

GTB[®] Certified Tester
Lehrplan
Foundation Level
Test Data Specialist

Version 1.1



Deutschsprachige Ausgabe
herausgegeben durch das German Testing Board e. V.

Urheberschutzvermerk

Copyright © German Testing Board (nachstehend als GTB® bezeichnet).

Urheberrecht © 2017 die Autoren der deutschen Originalausgabe: Eckehard Kruse, Klaus Franz, Tanja Tremmel

Dieser GTB® Certified Tester Lehrplan Foundation Level – Test Data Specialist, deutschsprachige Ausgabe, ist urheberrechtlich geschützt.

Inhaber der ausschließlichen Nutzungsrechte an dem Werk ist German Testing Board e. V. (GTB®).

Die Nutzung des Werks ist – soweit sie nicht nach den nachfolgenden Bestimmungen und dem Gesetz über Urheberrechte und verwandte Schutzrechte vom 9. September 1965 (UrhG) erlaubt ist – nur mit ausdrücklicher Zustimmung des GTB® gestattet. Dies gilt insbesondere für die Vervielfältigung, Verbreitung, Bearbeitung, Veränderung, Übersetzung, Mikroverfilmung, Speicherung und Verarbeitung in elektronischen Systemen sowie die öffentliche Zugänglichmachung.

Dessen ungeachtet ist die Nutzung des Werks einschließlich der Übernahme des Wortlauts, der Reihenfolge sowie Nummerierung der in dem Werk enthaltenen Kapitelüberschriften für die Zwecke der Anfertigung von Veröffentlichungen gestattet. Die Verwendung der in diesem Werk enthaltenen Informationen erfolgt auf die alleinige Gefahr des Nutzers. GTB® übernimmt insbesondere keine Gewähr für die Vollständigkeit, die technische Richtigkeit, die Konformität mit gesetzlichen Anforderungen oder Normen sowie die wirtschaftliche Verwertbarkeit der Informationen. Es werden durch dieses Dokument keinerlei Produktempfehlungen ausgesprochen.

Die Haftung des GTB® gegenüber dem Nutzer des Werks ist im Übrigen auf Vorsatz und grobe Fahrlässigkeit beschränkt. Jede Nutzung des Werks oder von Teilen des Werks ist nur unter Nennung des GTB® als Inhaber der ausschließlichen Nutzungsrechte sowie der oben genannten Autoren als Quelle gestattet.

Änderungsübersicht der deutschsprachigen Ausgabe

Version	Datum	Änderungen
1.0	Juni 2017	Initiale Erstellung der Version 1.0 zur Veröffentlichung auf der Webseite des GTB®
1.1	März 2021	<ul style="list-style-type: none">• Anpassung an den ISTQB® Certified Tester Foundation Level (CTFL) Lehrplan Version 2018 V3.1D• Aktualisierte Darstellung des Testprozesses in Abschnitt 4.1• Ergänzung der Anonymisierungsverfahren in Abschnitt 2.4.2• Aktualisierung des Abschnitts 3.1 im Hinblick auf die DSGVO und das Bundesdatenschutzgesetz• Einfügen des Abschnitts 6.3 mit neuem Text• Formale Verbesserungen, inhaltliche Straffung mehrerer Abschnitte• Einfügen einer Gliederungsebene in Abschnitt 2.4 ohne neuen Text• Anpassen der Lernziele in Kapitel 3.1 an den Inhalt des Kapitels• Aktualisierung der Quellenangaben• Ergänzen von Begriffen im Glossar dieses Lehrplans• Einfügen des Index am Ende des Dokuments• Übernahme in die Vorlage für Lehrpläne nach dem ISTQB® Schema

Inhaltsverzeichnis

Urheberschutzvermerk	2
Änderungsübersicht der deutschsprachigen Ausgabe	3
Inhaltsverzeichnis	4
Danksagung	7
Motivation	8
Steckbrief zum Test Data Specialist	9
0. Einführung in diesen Lehrplan	10
0.1 Zweck dieses Dokuments	10
0.2 GTB® Certified Tester Foundation Level Test Data Specialist	10
0.3 Geschäftlicher Nutzen	11
0.4 Prüfungsrelevante Lernziele	11
0.5 Empfohlene Unterrichtszeiten	12
0.6 Zulassung zur Prüfung	12
0.7 Informationsquellen	12
1. Testdatenmanagement: Herausforderungen, Risiken, Nutzen – 75 Min.	13
1.1 Ziele des Testdatenmanagements	14
1.2 Herausforderungen für das TDM	14
1.3 Kosten und Nutzen des TDM	15
1.3.1 Nutzen des TDM	15
1.3.1.1 Wiederverwendbarkeit und Verfügbarkeit	15
1.3.1.2 (Datenschutz-) Rechtliche Absicherung	16
1.3.2 Kosten des TDM	16
1.4 Risiken durch unzureichendes TDM	17
1.4.1 Risiko Datenschutz	17
1.4.2 Risiko Datensicherheit	17
1.4.3 Risiko Projekterfolg	17
1.4.4 Risiko Testqualität	18
2. Grundlagen und Verfahren – 300 Min.	19
2.1 Definition TDM	20
2.2 Testfall und Testdaten	20
2.2.1 Abhängigkeit abstrakter Testfall, konkreter Testfall, Testdaten	20

2.2.2	Definition Testdaten.....	21
2.2.3	Struktur von Testdaten	21
2.3	Testdatenarten	22
2.3.1	Klassifizierung der Testdaten nach Verwendungsart	22
2.3.2	Klassifizierung der Testdaten nach Gewinnungsart	22
2.3.3	Eigenschaften von Testdaten	24
2.4	Erzeugung von Testdaten	24
2.4.1	Produktionsabzug.....	24
2.4.2	Anonymisierung.....	26
2.4.3	Synthetische Erzeugung.....	28
3.	Gesetzliche Regeln und Normen – 90 Min.....	30
3.1	Gesetzliche Regeln und Normen.....	31
3.1.1	Gesetze & Normen	31
3.1.2	Datenschutz (rechtliche Konformität).....	32
3.1.3	Datensicherheit	33
3.1.4	Auditierbarkeit	34
3.2	Branchenspezifische Regeln (rechtliche Konformität)	34
3.3	Unternehmensinterne Regeln (IT-Governance)	35
4.	Testdatenmanagement-Prozess – 225 Min.....	36
4.1	Der TDM-Prozess.....	37
4.2	Hauptaktivitäten im TDM-Prozess	39
4.2.1	Planung (TDM)	39
4.2.2	Steuerung (TDM).....	40
4.2.3	Testdaten spezifizieren.....	40
4.2.4	Testdatenerzeugung konzipieren	41
4.2.5	Testdaten bereitstellen	41
4.2.6	Testdatenabschlussbericht erstellen	42
4.2.7	Testdaten archivieren	42
4.3	Arbeitsergebnisse im TDM-Prozess	43
4.3.1	Testdatenrichtlinie	43
4.3.2	Testdatenkonzept.....	43
4.3.3	Testdatenspezifikation.....	45
4.3.4	Testdatenbereitstellungsroutinen	45

4.3.5	Testdatenbereitstellungsplan.....	46
4.3.6	Testdaten	46
4.3.7	Testdatenbereitstellungsbericht.....	46
4.3.8	Testdatenabschlussbericht.....	46
4.3.9	Testdatenbewertungsbericht	47
4.4	Rollen im TDM-Prozess	47
4.4.1	Testdatenmanager	47
4.4.2	Testdatenmodellierer.....	49
4.4.3	Testdatenrealisierer.....	50
5.	Einbindung in die Organisation – 60 Min.	52
5.1	Einbindung in die Organisation	53
5.2	Verantwortlichkeiten	53
5.3	Einbindung in die Prozesse	53
5.4	Einbindung in die IT.....	54
5.5	Einführung des TDM in die Organisation.....	54
6.	Werkzeuge – 45 Min.....	56
6.1	Typen von Werkzeugen	57
6.2	Werkzeuge im TDM-Prozess	57
6.2.1	Planung und Steuerung.....	57
6.2.2	Testdaten spezifizieren.....	57
6.2.3	Testdatenerzeugung konzipieren	58
6.2.4	Testdaten bereitstellen	59
6.2.5	Testdatenabschlussbericht erstellen	59
6.2.6	Testdaten archivieren	60
6.3	Besondere Betrachtung von Testdaten für die Testautomatisierung.....	60
7.	Glossar	61
8.	Literaturhinweise	66
8.1	Normen und Standards	66
8.2	ISTQB®-Dokumente	67
8.3	Fachliteratur	67
8.4	Internetquellen	67
9.	Index.....	68

Danksagung

Das Autorenteam der deutschen Originalausgabe dieses Dokuments dankt den Reviewern der Version 1.0 für ihre Kommentare und Beiträge (in alphabetischer Reihenfolge):

Graham Bath, Armin Beer, Armin Born, Christian Brandes, Renzo Cerquozzi, Dr. Klaudia Dussa-Zieger, Manuel Fischer, Florian Fieber, Dr. Matthias Hamburg, Günter Jung, Michael Knecht, Anke Löwer, Dr. Armin Metzger, Jörn Münzel, Horst Pohlmann, Eike Riedemann, Prof. Dr. Ina Schieferdecker, Dr. Frank Simon, Stefan Stefan, Sandro Zazzali.

Besonderer Dank gilt den Mitwirkenden an der vorliegenden Überarbeitung und Aktualisierung zur Version 1.1 (in alphabetischer Reihenfolge):

Klaus Franz, Jan Giesen, Eckehard Kruse, Elke Mai, Markus Niehammer, Tanja Tremmel.

Motivation

Seit 2012 gibt es den ASQF® Arbeitskreis „Testdatenmanagement“, der zum Thema Testdatenmanagement öffentlich zugängliche Ausarbeitungen erstellt und auf diversen Veranstaltungen zu dem Thema referiert.

Ein Ziel des Arbeitskreises ist, einen (Quasi-)Standard zu erarbeiten.

Ein Teil der Arbeitsgruppe möchte das Thema Testdatenmanagement mit einer standardisierten Schulung in die Breite bringen, um das Testdatenmanagement in den Unternehmen zu vereinheitlichen. Der ISTQB®-Testprozess soll durch den Testdatenmanagementprozess ergänzt, nicht geändert werden.

Steckbrief zum Test Data Specialist

Dauer:	2 Tage
Einordnung:	Ergänzung zum Certified Tester Foundation Level
Voraussetzung der Teilnehmer:	Zertifikat Certified Tester Foundation Level
Voraussetzung der Trainer:	Mindestens ein Zertifikat Certified Tester Advanced Level
Zielgruppe:	<p>Personen die zu Testzwecken mit sensiblen, schützenswerten Daten, Massendaten, Daten im Zusammenhang mit Testautomation, spezifischen Anforderungen an Testdaten arbeiten.</p> <p>Tester, Testmanager, Projektmanager, QS-Verantwortliche, QS-Beauftragte, Datenschutzbeauftragte, Entwickler</p>

0. Einführung in diesen Lehrplan

0.1 Zweck dieses Dokuments

Dieser Lehrplan bildet die Grundlage für die Qualifikation als Test Data Specialist auf der Aufbaustufe Foundation Level des Softwaretest-Ausbildungsprogramms des International Software Testing Qualifications Board (im Folgenden ISTQB® genannt).

