

Testdatenmanagement mit formalen Testdatenspezifikationen in einem großen Migrationsprojekt

Dierk Ehmke¹

Abstract: Für ein großes Migrationsprojekt in der Finanzindustrie wurde ein spezifikationsbasierter Testdatenmanagementansatz entwickelt. Die Testdaten werden automatisch im Ticket-system Jira bereitgestellt. Beschrieben werden Ansatz, Resultate und Voraussetzungen für eine erfolgreiche Einführung.

Keywords: formal specification; test data; test data management; migration project test

1 Einleitung

Aus einer jahrzehntealten Mainframe-Anwendung werden Millionen Verträge in ein bestehendes, modernes Zielsystem migriert. Dort existiert ein Bestand, zu dem die migrierten Daten hinzukommen. Nur im Altsystem vorhandene Geschäftsprozesse müssen implementiert werden. Wichtigstes Testziel ist die Korrektheit der migrierten Verträge sowie der neuen und der bestehenden Geschäftsprozesse.

Mehrere Testteams testen gemeinsam oder exklusiv auf den Testumgebungen. Das Testmanagement wird von erfahrenen internen und externen Mitarbeitern durchgeführt. Testfälle werden von besonders erfahrenen Sachbearbeitern erstellt. Sachbearbeiter sind es auch, die die Tests ausführen. Dem Testdatenmanagement steht eine neunmonatige Vorbereitungsphase zur Verfügung.

Als Ticketsystem wird Jira mit dem Testmanagementwerkzeug XRay eingesetzt. Für Regressionstests werden Testfälle automatisiert. Die Datenbanken des Zielsystems sind auf den Testumgebungen mit SQL erreichbar. Der Hersteller des Zielsystems unterstützt bei der Umsetzung fachlicher Spezifikationen in Datenbankabfragen.

Das Mengengerüst umfasst Millionen Verträge, hunderte Geschäftsprozesse, hunderte Projektbeteiligte (teilweise temporär, geschätzt ein Drittel davon im Test) und tausende Testfälle.

Die Aufgabenstellung besteht darin, zu jedem Testfall für die Testdurchführung nach fachlichen Kriterien passende Verträge auszuwählen. Wegen etwaiger Testwiederholungen müssen es mehrere sein. Da die Testdurchführungen oft die Daten so verändern, dass sie

¹ Freiberuflicher Softwaretestberater mit Schwerpunkt Testdatenmanagement, Herdweg 76, 64285 Darmstadt, mail@d-ehmke.de, <https://www.d-ehmke.de>

für folgende Testdurchführungen unbrauchbar werden, müssen die Verträge jeweils für genau eine Durchführung eines Testfalls reserviert werden.

Testdaten sind die migrierten anonymisierten Produktionsdaten. In einem Migrationsprojekt sind die migrierten Daten selbst Gegenstand des Tests, die Verwendung synthetischer Daten ist hier nicht zielführend, denn dann würde man durch die Migration verursachte Datenfehler nicht entdecken.

Dem Stand der Technik entsprechende Lösungsalternativen wurden in Betracht gezogen [Tr18]. A: die Tester suchen sich zu Beginn der Testdurchführung passende Verträge. B: in die Testfallbeschreibung wird die Nummer eines passenden Vertrages aufgenommen. C: Ein Testdatenmanager führt manuell Listen von passenden Verträgen zu jedem Testfall. D: Anschaffung eines geeigneten Testdatenmanagementsystems. E: Implementierung einer automatisierten Individuallösung.

A, B und C wurden verworfen. Wenn sich die Datenbasis ändert, weil zum Beispiel der Vertrag im Bestand verändert wurde und man (neue) Verträge suchen muss, ist nicht sichergestellt, dass man die Spezifikationen der erforderlichen Testdateneigenschaften hat. Für automatische Regressionstests ist das ein großes Problem, dass dadurch verschärft wird, dass man für die Testwiederholungen viele Verträge je Testfall benötigt. Alternative D – die Anschaffung eines passenden Testdatenmanagementsystems wurde verworfen, da dies zeitlich und organisatorisch nicht machbar war und nicht sichergestellt war, dass man am Ende ein im Budgetrahmen befindliches Produkt findet, dass alle Anforderungen erfüllt. Eine mit kleinen Schritten einführbare Individuallösung (E) wurde entwickelt und hat sich bewährt.

