestition in die Datenmodellierung und in die Ausgestaltung der Prozesse, können dann aber zu einem sehr robusten Service für die Datenbereitstellung ausgebaut werden.

Letztendlich ist es also keine Frage des Toolsets, sondern der Prozesse für die Erstellung und Benutzung von Testdaten. Ob Open-Source- oder kommerzielle Lösung – verschiedene Aktivitäten wollen unterstützt werden.



Matthias Rasking leitet den Bereich Accenture Test Services in Deutschland, Österreich und der Schweiz sowie Accentures globale Testing Community of Practice mit über 13 000 Testspezialisten. Er ist seit mehr als 10 Jahren tätig für Kunden in verschiedenen Industrien in den Bereichen Software- und Systementwicklung und bei der Gestaltung und Implementierung von Test- und IT-Prozessen. Neben seinen Projektaufgaben bei Accenture leitet er als Working Group Manager den Bereich Model Development and Maintenance der TMMi foundation, einer Non-Profit-Organisation zur Etablierung von TMMi als Prozessstandardmodell im Testing.



Stephan Oswald war nach Ausbildung und Wanderjahren in der IT, von 1993 – 2007 Geschäftsführender Gesellschafter der Firmen ProConnect Consulting GmbH und PMWarehouse Solution GmbH in den Bereichen Service Delivery RZ und Hochverfügbarkeitslösungen, sowie einer der ersten Anbieter einer Cloud-Lösung für E-Mail-Marketing. Seit Anfang 2010 verantwortet er bei der GFB Softwareentwicklung den Aufbau Marketing und Vertrieb DACH. Mit Q-up, dem Testdatengenerator, treibt er auch hier wieder eine hochinnovative Lösung zum Erfolg.

Anzeige

m 360 bile

Der frische 360°-Blick auf mobile Entwicklungen!

PROVIDER

OS

ANDROID

APPS

EDITORS PICK

SMARTPHONES

IPHONE

E-READER

Der frische 360°-Blick auf mobile Entwicklungen

iPhone, Android und BlackBerry sind nicht mehr wegzudenken aus dem mobilen Leben der Menschen. Tendenz steigend. Mit immer neuen Geräten, neuen Anbietern und einem weiter steigenden Bedarf an Applikationen ist der Hunger nach Informationen in diesem Bereich schier unendlich. Dieser Entwicklung trägt das neue Webportal mobile360.de Rechnung, das über sämtliche Plattformen und Betriebssysteme hinweg berichtet. Zum Fokus des Portals gehören neben der Softwareentwicklung auch der Markt, Trends und Neuigkeiten im Bereich der App Stores sowie die verschiedenen Smartphones und Tablets – ein 360°-Blickwinkel auf das Mobile Development.

mobile360.de