Das GTB® stellt den Lehrplan folgenden Adressaten zur Verfügung:

- Prüfungsinstitutionen zur Erarbeitung von Prüfungsfragen in der jeweiligen Landessprache, die sich an den Lernzielen der jeweiligen Lehrpläne orientieren.
- Ausbildungsanbietern zur Erstellung ihrer Kursunterlagen und zur Bestimmung einer geeigneten Unterrichtsmethodik.
- Prüfungskandidaten zur Vorbereitung auf die Prüfung (im Rahmen eines Schulungskurses oder des freien Lernens).
- Allen Personen, die im Bereich Software- und Systementwicklung tätig sind und ihre fachliche Kompetenz beim Testen von Software verbessern möchten, sowie als Grundlage für Bücher und Fachartikel.

Das GTB® kann die Nutzung dieses Lehrplans auch anderen Personenkreisen oder Institutionen für andere Zwecke genehmigen, sofern diese vorab eine entsprechende schriftliche Genehmigung einholen.

0.2 GTB® Certified Tester Foundation Level Test Data Specialist

Die Qualifikation zum Foundation Level Test Data Specialist richtet sich an alle Personen im Bereich Softwaretesten, die ihr Wissen über Testdatenmanagement vertiefen möchten, bzw. an Personen, die sich in ihrer beruflichen Laufbahn auf das Management von Testdaten spezialisieren wollen.

Dieser Lehrplan behandelt die folgenden Hauptaspekte des Testdatenmanagements:

- Testdatenmanagement ist die Menge aller methodischen, konzeptionellen, organisatorischen und technischen Maßnahmen und Verfahren zur Behandlung der Ressource Testdaten
- mit dem Ziel, die Testdaten mit ihrem maximalen Nutzungspotenzial unter Berücksichtigung von Datenschutzerfordernungen
- in die Testprozesse einzubringen und die effiziente Nutzung der Testdaten zu gewährleisten.

0.3 Geschäftlicher Nutzen

Dieser Abschnitt zeigt den geschäftlichen Nutzen auf, der von Kandidaten mit Zertifizierung im Foundation Level Test Data Specialist erwartet werden kann.

Business Outcomes	Kapitel
TDM-1. Die mit Testdatenmanagement verbundenen Ziele, Herausforderungen, Kosten, Nutzen und Projektrisiken überzeugend darstellen.	1
TDM-2. Die Anforderungen an Testdaten formulieren und Verfahren zur Auswahl oder Erstellung von Testdaten anwenden.	2
TDM-3. Ein Projekt bezüglich TDM-relevanter regulatorischer Vorgaben beraten.	3
TDM-4. Das Projekt bezüglich TDM im Testprozess beraten und unterstützen.	4
TDM-5. Testdatenkonzepte und andere TDM-Ergebnisartefakte erstellen.	4.3
TDM-6. Die im TDM-Prozess zu besetzenden Rollen kennen und einplanen.	4.4
TDM-7. Dazu beitragen, den TDM-Prozess in die Unternehmensorganisation und deren Prozesse und IT zu integrieren.	5
TDM-8. Die für den Kontext geeigneten Arten von Testdatenwerkzeugen vorschlagen.	6

0.4 Prüfungsrelevante Lernziele

Die Lernziele unterstützen die geschäftlichen Ziele und dienen zur Ausarbeitung der Prüfung für die Zertifizierung zum Foundation Level Test Data Specialist. Den einzelnen Lernzielen ist jeweils eine kognitive Wissensstufe (K-Stufe) zugeordnet.

Die K-Stufe (bzw. kognitive Wissensstufe) dient dazu, Lernziele gemäß der überarbeiteten Taxonomie von Bloom [Anderson01] zu klassifizieren. Das ISTQB® verwendet diese Taxonomie bei der Erstellung der Prüfungen zu den Lehrplänen.

Dieser Lehrplan berücksichtigt folgende der insgesamt sechs möglichen kognitiven Wissensstufen:

Stufe	Kennwort	Beschreibung
1	Kennen	Der Lernende kann einen Begriff oder ein Konzept erkennen, abrufen und sich daran erinnern.
2	Verstehen	Der Lernende kann Aussagen zum Thema begründen oder erläutern.
3	Anwenden	Der Lernende kann erworbenes Wissen auf gegebene neue Situationen übertragen oder zur Problemlösung anwenden.
4	Analysieren	Der Lernende kann Informationen mit Bezug auf eine Vorgehensweise oder ein Verfahren zum besseren Verständnis in ihre einzelnen Bestandteile aufgliedern sowie zwischen Fakten und Rückschlüssen unterscheiden.

Allgemein gilt, dass der gesamte Inhalt des vorliegenden Lehrplans entsprechend der Lernziele der kognitiven Stufe K1 geprüft werden kann. Dies bedeutet, dass die Prüfungskandidaten Begriffe oder Konzepte erkennen, sich an sie erinnern und sie wiedergeben können. Die relevanten Lernziele der kognitiven Stufen K2 und höher werden immer zu Beginn des jeweiligen Kapitels angegeben.

0.5 Empfohlene Unterrichtszeiten

Für jedes Lernziel dieses Lehrplans wurde die Mindestschulungszeit festgelegt. Die gesamte Unterrichtszeit für jedes Kapitel ist in der Kapitelüberschrift angegeben.

0.6 Zulassung zur Prüfung

Voraussetzung für die Prüfung zum Foundation Level Test Data Specialist ist das erworbene Zertifikat zum ISTQB® Certified Tester Foundation Level (CTFL).

0.7 Informationsquellen

Die im Lehrplan verwendeten Fachbegriffe des Testens in Deutsch sind im GTB®/ISTQB® Standardglossar der Testbegriffe definiert [GTB-GLOSSAR].

Kapitel 8 enthält eine Liste empfohlener Bücher und Fachartikel zum Thema Testdatenmanagement.

1. Testdatenmanagement: Herausforderungen, Risiken, Nutzen – 75 Min.

Schlüsselbegriffe

Datenschutz, Datensicherheit, Testdaten, Testinfrastruktur, Testumgebung

Lernziele

1.1 Ziele des Testdatenmanagements

TDM - 1.1 (K1) Die Ziele von Testdatenmanagement nennen können.

1.2 Herausforderungen des Testens bezüglich Testdaten

TDM - 1.2 (K1) Die Herausforderungen des Testens bezüglich Testdaten beschreiben können.

1.3.1 Nutzen des TDM

TDM - 1.3.1 (K1) Den Nutzen von Testdatenmanagement nennen können.

1.3.2 Kosten des TDM

TDM - 1.3.2 (K1) Die Kostenfaktoren von TDM nennen können.

1.4 Risiken durch unzureichendes TDM

TDM - 1.4 (K1) Risiken kennen, die durch unzureichendes TDM entstehen können.

1.1 Ziele des Testdatenmanagements

Ein erfolgreiches systematisches Testdatenmanagement (TDM) wird es einem Unternehmen ermöglichen, Testdaten effizienter zu verwalten, projektübergreifend Synergien bei der Definition, Bereitstellung und Nutzung von Testdaten zu heben, bereits bestehende Testdatenbestände zu schützen sowie neue Bestände zielgerechter bereitzustellen und dabei den rechtlichen und regulatorischen Anforderungen zu genügen.

In diesem Zusammenhang sollte eine effiziente und sparsame Datenhaltung genauso abgewogen werden wie eine geeignete Werkzeugunterstützung.

Die definierten und beschriebenen Rollen des Testdatenmanagements zeigen, welche organisatorischen, fachlichen und technischen Kenntnisse und Erfahrungen erforderlich sind. Um Testdatenmanagement erfolgreich in eine Organisation zu integrieren, ist es empfehlenswert zusätzlich zum Testbereich verschiedene Fachbereiche ins Boot zu holen, wie Datenschutz, Recht und Compliance.

Testdatenmanagement – Ergänzend zum ISTQB®-Glossar:

Testdatenmanagement ist die Menge aller methodischen, konzeptionellen, organisatorischen und technischen Maßnahmen und Verfahren zur Behandlung der Ressource Testdaten mit dem Ziel, die Testdaten mit ihrem maximalen Nutzungspotenzial in den Testprozess einzubringen und die effiziente Nutzung der Testdaten zu gewährleisten.

Darüber hinaus muss ein professionelles Testdatenmanagement auch die Aspekte der Aussagequalität von Tests sowie des Datenschutzes berücksichtigen.

1.2 Herausforderungen für das TDM

In der Definition, Bereitstellung und im Umgang mit Testdaten gilt es, zahlreiche Herausforderungen zu bewältigen. Dazu zählen:

- Zielgerechte Testdatenbereitstellung
- Rechtzeitige Planung und (bei Bedarf) Beauftragung der Testdatenbereitstellung
- Klare Verantwortlichkeiten
- Verbindung zwischen Testumgebung und Testdatenanforderung
- Verknüpfung von Testfall und Testdaten
- Auflösung von Wissensmonopolen

- Wiederverwendbarkeit von Testdaten
- Effiziente Verwaltung und Pflege von Testdaten
- Beachtung regulatorischer Vorgaben
- Einbindung in den Testprozess
- Sehr umfangreiche oder komplexe Daten
- Exklusivität oder Mehrfachnutzung von Daten und Infrastruktur

Bei der Planung und Beauftragung von Testdaten in Zusammenhang mit der Bereitstellung der Testumgebung muss beispielsweise darauf geachtet werden, dass die Daten, vor allem wenn sie aus verschiedenen Schnittstellen kommen, zum Zeitpunkt der Testausführung konsistent sind.

1.3 Kosten und Nutzen des TDM

Systematisches Testdatenmanagement bietet nützliche Aspekte, die sich in Kostenersparnis und effizienteren Prozessen auswirken. Bei der Einführung und beim Betrieb von TDM entstehen aber auch Kosten, die dem erzielbaren Nutzen gegenübergestellt werden sollten.

1.3.1 Nutzen des TDM

TDM muss sich daran messen lassen, wie gut es Teststrategie und Testprozess unterstützt. Der Nutzen des TDM liegt sowohl in der Optimierung der Vorgehensweise in einem einzelnen Projekt als auch in einer Organisation. Als weiterer wichtiger Nutzenfaktor ist die rechtliche Absicherung zu nennen.

1.3.1.1 Wiederverwendbarkeit und Verfügbarkeit

Wenn durch das TDM sichergestellt ist, dass die Testdaten wiederverwendbar abgelegt werden und/oder jederzeit reproduziert werden können, kann dadurch eine höhere Effizienz in den Regressionstests erzielt werden.

Wenn Testdaten (abstrakte bzw. konkrete) Testfälle referenzieren, die wiederum mit Anforderungen verknüpft sind, kann mit wohldefinierten Testdatenbeständen die Überdeckung von Anforderungen durch Testfälle nachgewiesen werden.

Werkzeuge zur Generierung synthetischer Testdaten, zur Anonymisierung produktiver Daten und zur Verwaltung von Testdaten im Zusammenhang mit Testfällen können den Testprozess deutlich beschleunigen.