2 Lösung

Kern der Lösung sind formale Testdatenspezifikationen, die automatisch in Datenbanksuchabfragen für die Vertragssuche konvertiert werden. Die Notation ist eine einfach gehaltene Domain Specific Language und orientiert sich an der Fachterminologie in der Organisation. Beispiel:

```
TESTDATENSPEZIFIKATION
  benötigt wird ein Vertrag mit den Merkmalen
    Tarif = XYZ_2012
    Zahlart = 57
    Beginn >= 01.01.1970
    Beitrag >= 100
ENDE TESTDATENSPEZIFIKATION
```

Jeder Testfall beinhaltet eine Testdatenspezifikation, für die in Jira ein Feld implementiert wurde. Werden ausnahmsweise keine Testdaten benötigt, wird mit einem speziellen Text

ausgedrückt, dass dies intendiert ist und kein Versehen vorliegt. Aktuell werden etwa 75 Attribute und 1000 Schlüssel unterstützt. Vereinfacht verläuft der Prozess von der Testdatenspezifikation bis zur Testdatenbereitstellung wie folgt:

Erstellung: Die Testfallersteller müssen für jeden Testfall Testdatenspezifikationen erstellen, die in Jira eingestellt werden, siehe Abb. 1. Da die Erstellung einer korrekten formalen Spezifikation schwierig ist, wird sie mit einem Erfassungsformular unterstützt, in dem formale Fehler abgefangen werden. Realisiert wurde es mit einem marktgängigen Tabellenkalkulationsprogramm.

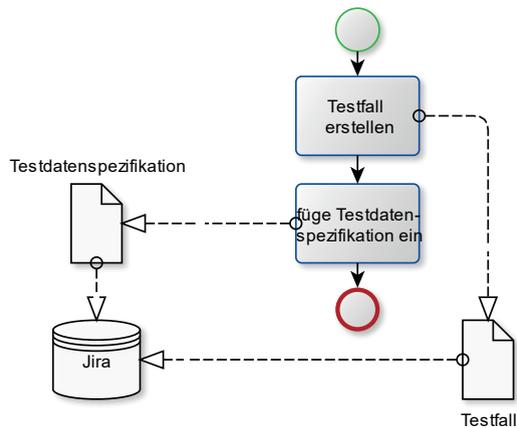


Abb. 1: Prozess der Erstellung einer Testdatenspezifikation

Reservierung: Vor Beginn einer Testkampagne stellen die Testmanager die durchzuführenden Testfälle zusammen und exportieren deren Testdatenspezifikationen aus Jira für die Reservierung durch das Testdatenmanagement. Ein Compiler übersetzt diese in ein SQL-Programm, das die Suche und Reservierung im Datenbestand der zugewiesenen Testumgebung durchführt.

Ein Compiler übersetzt diese in ein SQL-Programm, das die Suche und Reservierung im Datenbestand der zugewiesenen Testumgebung durchführt. Folgender SQL-Code setzt die Testdatenspezifikation weiter oben um:

```

select V.NR from VERTRAG as V
  where V.TARIF='XYZ_2012' AND V.BEGINN >= '01.01.1970'
        AND V.ZAHLART=57 AND V.BEITRAG = 100
        FETCH FIRST 5 ROWS ONLY
  
```

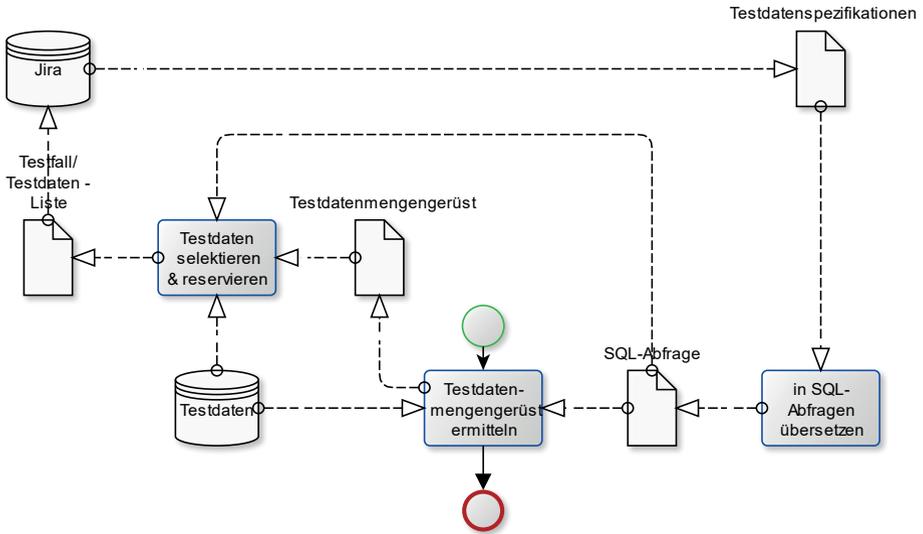


Abb. 2: Prozess der Testdatenbereitstellung

Bereitstellung, siehe Abb. 2: Ergebnis ist eine Liste von Testfallschlüsseln mit zugeordneten Verträgen, die so aufgebaut ist:

Testfall	Vertrag
MIG-2799	V52643921
MIG-2799	V20363262
MIG-9156	V98879715
MIG-9156	V59287826

Sie wird in Jira abgelegt.

Zugriff, siehe Abb. 3: Vor Teststart werden sogenannte Testruns für die Durchführung der Testfälle angelegt, das ist der Prozess in XRay. Automatisch wird jedem Testrun ein hinterlegter Vertrag zugewiesen, indem im Kommentarfeld ein Text eingefügt wird: "Dieser Testausführung ist Vertrag V52643921 zugewiesen".

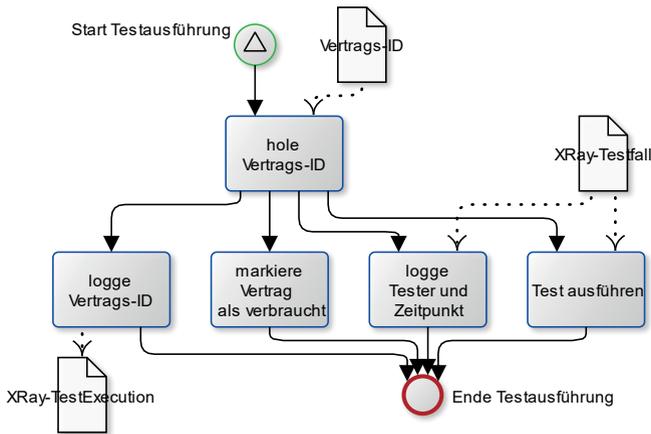


Abb. 3: Zugriff auf Testdaten während der Testausführung

In der Systemübersicht in Abb. 4 sieht man die beteiligten Systeme und den Datenfluss.