Nicht selten wird die Erzeugung von Testdaten in den Prozess der Testautomatisierung einbezogen. Dies spielt besonders dann eine Rolle, wenn die

einem Testfall zugrunde liegenden Daten synthetisch erzeugt werden und diese erst im Moment der Durchführung eines Testfalls entstehen.

1.3.1.2 (Datenschutz-) Rechtliche Absicherung

Die Konformität zu Datenschutzrichtlinien entsteht dadurch, dass personenbezogene Daten beim Testen entweder gar nicht oder nur in pseudonymisierter bzw. anonymisierter Form verwendet werden (Abschnitt 2.3.2). Außerdem kann Revisionssicherheit erreicht werden, indem der Lebenszyklus von Testdaten nachvollziehbar gemacht wird.

Dazu müssen die Testdaten der Versionskontrolle des Projekts unterliegen und die Verfolgbarkeit sichergestellt sein. D. h. die folgende Frage kann jederzeit beantwortet werden: Welche Anforderung wurde mit welchen Testfällen und Testdaten getestet und welche Abweichungen wurden dabei entdeckt?

Indem zusätzlich zur Verknüpfung von Anforderung, Testfall und Abweichung auch die Information über die betreffenden Testdaten zur Verfügung gestellt wird, unterstützt das TDM den Testprozess hinsichtlich seiner Revisionssicherheit.

Gesetzliche Regeln und Normen werden in Kapitel 3 beschrieben.

1.3.2 Kosten des TDM

Folgende Kostenfaktoren bringt TDM mit sich:

- (Selbst-) Kosten des Prozesses:
 - Kosten für die Implementierung des Prozesses
 - Kosten für etwaigen personellen Mehraufwand
 - Kosten für den gesamten Eigenaufwand der Testdatenproduktion
- Kosten für Werkzeuge:
 - Evaluierungskosten
 - Lizenzkosten
 - Schulungskosten einschließlich der personellen Kosten
 - Zeit und Aufwand, um einen signifikanten Nutzen aus der Anwendung des Werkzeugs ziehen zu können
 - Wartungskosten
- Kosten für Hardware

Mit der optimal gewählten Größe und Beschaffenheit des Testdatenbestands können durch effizientes Testdatenmanagement die Folgekosten der Testdaten-Bereitstellung reduziert werden. Solche Folgekosten können sich beispielsweise in der Bereitstellung und Wartung der Testinfrastruktur auswirken, z. B. im kostspieligen Unterhalt von zahlreichen Testumgebungen mit umfangreichem Datenbestand.

1.4 Risiken durch unzureichendes TDM

Wenn die Herausforderungen (siehe oben) des TDM nicht gemeistert werden, entstehen durch "schlechte" Testdaten und "ungeeignetes" Testdatenmanagement Risiken im rechtlichen Bereich, für das Projekt und für das Produkt. „Schlechte“ Testdaten sind fehlende, ungenaue, falsche, unzureichend verschlüsselte, duplizierte oder nicht zeitnah verfügbare Daten sowie sensible interne Informationen, die unverfälscht aus der Produktion in Testsysteme übernommen wurden. Ein „ungeeignetes“ Testdatenmanagement stellt durch ineffiziente Ablage und fehlende Strukturierung Daten nicht so bereit, dass ein an der Testdatenbereitstellung unbeteiligter Dritter diese versteht oder findet.

1.4.1 Risiko Datenschutz

Werden die Vorgaben zum Schutz personenbezogener Daten z. B. laut Datenschutz-Grundverordnung (DSGVO) nicht eingehalten und beispielsweise die Daten nicht anonymisiert, kann dies im Fall der unbefugten Nutzung zu einem Imageschaden und zu Vertrauensverlust führen. Im Schadensfall kann dies auch Geldstrafen nach sich ziehen (siehe auch Kapitel 3).

1.4.2 Risiko Datensicherheit

Das Risiko einer Sicherheitslücke für unternehmensbezogene sensible Daten besteht, wenn interne Regelungen zum Umgang mit diesen Daten nicht existieren oder nicht eingehalten werden. Solche Regeln können das Risiko nur reduzieren und nie ganz ausschalten, da es auf technischer Ebene fast immer noch Möglichkeiten zur Umgehung von Beschränkungen gibt, zum Beispiel für den Administrator eines Systems.

1.4.3 Risiko Projekterfolg

Wenn Testdaten nicht rechtzeitig zur Verfügung stehen, kann die Testdurchführung möglicherweise nicht wie geplant beginnen.

Wird bei der Bereitstellung von Testdaten nicht auf ein strukturiertes Vorgehen zum Datenmanagement geachtet, ist eine richtige Interpretation und erneute Verwendung erschwert (z. B. durch veraltete, schlecht gewartete, fehlerhafte Testdaten).

Beim Regressionstest wird erneut auf Testdaten zurückgegriffen. Sind die zugrundeliegenden Testdaten nicht reproduzierbar oder wiederauffindbar, können unter Umständen Regressionstests nicht oder nur ineffizient ausgeführt werden.

Wenn viele Testdaten manuell erzeugt werden, können Fehler in den Daten auftreten oder ungeplante Mehrarbeit entstehen. Wenn die beteiligten Personen nicht mit dem TDM-Prozess und den TDM-Tools vertraut sind, werden zudem unnötig Ressourcen verschwendet.

1.4.4 Risiko Testqualität

Die gewünschte Qualität der zu testenden Software kann nicht gewährleistet werden, wenn die Testdaten nicht zu den Testaufgaben passen, unzureichend sind oder wenn diese schlicht fehlen.

Beim kompletten oder auszugsweisen Kopieren von Daten aus dem Produktivsystem in ein Testsystem, dem sogenannten „Produktionsabzug“ bestehen einige Risiken. Produktionsabzüge haben nicht nur Risiken bezüglich des Datenschutzes und der Datensicherheit. Darüber hinaus haben sie Einfluss auf die Testqualität. Oft entsprechen die Datenbestände nicht dem Gebot der Datenminimierung, sind veraltet oder enthalten nicht die zu den Testzielen passenden Daten bzw. diese sind nur mit hohem Aufwand zu finden.

Im Gegensatz zu produktiven Daten werden synthetische Daten künstlich manuell oder mit Hilfe von Werkzeugen erzeugt. Synthetische Testdaten beeinflussen die Testqualität in Bezug auf ihre Realitätsnähe, die entsprechend der Testziele gefordert sein kann. Für einen Negativtest müssen ggf. Daten gezielt manipuliert werden und Realitätsnähe ist hierbei eher hinderlich.

2. Grundlagen und Verfahren – 300 Min.

Schlüsselbegriffe

abstrakter Testfall, konkreter Testfall, synthetische Daten, Produktivdaten, Anonymisierung, Anonymisierungsgrad, Pseudonymisierung, pseudonymisierte Daten, Testdatenart, Testdatenverwendungsart, Testdatengewinnungsart, Testdatenkoeffizient, Testdatenmanagement, Metadaten, Datenminimierung, Datenmaskierung

Lernziele

2.2 Testfall (abstrakter, konkreter) und Testdaten

TDM - 2.2.1 (K3) Aus abstrakten Testfällen Anforderungen an Testdaten ableiten.

TDM - 2.2.2 (K3) In der Lage sein, für konkrete Testfälle Testdaten zu erstellen.

2.3 Testdatenarten

TDM - 2.3.1 (K2) Testdatenarten nach Verwendungsart einordnen können.

TDM - 2.3.2 (K2) Testdatenarten nach Gewinnungsart einordnen können.

TDM - 2.3.3 (K4) Testdatenarten in einem Szenario erkennen können, aus einem Szenario die benötigten Testdatenarten ableiten können.

2.4 Erzeugung von Testdaten

TDM - 2.4.1 (K2) Anwendbarkeit von synthetischen, produktiven, anonymisierten und pseudonymisierten Testdaten in einem Szenario begründen.

TDM - 2.4.2 (K3) Methoden der Datenmaskierung bei der Erzeugung von Testdaten anwenden können.

2.1 Definition TDM

Die folgenden Definitionen für den Begriff „Testdatenmanagement“ des ISTQB® und der Arbeitsgruppe TDM befinden sich aktuell noch in einem Abstimmungsprozess:

- The process of analyzing test data requirements, designing test data structures, creating and maintaining test data.

Quelle: ISTQB®/GTB® V3.1

- Testdatenmanagement ist die Menge aller methodischen, konzeptionellen, organisatorischen und technischen Maßnahmen und Verfahren zur Behandlung der Ressource Testdaten mit dem Ziel, die Testdaten mit ihrem maximalen Nutzungspotenzial unter Berücksichtigung von Datenschutzerfordernungen in die Testprozesse einzubringen und die effiziente Nutzung der Testdaten zu gewährleisten.

Quelle: ASQF® AG TDM 2014

2.2 Testfall und Testdaten

2.2.1 Abhängigkeit abstrakter Testfall, konkreter Testfall, Testdaten

Für das Testobjekt werden abstrakte Testfälle beschrieben. Diese lassen sich gegen die konkreten Testfälle dadurch abgrenzen, dass nur in konkreten Testfällen Testdaten zu finden sind.

- Ein abstrakter Testfall ist ein Testfall mit abstrakten Vorbedingungen, Eingabedaten, erwarteten Ergebnissen, Nachbedingungen und Aktionen (falls anwendbar).
- Testdaten werden in einem abstrakten Testfall durch „Parameter“ als Platzhalter ersetzt. In einem weiteren Schritt der Testspezifikation werden formale Parameter durch das Einfügen konkreter Werte zu konkreten Testfällen. Abstrakte Testfälle werden beispielsweise dann mit formalen Parametern versehen, wenn das Ersetzen durch echte Werte einen Schritt zur Testautomatisierung darstellt. Die Vorbereitungen zugunsten der Testautomatisierung sind hierbei von Anfang an einzuplanen.
- Ein konkreter Testfall ist ein Testfall mit konkreten Werten für Vorbedingungen, Eingaben, erwartete Ergebnisse und Nachbedingungen sowie eine detaillierte Beschreibung der Aktionen (falls anwendbar). Durch Einfügen verschiedener Werte innerhalb des erlaubten Wertebereichs eines Parameters wird aus einem einzelnen konkreten Testfall eine Reihe von

Testfallvarianten. Diese unterscheiden sich lediglich in den Testdaten (und hier nur in wenigen oder einzelnen Parametern), die jeweils in den gleichen Testschritten verarbeitet werden. Statt eines Parameters, der durch einen absoluten Wert ersetzt wird (z. B. 31.12.2017), kann auch eine Funktion verwendet werden, die einen relativen Wert einsetzt (z. B. Funktion „heute“ -1).

- Testdaten werden für die Verwendung in konkreten Testfällen entworfen. Die Qualität der Testdaten, die im konkreten Testfall verwendet werden, ist entscheidend für den Erfolg des Testvorhabens. Bei der Auswahl von Testdaten sollte auf realistische Kombinationen von Wertebereichen der Parameter geachtet werden, mit Ausnahme von Negativtests, bei denen dies nicht erwünscht ist.

2.2.2 Definition Testdaten

Siehe Definition ISTQB® im Glossar. Ergänzend:

Testdaten sind konkrete Werte von Datenobjekten, die zur Ausführung eines Testobjekts benötigt werden. Sie können z. B. in Dateiform, in Form von Datenbanktabellen oder in Listenform vorliegen.

- Unter Testdaten sind Werte zu verstehen, die physikalisch vorliegen können oder
- in anderer Weise mit konkreten Testfällen in Assoziation stehen können.

Testdaten können auch als dynamische Objekte vorliegen, z. B. in einer definierten Einheit aus Datenquelle, Datenstruktur und Regelwerk (bzw. sogar ausführbarem Code) zur Verarbeitung / Interpretation.

2.2.3 Struktur von Testdaten

Testdaten liegen in einfachster Form als 1 bis n strukturidentische Einzeldatensätze vor. Die jeweilige Struktur ist abhängig von den Inhalten und Typen der Datenfelder innerhalb eines Datensatzes. Die konkrete Ausprägung eines einzelnen Datensatzes wird durch die Gesamtheit der Werte der einzelnen Datenfelder repräsentiert. Durch Abhängigkeiten und Verbindungen unterschiedlicher Datensätze zueinander werden komplexe Struktureinheiten (Datenmodelle) gebildet.

Testdaten können als strukturierte Daten (in Datenbanken) sowie auch als unstrukturierte Daten (z. B. Text) vorliegen.

2.3 Testdatenarten

Testdaten können nach unterschiedlichen Kriterien klassifiziert werden. Mögliche Klassifizierungen – nach Testdaten-Verwendungsart und Testdaten-Gewinnungsart werden in den folgenden Abschnitten vorgestellt.

2.3.1 Klassifizierung der Testdaten nach Verwendungsart

- **Eingabedaten:** Die Daten, die bei der Ausführung eines Testfalls an ein Testobjekt übergeben (eingegeben, eingelesen oder auf anderem Weg übertragen) werden.
- **Zustandsdaten:** Die Daten, die den "inneren Zustand" eines Testobjekts enthalten (in Datenbanken, Dateien oder Umsystemen vorhanden sind) und von Testfällen als wiederverwendbarer Startzustand des Testobjekts benötigt/vorausgesetzt werden.
- **Ausgabedaten:** Die Daten, die nach Ausführung eines Testfalls von einem Testobjekt zurückgegeben (ausgegeben, gedruckt, gesendet oder auf anderem Weg übertragen) werden.
- **Soll-Daten:** Die Daten, die vor Ausführung eines Testfalls als erwartetes Ergebnis (Testorakel) bereitgestellt werden.

2.3.2 Klassifizierung der Testdaten nach Gewinnungsart

Testdaten können in die folgenden vier Gewinnungsarten unterteilt werden:

1. **synthetische Testdaten:** Synthetische Testdaten sind Daten, die durch verschiedene Generierungsverfahren erzeugt werden und keinerlei inhaltlichen Bezug zu Produktivdaten aufweisen. Die Erzeugung synthetischer Testdaten bedarf tiefer Kenntnisse des jeweiligen Datenmodells und der sich aus den Geschäftsprozessen ergebenden Datenvarianz. Es muss dabei sichergestellt werden, dass in Abhängigkeit von Teststufe und Testziel die Bandbreite wie auch Verteilung der realen Daten abgedeckt wird. Die Erzeugung synthetischer Testdaten ist nicht nur aus Gründen der Reproduzierbarkeit der Ergebnisse ohne Tool-Unterstützung kaum leistbar. Synthetische Testdaten werden häufig in frühen Teststufen eingesetzt.
2. **Produktivdaten:** Produktivdaten (Synonyme: Produktionsdaten, Wirkdaten, Echtdateien) sind reale, gegebenenfalls schützenswerte und unveränderte Daten, die aus dem produktiven Einsatz eines Systems stammen oder in produktiven Systemen verwendet werden.

3. **anonymisierte Daten:** anonymisierte Testdaten sind Daten, die durch verschiedene Verfahren zur Anonymisierung entstehen, bei denen Produktivdaten manipuliert werden, bis der gewünschte Anonymisierungsgrad erreicht ist. Mit Anonymisierungsgrad wird die Gegenwahrscheinlichkeit zur Identifizierungswahrscheinlichkeit bezeichnet.

Beispiel: Eine Zahl zwischen 1 und 100 soll durch eine Zufallszahl zwischen 1 und 100 ersetzt werden. Die Identifizierungswahrscheinlichkeit beträgt somit 1%, der Anonymisierungsgrad 99%.

Für die Ermittlung des Anonymisierungsgrads müssen Informationen vorliegen, die ein vermeintlicher Angreifer nicht zwangsweise hat.

Man unterscheidet Anonymisierungsgrade von „nicht anonymisiert“ bis hin zur höchstmöglichen Anonymisierung:

- Formale Anonymität: Nur Entfernung der direkten Identifikatoren
 - Faktische Anonymität: Deanonymisierung nicht gänzlich ausgeschlossen. Bei personenbezogenen Daten sind dies pseudonymisierte Daten.
 - Absolute Anonymität: Identifizierung der Auskunftgebenden unmöglich
4. **pseudonymisierte Daten:** „Wie das Anonymisieren dient das Pseudonymisieren dazu, aus Produktivdaten den Personenbezug „auszuschließen oder wesentlich zu erschweren“. Beim Pseudonymisieren gibt es jedoch eine Möglichkeit, die Verknüpfung mit dem echten Namen wiederherzustellen.“¹

Über diese vier Möglichkeiten hinaus gibt es Mischformen aus anonymisierten, pseudonymisierten und synthetischen Testdaten.

Im Testdatenmanagement werden Anonymisierung und Pseudonymisierung weitergefasst als nur auf den Datenschutz bezogen.

Unabhängig davon, ob es sich um personenbezogene oder personenbeziehbare Daten handelt, ist eine Klassifizierung vorzunehmen, je nachdem welcher Anonymisierungsgrad jeweils erreicht wird.

Der englische Fachbegriff für die Anonymisierung bzw. Verfremdung von Daten ist „Data Masking“ bzw. deutsch „Datenmaskierung“. Data Masking bezieht sich nicht allein auf personenbezogene Daten und ist daher weiter gefasst als die reine Anonymisierung und Pseudonymisierung von personenbezogenen oder personenbeziehbaren Daten. Entscheidend ist, dass die Daten nicht verschlüsselt, sondern nur verfremdet, d. h. „maskiert“ werden – ihre Lesbarkeit bleibt in der Regel erhalten.

¹ Quelle: Datenschutz Wiki: http://www.bfdi.bund.de/bfdi_wiki/index.php/3_BDSG_Kommentar_Absatz_6a

2.3.3 Eigenschaften von Testdaten

Testdaten können reale oder realistische Daten sein oder Daten, die je nach angewandter Testmethode einem bestimmten Testziel dienen. Testdaten sind im Allgemeinen realistische Daten, aber keine realen Daten. Es können auch unrealistische Daten erzeugt werden, um Fehlerzustände zu provozieren. Die gespeicherten Testdaten sind unauflöslich mit den Testumgebungen (und den Testzielen der dort auszuführenden Tests) verbunden. Deshalb ist der Umgang mit Testdaten immer (test-)umgebungsabhängig.

2.4 Erzeugung von Testdaten

Die Erzeugung von Testdaten orientiert sich an den zugrundeliegenden Testzielen.

2.4.1 Produktionsabzug

Die Verwendung eines Produktionsabzugs zur direkten Nutzung im Testen ist – sofern sich personenbezogene Daten in dem Produktionsabzug befinden - aus datenschutzrechtlichen Gründen problematisch, aber nicht zwangsweise unzulässig.

Personenbezogene Daten sind Informationen, die eindeutig auf eine natürliche Person schließen lassen, zum Beispiel: Name, Vorname, Telefonnummer oder E-Mail.

Datenschutzaspekte (speziell rechtliche) variieren je nach Bereich und sind von Land zu Land unterschiedlich. Dies stellt häufig bei multinational tätigen Firmen zusätzliche Herausforderungen, speziell seitdem aus Kosteneffizienzgründen die stärkere Integration gekoppelt mit "Offshoring" von Entwicklungs- & Testaktivitäten immer häufiger wird. Aus wirtschaftlichen Gründen (Prinzip der Verhältnismäßigkeit) oder aus Gründen der Testbarkeit (z. B. verschlüsselt vorliegende Daten / Nachstellen von Produktionsfehlern) kann das Testen mit Produktivdaten durchaus legitim und notwendig sein. Darüber hinaus sind bei verschiedenen Produktionskopien immer wieder andere Konstellationen der Daten enthalten, denn Produktionsdaten unterliegen ständiger Veränderung.

Weitere Begriffe (nicht Deckungsgleich): Produktionsklone, Stichtagsabzug

Eigenschaften:

- Produktivdaten enthalten nicht zwangsläufig zu den Testzielen passende Daten, oder diese sind schwer zu finden
- Daten sind technisch von der Produktion entkoppelt (Änderungen in Produktion wirken sich nicht unmittelbar auf den Produktionsabzug aus)

- Daten sind chronologisch von der Produktion entkoppelt (solange die abgezogenen Daten nicht geändert werden, bleibt für diese Daten die Zeit stehen)

Daten können mit der Zeit inkonsistent werden, da Anwendungen (durch zeitliche Auslöser) Änderungen der Daten erwarten, die in einer Testumgebung nicht immer gegeben sind. Wenn in der Produktivumgebung im Tagesgeschäft regelmäßig bestimmte Jobs laufen, die Daten verändern, werden diese in der Testumgebung mitunter nur auf Anfrage ausgelöst.

Mögliches Entgegenwirken durch

- Verwendung synthetischer Daten
- **Datenalterung:** Manipulation von zeitrelevanten Feldern (z. B. durch SQL), um die Konsistenz der Daten zu erhalten, wenn diese von der Produktion entkoppelt vorliegen,
- **Zeitreise / Zeitraffer:** (zeitlich gerafftes) Ausführen von (fachlichen) Anwendungen, um die Konsistenz der Daten zu erhalten. Nicht immer kann man im Test die Fristen für eine Verarbeitung abwarten, die in der Produktion gelten, weil dadurch der Test verzögert und inkonsistente Daten entstehen würden. Hier gilt es, die Systemkonfiguration entsprechend anzupassen.

Beispiel: In einem System werden unverarbeitete Vorgänge nach 99 Tagen gelöscht. Ein Vorgang besteht dabei immer aus zwei Nachrichten (ein Kauf- und ein Verkaufsauftrag) die in mehreren Kriterien übereinstimmen müssen (z. B. Preis, Menge, Produkt, Kauf- und Verkaufszeitpunkt). Geht nur eine Nachricht ein, oder wird keine passende zweite Nachricht gefunden, bleibt diese für 99 Tage im System stehen, bevor sie gelöscht wird. In der Konfiguration des Testsystems wird diese Frist zum Löschen einzelner Nachrichten auf 5 Tage herabgesetzt.