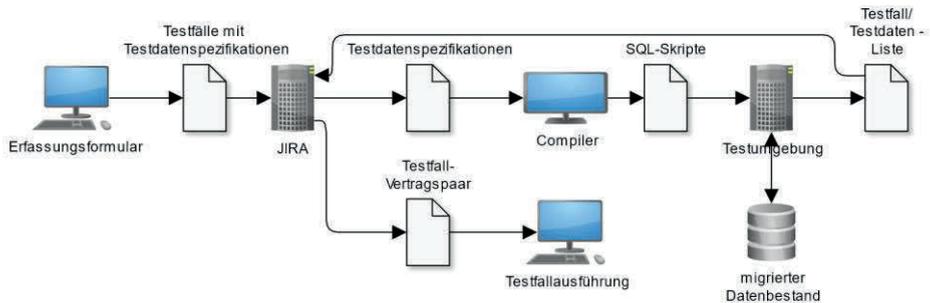


Abb. 4: Systemübersicht

Bei Inbetriebnahme des Verfahrens stellte sich heraus, dass für viele Testdatenspezifikationen keine passenden Testdaten gefunden wurden, weil die Testfallersteller nicht darauf geachtet hatten, dass es zu den Testdatenspezifikationen Verträge im Bestand gibt. Da die zugehörigen Testfälle ohne Testdaten nicht ausgeführt werden, konnte die geplante Testabdeckung nicht erzielt werden. Mit folgenden Maßnahmen wird gegengesteuert:

Schulung der Testfallersteller, damit sie nur die für den Testfall relevanten Attribut/Schlüssel-Kombinationen definieren.

Automatische Vereinfachung der Testdatenspezifikationen während der Testdatensuche: es gibt zum Beispiel Vertragseigenschaften, die durch das Zielsystem hergestellt werden

können. Die zugehörigen Suchkriterien dürfen eliminiert werden, wenn sie Treffer verhindern. Der Vorgang wird protokolliert und ist transparent für die Testmanager. Beispiel: der Vertrag soll im Zustand „beendet“ sein. Dieses Kriterium kann entfernt werden und der Tester führt zu Beginn des Tests den Geschäftsvorfall „Vertrag beenden“ durch, damit der Vertrag für den Testfall geeignet ist.

Erweiterung des Erfassungsformulars um eine Zählfunktionalität, mit der die Anzahl der Treffer für eine Testdatenspezifikation ermittelt wird.

Seltene Datenkonstellationen müssen für die Testfälle reserviert werden, die sie benötigen. Es hat sich bewährt, die Zuteilung bei den Spezifikationen mit den wenigsten Treffern zu beginnen.

Teilen sich mehrere Testfälle eine Testdatenspezifikation und können sie nicht alle die benötigte Anzahl Testdaten mit den geforderten Eigenschaften bekommen, werden die verfügbaren Verträge gleichmäßig an sie verteilt.

Nach Teststart benötigen Testmanager oder Tester aus verschiedenen Gründen spontan Testdaten mit bestimmten Eigenschaften. Dazu gibt es im Erstellungsformular eine Reservierungsfunktionalität, mit der sie passende, konsistente Verträge reservieren.

3 Resultate

Folgende Erfahrungen gibt es; auf die grundlegenden Funktionen, wie Reservierung, Bereitstellung passender Verträge usw. wird nicht näher eingegangen.

Ohne Akzeptanz gibt es keine Spezifikationen, das musste manches modell- oder spezifikationsbasierte Verfahren in der Praxis erfahren (wenn er nicht grundsätzlich an der Komplexität der Aufgabe gescheitert ist). Obendrein ist eine aus Sicht eines Informatikers einfache Notation dies noch lange nicht für die Mitarbeiter der Fachabteilung. Der vorgestellte Ansatz wird in fast allen Testteams konsequent umgesetzt. Testteams, die anfangs ein anderes Verfahren wählten, wechselten im Laufe des Projekts zum vorgestellten Ansatz.

Ein Datenbereitstellungsverfahren in einem solch großen Projekt muss skalieren. Mit steigender Anzahl Testfälle nimmt lediglich die Laufzeit der Skripte linear zu. Zusätzliche Testteams werden bedient, ohne dass das Testdatenmanagement personell aufgestockt werden muss.

Das Verfahren liefert für die automatisierten Regressionstestfälle beliebig - sofern es der Bestand hergibt - viele passende Testdaten, mit denen Regressionstestläufe ausgeführt werden und schafft so die Voraussetzung für langfristig stabil laufende automatische Tests.

Zwei wichtige Testdokumentationsaspekte werden durch das Verfahren sichergestellt. Zum einen werden die Testdateneigenschaften im Testfall exakt mit den formalen Testdatenspezifikationen dokumentiert und zum anderen wird lückenlos automatisch dokumentiert, welche Verträge im Testlauf verwendet werden.