- **Simulation / Spiegelung der Produktion:** (zeitversetztes) Nachfahren der Produktion, um die Produktionsnähe der Daten zu erhalten. Die testrelevanten systemseitigen Prozesse der Produktion sollten in der Testumgebung nachgefahren oder simuliert werden.
- **Setzen von (Test)Umgebungszeit und Datum:** Setzen der (Test)Umgebungszeit und des Datums auf den Zeitpunkt des Produktionsabzugs, um die Konsistenz der Daten zu erhalten. Damit wird immer wieder auf den Anfangsstand der Daten zurückgerollt, wie auf einen „Haltepunkt“.

Beim Produktionsabzug kann unterschieden werden in

- einen Abzug der kompletten Produktivumgebung und
- einen teilweisen Abzug der Produktivdaten.

Ein Teildatenabzug wird beispielsweise bei sehr großen Datenbeständen im Produktivsystem vorgenommen. Eine Herausforderung dabei ist die Beibehaltung der referentiellen Integrität von strukturierten Daten in Datenbanken.

Produktivdaten sind bei nur wenigen Testaktivitäten zur Erreichung von Testzielen notwendig. Zur Vorbereitung von Migrationsprüfungen können Teilabzüge von Produktivdaten jedoch sehr hilfreich sein. Diese können sich beispielsweise auf wenige Stammdatenobjekte erstrecken, bzw. bei gestaffelt geplanten Migrationen entsprechend in mehrere funktionale Migrationsblöcke sequentiell vorgeprüft (verifiziert) werden.

Testdatenbestände, die zum Aufbau einer integrierten Systemkette aus mehreren Anwendungen benötigt werden, müssen mit einem einheitlichen Zeitstempel im Datenabzug versehen sein, um Datenverlust und Dateninkonsistenzen an den Schnittstellen zu vermeiden und soweit notwendig zu reduzieren.

2.4.2 Anonymisierung

Anonymisierung bezeichnet den Prozess der Verfremdung von Produktivdaten zu dem Zweck, die Daten selbst oder ihren Bezug zur Realität unkenntlich zu machen. Im Kontext des Testdatenmanagements dient die Anonymisierung der Bereitstellung von realitätsnahen Testdaten für die Durchführung von Tests, deren zu Grunde liegende Produktivdaten dem Test nicht zur Verfügung stehen. Der Grund hierfür kann der Datenschutz sein, aber auch die Wahrung von Betriebsgeheimnissen.

Eigenschaften:

- Der Prozess der Anonymisierung kann bedarfsgerecht durch verschiedene Anonymisierungsverfahren zu definierten Anonymisierungsgraden erfolgen.
- Ein probates Vorgehen zur Generierung von Testdaten stellt die Anonymisierung eines Produktionsabzuges dar. Hierbei werden personenbezogene und auch unternehmenskritische Daten derart verändert, dass es nicht mehr möglich ist Rückschlüsse auf den Ursprungsdatensatz zu ziehen. Dieses Vorgehen ist bei großen Datenmengen natürlich nur mittels Script- bzw. Toolunterstützung möglich. In komplexen Systemen wird eine Verfremdung zudem durch die notwendige Erhaltung der Datenintegrität erschwert. Zusätzlich müssen bei diesem Vorgehen die Themen Alterung und Verbrauch von Testdaten berücksichtigt werden.

- Der Verbrauch von Testdaten wird vor allem bei selten vorkommenden Datenkonstellationen („Sonderfall“) spürbar, wenn diese Daten in einen Systemzustand versetzt werden, der keine weitere Nutzung für den Test zulässt, weil dieser Zustand nicht oder nur mit großem Aufwand rückgängig gemacht werden kann.
- Wie auch bei der Produktion synthetischer Testdaten steigt bei der Anonymisierung mit zunehmenden Anonymisierungsgrad das Risiko von inkonsistenten Datensätzen in komplexen und/oder verteilten Systemen. Solche Systeme sind schwerer testbar und stellen besonders hohe Anforderungen an das Testdatendesign.
- In einigen Systemen hängen Alterung und Verbrauch der produzierten Testdaten auch von deren Beschaffenheit ab. Dies ist bei der Auswahl der Anonymisierungsverfahren zu berücksichtigen, so dass die anonymisierten Testdaten im System äquivalent zu den zu Grunde liegenden Produktivdaten altern und verbraucht werden.
- Wenn die Anonymisierung von Testdaten jedoch dazu führt, dass grundlegende Eigenschaften von Datenkonstellationen nicht mehr erkennbar sind, können Testdaten wertlos werden, weil Konstellationen nicht mehr erkennbar / auffindbar sind. Deshalb sollte bei der Anonymisierung ein sinnvoller Umfang festgelegt werden, um den Wert von Testdaten nicht zu mindern.

Anonymisierungsverfahren

Generelle Verfahren der Anonymisierung bzw. Pseudonymisierung sind Informationsreduktion und Informationsveränderung. Bei der Informationsreduktion werden keine Inhalte verändert, sondern weggelassen, unterdrückt oder eingeschränkt. Bei der Informationsveränderung werden Inhalte nach vorgegebenen Regeln verändert.

Beispiele für Informationsreduktion sind, den Geburtstag auf das Geburtsjahr zu reduzieren, die Postleitzahl nur mit der ersten Ziffer zu belegen oder Daten mit exponierten Werten wegzulassen, wie den einzigen 97-jährigen Versicherungsnehmer.

Informationsverändernde Verfahren sind zum Beispiel Generalisierung, Maskierung/Substitution, Shuffling/Perturbation oder Varianzbildung.

- Bei der **Generalisierung** wird das Datenelement mit einem weniger spezifischen aber semantisch korrekten Wert ersetzt.

Beispiel: Alle Postleitzahlen im Bereich 10000 bis 19999 werden mit dem Wert 11000 ersetzt.

- Bei einer **Maskierung/Substitution** wird das Datenelement durch ein definiertes Muster ersetzt.

Beispiel: Für die Maskierung wurde definiert, dass nur der erste Buchstabe des Nachnamens erhalten bleiben soll. Die restlichen Zeichen werden durch den Buchstaben „x“ ersetzt. Aus dem Nachnamen „Müller“ würde dann „Mxxxxx“.

- Beim **Shuffling/Perturbation** werden gleiche Datensätze durcheinandergewirbelt, d. h. die enthaltenen Daten werden zwischen den Datensätzen nach einem vorgegebenen Schlüssel getauscht. Der Bezug zu den Ursprungsdaten (z. B. einer Person) muss beim Shuffling aufgelöst werden und eindeutige Informationen (z. B. Telefonnummer, Kreditkartennummer) werden zusätzlich noch mit einer weiteren Methode verfremdet. Dabei ist sicherzustellen, dass ein Datum, welches eigentlich verfremdet werden soll, nicht zufällig wieder eingesetzt wird. Shuffling sollte nur bei einer größeren Menge an Daten angewendet werden.
- Durch **Varianzbildung** werden auf Zahlen basierende Daten so verfremdet, dass diese Zahlenwerte in festgelegten Streuungsintervallen verändert werden. Dabei wird das Streuungsintervall zufällig erhöht oder verringert.

Beispiel: Durch das Varianzintervall [-20%, +20%] für Gehälter wird aus dem Ursprungswert 5.000 € das Testdatum 4.432 € oder aus 6.010 € werden 6.009 €.

Zur **Anonymisierung von Produktionsdaten** werden häufig sogenannte ETL-Tools (Extract, Transform, Load) eingesetzt (siehe Kapitel 6).

2.4.3 Synthetische Erzeugung

Testdaten werden synthetisch erzeugt, wenn keine Produktivdaten verwendet werden. Um dennoch zu realistischen Testdaten zu kommen, können Metadaten als alternative Informationsquellen herangezogen werden. Metadaten sind Daten, die Informationen über Merkmale anderer Daten enthalten, aber nicht diese Daten selbst.

Liegen diese Metainformationen in elektronisch verarbeitbarer Form vor, ist es oftmals möglich, mit Hilfe von Werkzeugen die Testdaten zu definieren, zu spezifizieren und dem Testprozess bereitzustellen. Es handelt sich dann um ein Metadatenmodell für Testdaten.

Eigenschaften:

- Die Qualität der Testdaten hängt entscheidend von den zur Verfügung stehenden Metainformationen ab.

- Die Komplexität des zu testenden Systems bestimmt den Aufwand im Testdatendesign wesentlich. Hier kann eine Abwägung zur alternativen Nutzung von Produktivdaten einen effizienten Weg aufzeigen. Synthetische Testdaten haben jedoch den Vorteil, dass sie genau auf die Anforderungen der Systeme und Schnittstellen herstellbar sind und in der Regel eine höhere Automatisierbarkeit erlauben.
- Bei richtiger Anwendung bestehen keine regulatorischen Hinderungsgründe für den Einsatz synthetischer Testdaten. Auch eine Anonymisierung ist dann nicht erforderlich.
- Synthetische Testdaten werden eingesetzt, wenn entsprechende Daten im Produktivbestand nicht oder noch nicht verfügbar sind oder nur mit hohem Aufwand selektiert werden können.

Synthetisch erzeugte Datenmengen sind in der Regel kleiner als bei den übrigen Produktionsverfahren für Testdaten und haben einen positiven Effekt auf den Testdatenkoeffizienten.

Der **Testdatenkoeffizient** beschreibt das Verhältnis aus vorhandenen zu genutzten Testdaten. Er ist eine Messgröße für die Datenschutzerfordernung nach Datenminimierung.

Beispiel: 100 GB Testdaten sind vorhanden, genutzt werden 20 GB: Der Koeffizient ist 20%.

Mischformen aus anonymisierten und synthetischen Testdaten

In der Praxis hat sich häufig eine Mischform aus anonymisierten Testdaten und synthetischen Testdaten bewährt.

Hierbei werden die Vorteile beider Verfahren kombiniert. So können auf der einen Seite Daten, die dem Verbrauch und der Datenalterung unterliegen, immer wieder neu synthetisch erzeugt werden. Auf der anderen Seite können Daten mit komplexen Abhängigkeiten durch anonymisierte Produktivdaten einmalig bereitgestellt werden.

3. Gesetzliche Regeln und Normen – 90 Min.

Schlüsselbegriffe

personenbezogene Daten, Auditierbarkeit, Konformität, IT-Governance

Lernziele

3.1 Gesetzliche Regeln und Normen

TDM – 3.1.1 (K1) Gesetzliche Regeln und Normen im Zusammenhang mit Testdaten benennen können.

TDM – 3.1.2 (K1) Vorgaben des Datenschutzes, im Umgang mit personenbezogenen und sensiblen Daten im Zusammenhang mit Testdaten, benennen können.

TDM – 3.1.3 (K1) Vorgaben der Datensicherheit, im Umgang mit personenbezogenen und sensiblen Daten im Zusammenhang mit Testdaten, benennen können.

TDM – 3.1.4 (K2) Auditierbarkeit im Zusammenhang mit Testdaten verstehen und beschreiben können.

3.2 Branchenspezifische Regeln (rechtliche Konformität)

TDM – 3.2 (K2) Rechtliche Konformität im Zusammenhang mit Testdaten und die Relevanz für den TDM-Prozess verstehen.

3.3 Unternehmensinterne Regeln (IT-Governance)

TDM – 3.3 (K2) IT-Governance im Zusammenhang mit Testdaten und die Relevanz für den TDM-Prozess verstehen.

3.1 Gesetzliche Regeln und Normen

3.1.1 Gesetze & Normen

Die Regeln für die Verarbeitung personenbezogener Daten sind vorgegeben durch die Datenschutz-Grundverordnung (DSGVO), bereichsspezifische Regeln und durch das Bundesdatenschutzgesetz (BDSG).

Jegliche Verarbeitung personenbezogener Daten bedarf einer ausdrücklichen Erlaubnis, sei es durch eine gesetzliche Regelung oder durch die Einwilligung des Einzelnen. Die gesetzlichen Regeln gelten in Deutschland in folgender Rangfolge:

1. DSGVO
2. Bereichs-spezifisches Datenschutzrecht. Es geht dem BDSG, nicht aber der DSGVO vor. Bereichsspezifische Regeln finden sich beispielsweise
 - in den Büchern des Sozialgesetzbuchs,
 - im BSI-Gesetz,
 - in der Abgabenordnung.
3. Das BDSG gilt ergänzend, wenn es kein bereichs-spezifisches Datenschutzrecht gibt oder wenn die Vorschriften nicht abschließend sind.

In Deutschland gilt die DSGVO vorrangig, denn als europäische Verordnung gilt sie vor dem Recht der europäischen Mitgliedsstaaten. Die Datenschutz-Grundverordnung (DSGVO) ist ein in der Europäischen Union (EU) gültiger einheitlicher Rechtsrahmen. Damit wird eine möglichst einheitliche Anwendung des Rechts für Datenschutz in der EU angestrebt.

Das neue BDSG ist nachrangiges Recht. Seine Regelungen gelten nur dann, soweit die DSGVO nicht gilt. Aber das BDSG gilt auch dann nicht, wenn es andere Rechtsvorschriften des Bundes zum Datenschutz gibt. Diese bereichsspezifischen Rechtsvorschriften haben Vorrang vor dem BDSG.

Die Grundprinzipien des Datenschutzrechts sind:

1. Rechtmäßigkeit, Verarbeitung nach Treu und Glauben, Transparenz:
 - Rechtmäßigkeit der Verarbeitung: Der Besitzer der personenbezogenen Daten hat der Verarbeitung seiner Daten zugestimmt.
 - Verarbeitung nach Treu und Glauben: Daten sollen direkt bei der Person eingeholt werden und nicht von dritter Stelle eingeholt werden
 - Transparenz: garantiert im weiteren Sinne das Recht eines Menschen auf informationelle Selbstbestimmung

2. Zweckbindung: Bereits bei der Erhebung von Daten muss der Zweck der Erhebung offengelegt werden
3. Datenminimierung: so wenig personenbezogene Daten wie möglich erheben, nur für den Zweck benötigte Daten erheben, nicht darüber hinaus
4. Richtigkeit: sachliche richtige Daten, die auch aktuell sind
5. Speicherbegrenzung: sobald die personenbezogenen Daten nicht mehr für den Verarbeitungszweck genutzt werden, müssen sie gelöscht werden
6. Integrität und Vertraulichkeit: eine angemessene Sicherheit und Schutz vor unberechtigtem Zugriff muss gewährleistet sein
7. Rechenschaftspflicht: wer personenbezogene Daten speichert muss die Einhaltung der sechs oben angegebenen Prinzipien nachweisen können.

Die Neuerung im Datenschutzrecht, die mit der DSGVO gültig wurde, ist das Marktortprinzip: Das EU-Datenschutzrecht gilt auch für Wirtschaftsunternehmen außerhalb der EU. Der Anwendungsbereich erstreckt sich damit auch auf Unternehmen außerhalb Europas, die auf dem europäischen Markt tätig sind.

Daten ohne Bezug zu einer Person fallen nicht in den Anwendungsbereich der DSGVO.

Wenn ein Verstoß gegen Datenschutzrecht vorliegt und es zu einem Schaden gekommen ist, ergeben sich daraus Konsequenzen:

Das Datenschutzrecht kann durchgesetzt werden durch Warnungen, Verwarnungen, Anordnungen und Untersagungen. Aufsichtsbehörden können für nahezu alle Verstöße gegen die DSGVO Geldbußen verhängen. Dies kann bis zu 4% des weltweiten Umsatzes eines Jahres für ein Unternehmen betragen.

3.1.2 Datenschutz (rechtliche Konformität)

Der Datenschutz ist ein wichtiges Themengebiet des Testdatenmanagements und verfolgt die Absicht, personenbezogene und sensible Daten vor der unautorisierten Verwendung in Softwaretestvorhaben zu schützen.

- **Problem:** Als Bestandteil des Grundrechtes aus Art. 2 Abs. 1 GG hat jeder Mensch das Recht auf informationelle Selbstbestimmung.
 - Verantwortlich für den ordnungsgemäßen Umgang mit personenbezogenen Daten ist, wer die Daten zu einem bestimmten Zweck erhoben oder in sonstiger Weise genutzt oder verarbeitet hat.

- Da es vorgeblich einfacher und kostengünstiger erscheint, Produktivdaten im Test zu verwenden, werden diese in das Testsystem kopiert. Dabei wird aber die Zweckbestimmung für die Daten verletzt. Laut dem Bundesbeauftragten für den Datenschutz und die Informationsfreiheit ist es in Ausnahmefällen zulässig, da die Zweckgebundenheit zur Nutzung von Daten unter Beachtung besonderer Schutzmaßnahmen auch bei verschiedenen Tests mit Produktivdaten greifen kann (“Der Test mit Produktivdaten muss aber aus Gründen der Datenminimierung und Erforderlichkeit die Ausnahme bleiben. Dabei ist zu beachten, dass nur von der verantwortlichen Stelle – also dem Auftraggeber – die Produktivdaten zu Prüfungszwecken verwendet werden dürfen.”).
- **Lösung:** Kurz- bis mittelfristig kann eine Anonymisierung oder Generierung der Daten das juristische Problem umgehen und vor allem das Risiko eines Datenverlusts minimieren.
 - Verschiedene rechtliche Konstruktionen (Auftragsdatenverarbeitung, Vertraulichkeitserklärungen) können verwendet werden, um juristische Fragestellungen zu präzisieren und so die Nutzung von Produktivdaten durch Dritte oder zu Testzwecken zu ermöglichen.
 - Allerdings sichern diese Konstruktionen nicht das Risiko eines Datenverlustes und damit verbundene Reputationsschäden ab.
 - Eine nachhaltige Lösung beginnt jedoch bei der Konzeption und Planung des Tests und beschäftigt sich frühzeitig mit der Frage: “Welche Daten benötige ich für meinen Test?”.

3.1.3 Datensicherheit

In der ISO 27001, IT-Sicherheitsrichtlinie, findet sich der Punkt „Umgang mit Produktionsdaten“. Auch hier wird der Bedarf deutlich, sich mit Testdatenmanagement zu beschäftigen:

- „A.10.1.4. Werden Entwicklungs- und Testumgebungen von Produktivsystemen getrennt, um das Risiko nicht autorisierter Zugriffe oder versehentliche Änderungen am Produktivsystem zu vermindern?“
- „A.12.4.2. Werden Testdaten sorgfältig ausgewählt geschützt und kontrolliert?“

Datensicherheit beschränkt sich, im Gegensatz zum Datenschutz nicht auf personenbezogene Daten. Die wesentlichen Ziele sind Vertraulichkeit, Integrität und Verfügbarkeit.

3.1.4 Auditierbarkeit

Das Audit eines Prozesses zum Testdatenmanagement betrachtet die qualitativen Eigenschaften des Prozesses an sich. Das Testdatenmanagement unterliegt durch gesetzliche, regulatorische und/oder datenschutzrechtliche Aspekte besonderen Anforderungen an den Prozess, das Rechte- und Rollenmanagement und an die IT-Sicherheit.

Audits der einzusetzenden Verfahren sind in der Startphase komplexer Projekte dringend anzuraten, um spätere "Show-Stopper"-Überraschungen zu vermeiden.

Beispiele für besonders hohe Anforderungen finden sich z. B. in Medizintechnik, Luft- und Raumfahrttechnik, im Bankenbereich oder bei Versicherungen.

Beispiele für sensible Bereiche des Testdatenmanagements im Rahmen einer Auditierung sind:

- Dokumentation über Verwendung und Löschung der Testdaten unter Berücksichtigung von Aufbewahrungsfristen
- Nachweis des Anonymisierungsgrades sowie der verwendeten Anonymisierungsverfahren von Testdaten
- Zugriffsberechtigungen von Personen und deren Rolle/Status im Unternehmen
- Zugriffe auf bestimmte Testdaten durch bestimmte Personen(-kreise)
- Berechtigte Personen müssen Dokumente über Geheimhaltungsvereinbarungen zum Verständnis und Umgang mit Testdaten und Testinfrastruktur vor Einsatz unterschreiben

Ziel der Auditierung ist die Sicherstellung der Einhaltung von Gesetzen und Regularien und dient daher auch der rechtlichen Absicherung bei der Verwendung von Daten im Test. Ohne ein solides Testdatenmanagement mit definiertem Testdatenmanagementprozess, der die Dokumentation der Verwendung von Testdaten reguliert, ist die Einhaltung des Datenschutzes nicht nachweisbar.

3.2 Branchenspezifische Regeln (rechtliche Konformität)

Aufgrund der Vielzahl von branchenindividuellen Regelungen empfiehlt sich die Zusammenarbeit mit einer Branchenfachkraft. Knowhow aus beiden Domänen kann so im Sinne eines branchen- oder unternehmensspezifischen Testdatenmanagements eingebracht werden.

3.3 Unternehmensinterne Regeln (IT-Governance)

Über die vorgenannten externen Regeln existieren auch innerhalb der Unternehmen Vorgaben, die im Testdatenmanagement betrachtet werden müssen. Führung, Organisationsstrukturen und Prozesse sind zu berücksichtigen, um sicherzustellen, dass die Unternehmensstrategie unterstützt wird. Um eine spätere Umsetzung zu vereinfachen, ist es erforderlich, bereits während der Planung und bei Analyse und Design die Verantwortung des Managements und möglichst der Unternehmensführung für das Vorhaben zu erreichen. Möglich wird dies durch Aufzeigen von Risikofaktoren, ausgedrückt in Kennzahlen oder pekuniär.

Zu diesen Risikofaktoren müssen risikobegrenzende Maßnahmen mit Nachweis/ Angabe der autorisierenden Personen/Stellen protokolliert werden.

4. Testdatenmanagement-Prozess – 225 Min.

Schlüsselbegriffe

Testdatenrichtlinie, Testdatenkonzept, Testdatenspezifikation, Testdatenmodellierer, Testdatenrealisierer, Testdatenmanager, Testdatenbereitstellungsroutine, Testdatenbereitstellungsplan, Testdatenbereitstellungsbericht, Testdatenbewertungsbericht, Testdatenabschlussbericht, Testprozess

Lernziele

4.1 Der TDM-Prozess

TDM - 4.1.1 (K1) Die Hauptaktivitäten des TDM-Prozesses wiedergeben können.