Die Datensparsamkeit und damit die Verwendung kleiner dimensionierter Testumgebungen werden unterstützt. Bei großen Datenbeständen kann es vorteilhaft werden, auf den Testumgebungen mit Teilbeladungen zu arbeiten. Die Ladeskripte laufen wesentlich schneller und Ausgaben für Lizenzen und Hardware für die Testumgebungen sind kleiner als bei der Verwendung des gesamten Produktionsdatenbestands. Mit dem Verfahren lässt sich vorab der exakte Testdatenbedarf ermitteln, was die Voraussetzung für diese Vorgehensweise schafft.

Da Testdaten automatisch genau einer Durchführung eines Testfalls zugeteilt werden, können viele Testteams eine Testumgebung ohne Konflikte nutzen.

Dateninkonsistenzen werden effizient abgefangen. Aufgrund von Migrationsproblemen in mehreren Teilsystemen waren bis zu ein Drittel aller Daten betroffen. Beispielsweise sind dann versicherte Personen nicht in der ERP-Software hinterlegt und dadurch gibt es bei der Testausführung einen Fehler. Die Kriterien bzw. die Schlüssel der betroffenen Verträge werden in den Skripten hinterlegt und diese bei der Datensuche ausgeschlossen.

Der Code-Umfang für die Implementierung beträgt über 12.000 Lines of Code in Groovy, SQL, Shell-Skript und VBA.

4 Erfolgsfaktoren

Neben der erfolgreichen handwerklichen Umsetzung bedarf es der Akzeptanz in der Organisation, damit ein solcher Ansatz wirksam wird. Folgende Faktoren haben dazu beigetragen:

- die Einbeziehung der betroffenen Mitarbeiter während der Konzeption
- die Wahl einer Domain Specific Language für die Notation
- das komfortable Erstellungsformular mit Zählfunktion für schnelles Feedback für die Ersteller und Reservierungsfunktion für die schnelle Lösung von Testdatenproblemen durch das Testmanagement

- die automatische Zuweisung der Verträge beim Anlegen der Testausführungen
- die Unterstützung der Testautomation

Es braucht Erfahrung, um gute Testdatenspezifikationen zu erstellen, die einerseits zum Testfall passen und die Testabdeckung erhöhen und andererseits passende Testdaten liefern. Daher müssen Schulungen erfolgen und eine Lernkurve einkalkuliert werden, bis die Trefferquote der Testdatenspezifikationen akzeptabel wird.

5 Zusammenfassung

Die in das spezifikationsbasierte Testdatenmanagement gesetzten Erwartungen werden erfüllt. Neben technischen Aspekten tragen auch „weiche“ Faktoren für die Herstellung der Akzeptanz bei den Beteiligten zum Erfolg bei. Der vorgestellte Ansatz orientiert sich dabei eng an den vorgegebenen Prozessen und der vorhandenen technischen Infrastruktur. Die exakten Testdatenspezifikationen erlauben die Pflege der Testfälle und das Beschaffen von Testdaten auch noch dann, wenn die Ersteller der Testfälle nicht mehr zur Verfügung stehen. Das schafft nachhaltige automatisierte Regressionstests über die Laufzeit des Migrationsprojektes hinaus.

6 Danksagungen

Besonders bedanke ich mich bei Mario Friske für den Anstoß zu diesem Beitrag, Peter Weber für anregende Diskussionen in der Konzeptionsphase, den Kollegen im GI-AK TOOP für inspirierende Diskussionen, sowie Tanja Tremmel, Eckehard Kruse, Klaus Franz und Michael Kurz aus der Arbeitsgruppe ASQF Testdatenmanagement für viele Diskussionen und Anregungen in den vergangenen Jahren.

7 Literatur

[Tr18] Franz, K.; Tremmel, T.; Kruse, E.: Basiswissen Testdatenmanagement. dpunkt.verlag, Heidelberg, S. 114-120, 2018.