TDM - 4.1.2 (K2) Die Hauptaktivitäten des TDM-Prozesses dem Testprozess zuordnen können.

4.2 Aktivitäten im TDM-Prozess

TDM - 4.2.1 (K1) Die Aktivitäten des TDM-Prozesses wiedergeben können.

TDM - 4.2.2 (K2) Die Aktivitäten beschreiben und durchführen können.

4.3 Arbeitsergebnisse im TDM-Prozess

TDM - 4.3.1 (K2) Die Ergebnisartefakte des TDM-Prozesses erläutern können.

TDM - 4.3.2 (K3) Inhalte, Autoren und Zielgruppe(n) den Ergebnisartefakten zuordnen können.

TDM - 4.3.3 (K3) Arbeitsergebnisse abschätzen (Testdatenkonzept, Testdatenspezifikation) können.

4.4 Rollen im TDM-Prozess

TDM - 4.4.1 (K1) Die Rollen im TDM-Prozess benennen können.

TDM - 4.4.2 (K2) Die mit den Rollen verbundenen Aufgaben, Verantwortlichkeiten, Kompetenzstufen und Qualifikationen beschreiben können.

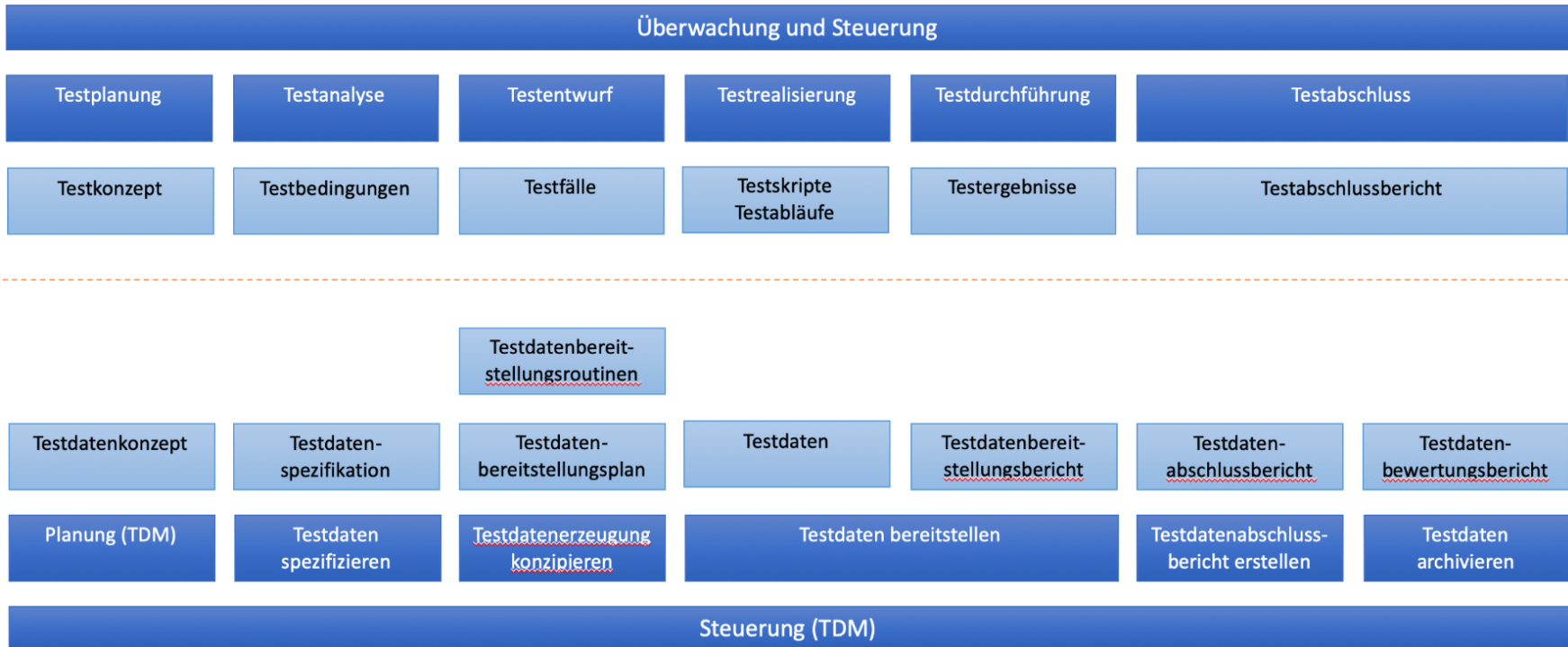
4.1 Der TDM-Prozess

Der Testdatenmanagement-Prozess (TDM-Prozess) orientiert sich am ISTQB®-Testprozess und besteht aus den beschriebenen einzelnen Hauptaktivitäten, die in der Summe das Testdatenmanagement ausmachen.

Diese Hauptaktivitäten können an den Hauptaktivitäten des Testprozesses ausgerichtet werden:

- Planung (TDM)
- Steuerung (TDM)
- Testdaten spezifizieren
- Testdatenerzeugung konzipieren
- Testdaten bereitstellen
- Testdatenabschlussbericht erstellen
- Testdaten archivieren

Der TDM-Prozess ist dadurch mit dem Testprozess verknüpft, wie in der folgenden Abbildung beschrieben.



4.2 Hauptaktivitäten im TDM-Prozess

Obgleich die Aufgaben des TDM-Prozesses in sequenzieller Reihenfolge angegeben sind und die einzelnen Aufgaben logisch aufeinander aufbauen, können und dürfen sie sich in der praktischen Anwendung überschneiden und teilweise auch gleichzeitig durchgeführt werden.

4.2.1 Planung (TDM)

Das Ergebnis der Planung für ein Projekt ist ein **Testdatenkonzept**, als Ergänzung bzw. Anlage zum Testkonzept.

Bei der Planung des TDM werden vorgegebene Verfahren eines Testdatenmanagementsystems an das Projekt angepasst und Ziele, Vorgehensweise, benötigte Ressourcen, Infrastruktur und Zeitplanung festgelegt. Eine Priorisierung der Testdaten findet statt und einzusetzende Werkzeuge werden bestimmt.

Dazu werden die für das TDM relevanten Anforderungen untersucht, um Aspekte zu beschreiben, die im Testdatenkonzept berücksichtigt werden müssen. Anforderungen an das TDM können in folgenden Dokumenten stehen:

- Testrichtlinie und Testdatenrichtlinie des Unternehmens
- Qualitätsziele des Unternehmens und des Projektes
- Testkonzept des Projektes
- Dokumente über Governance und rechtliche Konformität (Datenschutzbestimmungen, gesetzliche Anforderungen, branchenübliche Standards, betriebliche Vorgaben von z. B. Betriebsrat, ISO 9000)
- Dokumentation von Entwicklungsprozess oder Wartungsprozess

Eingangsartefakte

- Testdatenrichtlinie
- Testkonzept

Ergebnisartefakte

- Testdatenkonzept

4.2.2 Steuerung (TDM)

Die Steuerung umfasst die fortwährende Überwachung der aktuell durchgeführten Testdatenaktivitäten im Vergleich zur Planung, die Berichterstattung über die ermittelten Abweichungen und die Durchführung der notwendigen Aktivitäten, um die geplanten Ziele auch unter veränderten Situationen erreichen zu können. Die Aktualisierung der Planung muss auf Grundlage der veränderten Situation erfolgen.

Die Steuerung (TDM) entspricht der Steuerung aus dem ISTQB® Testprozess.

4.2.3 Testdaten spezifizieren

In dieser Hauptaktivität werden die Anforderungen an die zu erstellenden Testdaten analysiert und die betroffenen Datenquellen/Datensenzen untersucht.

Wichtigste Aktivität der Spezifikation ist die Festlegung des Designs der Testdaten in den notwendigen bzw. geforderten Ausprägungen.

Zudem sind unter anderem folgende Tätigkeiten durchzuführen:

- Planen des Einsatzzweckes (z. B. Zustandsdaten wie Währungen, Länder, Postleitzahlen) und Konfiguration des Testsystems
- Analysieren, ob Testdaten nach Testdurchführung verbraucht sind und Maßnahmen zur Korrektur festlegen.
- Analysieren, wie zeitliche Aspekte die Testdaten beeinflussen
- Festlegen der Bereitstellungsstrategie (z. B. Voll- oder Teilabzug, mit/ohne Anonymisierung, mit/ohne synthetische Anreicherung)
- Festlegen der benötigten Datenmengen
- Festlegen des Bereitstellungszeitpunktes
- Festlegen der Zielformate (XML, SQL, Messagequeue, etc.)
- Festlegen der Verantwortlichkeiten und Zugriffsrechte für die Testdaten
- Prüfen der rechtlichen Zulässigkeit und Freigabe

Die Ergebnisse dieser Aktivitäten werden in der Testdatenspezifikation dokumentiert.

Testdaten spezifizieren wird während der Testanalyse durchgeführt.

Eingangsartefakte:

- Anforderungsspezifikation

- Testfälle
- Testdatenkonzept

Ergebnisartefakte:

- Testdatenspezifikation

4.2.4 Testdatenerzeugung konzipieren

In dieser Hauptaktivität werden auf Basis der Testdatenspezifikation die Erzeugung der Testdaten konzipiert sowie Werkzeuge und Routinen (Skripte/Jobs) zur Bereitstellung und Erzeugung von Testdaten realisiert und geprüft.

Je nach Ergebnis der Analyse wird für die verschiedenen Datenquellen und -ziele die passende Erzeugungsstrategie (siehe Abschnitt 2.4) festgelegt. Im Testdatenbereitstellungsplan wird festgelegt, wann und von wem welche Testdaten mit den realisierten Bereitstellungsroutinen erzeugt werden.

Bereitstellungsroutinen müssen auf fachliche Korrektheit sowie regulatorische Zulässigkeit geprüft und freigegeben werden. Zudem müssen sie selbst funktional und nichtfunktional getestet werden.

Testdatenerzeugung konzipieren erfolgt während des Testentwurfs.

Eingangsartefakte:

- Testdatenkonzept
- Testdatenspezifikation

Ergebnisartefakte:

- Testdatenbereitstellungsroutinen
- Testdatenbereitstellungsplan

4.2.5 Testdaten bereitstellen

In der Hauptaktivität "Testdaten bereitstellen" werden die notwendigen Schritte zur Bereitstellung der erzeugten Testdaten festgelegt und ausgeführt.

In der Regel wird hierbei das Deployment vorhandener Daten in die Zielsysteme sowie die Aktualisierung von Zeitstempeln bzw. die Durchführung von „Zeitreisen“ für die betroffenen Testdaten oder auch die Umwandlung der vorhandenen Daten in andere Datenformate durchgeführt.

Erzeugte Testdaten müssen auf rechtliche Zulässigkeit geprüft und freigegeben werden. Im Testdatenbereitstellungsbericht werden die Ergebnisse dokumentiert.

Testdaten bereitstellen erfolgt während der Testrealisierung oder während der Testdurchführung.

Eingangsartefakte:

- Testfälle
- Testdatenbereitstellungsroutinen
- Testdatenbereitstellungsplan

Ergebnisartefakte:

- Testdaten
- Testdatenbereitstellungsbericht

4.2.6 Testdatenabschlussbericht erstellen

Während des Abschlusses der Testaktivitäten werden die gesammelten Daten und Informationen aus den abgeschlossenen Aktivitäten verwendet, um die Erfahrungen, Testmittel, Fakten und Zahlen zu konsolidieren.

Der Abschluss der Testaktivitäten umfasst die Konsolidierung der Testmittel und die Bewertung des Testdatenmanagement-Prozesses im Testdatenabschlussbericht. Der Testdatenabschlussbericht kann auch in den Testabschlussbericht integriert werden.

Eingangsartefakte:

- Testdatenbereitstellungsbericht
- Testprotokolle
- Fehlerberichte

Ergebnisartefakte:

- Testdatenabschlussbericht

4.2.7 Testdaten archivieren

Analog zur Testdatendokumentation ist die Testdatenarchivierung ein grundlegender Bestandteil des Testdatenmanagements hinsichtlich der wiederholten und nachvollziehbaren Ausführung zurückliegender Testszenarien, um z. B. in der

Vergangenheit durchgeführte Tests unter exakt denselben Bedingungen wiederholbar zu machen, etwa aufgrund gesetzlicher Anforderungen.

Der Abschluss der Testdatenaktivitäten umfasst

- das Archivieren der Testdatenmittel,
- die Bereinigung der Testumgebungen (Löschen nicht zu archivierender Testdatenmittel) und
- das Löschen nicht mehr benötigter Testdatenmittel (abgelaufene Archive).

Eingangsartefakte:

- Alle im Prozess erzeugten Testdatenartefakte
- Archiv

Ergebnisartefakte:

- Testdatenbewertungsbericht
- Archiv

4.3 Arbeitsergebnisse im TDM-Prozess

4.3.1 Testdatenrichtlinie

Eine Testdatenrichtlinie ist ein Dokument, das organisationsweit, projektübergreifend gültig ist und entspricht dem Testdatenkonzept auf einem umfassenderen, evtl. abstrakteren Level.

Die Testdatenrichtlinie entspricht der Testrichtlinie im ISTQB® Testprozess.

- **Zielgruppe:** Projektleitung, Test(daten)management und Stakeholder
- **Ersteller:** Testdatenmanager

4.3.2 Testdatenkonzept

Im Testdatenkonzept werden die Prozesse des Testdatenmanagements und die Schnittstellen zu anderen Prozessen, z. B. dem Anforderungs- und dem Testmanagement, für ein konkretes Projekt beschrieben.

Das Testdatenkonzept ist ein Dokument, in dem u. a. der Gültigkeitsbereich, die Vorgehensweise, die Ressourcen und die Zeitplanung der beabsichtigten Testdatentätigkeiten mit allen Aktivitäten für ein Projekt beschrieben werden.

Ein Testdatenkonzept muss unter anderem folgende Punkte enthalten:

- Anforderungen an die Testdateninfrastruktur,
- Regelung der Versionierung von Testdaten und Testdatenanforderungen,
- Angaben zur Löschung (Komplett-, Teillöschung) der Testdaten,
- Wiederherstellung von Testdaten (erneutes Bereitstellen eines Testdatenzustandes),
- Festlegen der Archivierungsregeln (Skripte, Testdaten, Prozessbeschreibungen, Vorhaltdauer der Archivbestände),
- Einzuhaltende rechtliche Anforderungen und Standards,
- Benötigte Testumgebungen,
- Testdatenbereitstellungsverfahren (mit Begründung der Auswahl) für jede im Testprozess spezifizierte Teststufe / Testart,
- Einzusetzende Werkzeuge,
- Festlegung der Verantwortlichkeiten,
- Dokumentation der Testdatenverwendung/Bereitstellung,
- Risiken, inkl. einer Planung für den Fall ihres Eintretens,
- Aufbewahrungsfristen und
- Festlegung geeigneter Metriken (s.u.).

Im Testdatenkonzept werden geeignete Metriken zur Bewertung des Prozesses festgelegt:

- Metriken zur Qualität der Testdaten
- Testfallbasierte Metriken
- Kostenbasierte Metriken
- Fehlerbasierte Metriken

Das Testdatenkonzept entspricht dem Testkonzept im ISTQB® Testprozess.

- **Zielgruppe:** Projektmanagement, Test(daten)management, Testdatenmodellierer, Testdatenrealisierer.
- **Ersteller:** Testdatenmanager

4.3.3 Testdatenspezifikation

In einer Testdatenspezifikation werden Ein- oder Ausgabewerte für ein Testobjekt formal beschrieben, sowie die Herkunft der Testdaten, die Qualität und die Menge vorgegeben. Einflussparameter sind das Testobjekt, Anforderungen an die Testdaten, Teststufe, Testteam, Testziel.

Ein Projekt kann mehrere Testdatenspezifikationen benötigen. Eine Testdatenspezifikation referenziert immer ein Testdatenkonzept.

In einer Testdatenspezifikation sollten folgende Fragen beantwortet werden, damit das definierte Testziel erreicht werden kann:

- Wie werden welche Testdaten bereitgestellt?
- Gibt es Abhängigkeiten von Testfällen und wie werden diese gehandhabt?
- Wer stellt die Testdaten mit welchem Werkzeug in welchem Format bereit?
- Wie wird sichergestellt, dass die Tester isoliert und ungestört testen können?
- Sind die Daten nach dem Test verbrannt und wie können sie wieder aufgefrischt werden?
- Sind Zeitreisen notwendig und wie können sie realisiert werden?
- Können Testdaten altern und „versterben“? Wie sind die Testdatenkonstellationen in diesen Fällen wiederherstellbar?
- **Zielgruppe:** Testdatenrealisierer, Tester
- **Ersteller:** Testdatenmodellierer

4.3.4 Testdatenbereitstellungsroutinen

Testdatenbereitstellungsroutinen bestehen aus Software (Tools, Skripte, Jobs, ...) und manuellen Prozessen sowie deren Dokumentation. Die Testdatenbereitstellungsroutinen müssen ausgeführt werden, um konkrete Testdaten für die Durchführung der Testfälle zu erzeugen und um Testdaten und Testumgebungen in einen definierten Zustand zu versetzen. Die Testdatenbereitstellungsroutinen setzen die Anforderungen der Testdatenspezifikation um und müssen diesbezüglich getestet und freigegeben sein.

- **Zielgruppe:** Testdatenrealisierer, Tester
- **Ersteller:** Testdatenrealisierer

4.3.5 Testdatenbereitstellungsplan

Im Testdatenbereitstellungsplan ist festgelegt, wann welche Testdatenbereitstellungsroutinen von wem ausgeführt werden müssen.

Der Testdatenbereitstellungsplan ist in der Regel Teil des Testausführungsplans.

- **Zielgruppe:** Testdatenrealisierer, Testdatenmanager, Tester
- **Ersteller:** Testdatenrealisierer, Testdatenmanager

4.3.6 Testdaten

Die eigentlichen Testdaten entsprechend der Definition in Abschnitt 2.2.2

4.3.7 Testdatenbereitstellungsbericht

Der Testdatenbereitstellungsbericht dient der Mitteilung einer erfolgten Testdatenbereitstellung und definiert, welche Freigaben für die Daten vorliegen. Dies können beispielsweise Einschränkungen auf bestimmte Personenkreise sein.

- **Zielgruppe:** Testdatenmanager, Testdatenrealisierer, Testdatenmodellierer
- **Ersteller:** Testdatenmanager

4.3.8 Testdatenabschlussbericht

Der Testdatenabschlussbericht fasst die Testdatenaktivitäten und -ergebnisse zusammen. Er enthält eine Bewertung der durchgeführten Tests gegen definierte Endkriterien im Hinblick auf Testdaten.

- Abweichungen: Darstellung der fehlerhaft bereitgestellten Testdaten
- Messen der Qualität der Testdaten mit geeigneten Metriken (siehe Abschnitt 4.3.2) oder Referenz auf den Testdatenbewertungsbericht
- Zusammenfassung der wichtigsten Testdatenaktivitäten
- Bewertung der benötigten Ressourcen, Personal, Zeitbedarf
- Darstellung der identifizierten und eingetretenen Projekt- und Produktrisiken

Der Testdatenabschlussbericht kann in den Testabschlussbericht integriert werden.

- **Zielgruppe:** Projektleitung, Testmanagement und Stakeholder
- **Ersteller:** Testdatenmanager

4.3.9 Testdatenbewertungsbericht

Der Testdatenbewertungsbericht enthält die Bewertung der Qualität der bereitgestellten Testdaten.

Dieser kann in den Testabschlussbericht integriert werden.

- **Zielgruppe:** Testdatenmanager (alle Kompetenzstufen)
- **Ersteller:** Testdatenmanager (Projekt)

4.4 Rollen im TDM-Prozess

Im Testdatenmanagement werden drei Rollen definiert:

- Der Testdatenmanager plant und steuert den Prozess.
- Der Testdatenmodellierer analysiert die Testdatenanforderungen und spezifiziert die Testdaten, er benötigt fachliches und methodisches Wissen.
- Der Testdatenrealisierer ist technisch orientiert (Programmierung, Toolspezialist) und stellt bzw. erzeugt die konkreten Testdaten.

Die Rollen orientieren sich an den Hauptaktivitäten des Testdatenmanagementprozesses. Sie können in der Realität in Personalunion besetzt werden. Sofern kein eigenständiges Testdatenmanagement im Projekt implementiert ist, können Rollen vom Testmanager und vom Tester bzw. Entwickler übernommen werden.

4.4.1 Testdatenmanager

Der Testdatenmanager ist für die Planung und Steuerung des TDM-Prozesses verantwortlich, um die im Testprozess benötigten Testdaten zeitgerecht und in der geforderten Qualität bereitzustellen.

Der Testdatenmanager sollte mehrjährige Testerfahrung besitzen und den Test(datenmanagement)-Prozess gut kennen. Er benötigt Kenntnisse über die Methoden und Werkzeuge zur Testdatenbeschaffung und -erzeugung sowie über regulatorische Anforderungen.

Der Begriff Testdatenmanager kann in kleinen Projekten als Oberbegriff für alle Rollen im Testdatenmanagement verwendet werden.

Aufgaben:

- TDM-Prozess planen und steuern
- Über den Status des TDM-Prozesses berichten
- Entscheidungen zum Einsatz von TDM-Werkzeugen treffen/unterstützen
- Testdatenrichtlinie (nicht projektspezifisch) erstellen
- Testdatenkonzept (projektspezifisch) erstellen
- Testdatenabschlussbericht erstellen und freigeben
- Abstimmung mit Beteiligten (z. B. Datenschutzbeauftragten, Rechenzentrum, anderen Testdatenmanagern, Fachbereichen, ...)
- Siehe auch Aktivitäten in den Prozessen „Planung (TDM)“ und „Steuerung (TDM)“ im TDM-Prozess

Verantwortlichkeit/Befugnis:

- Einhaltung der regulatorischen Vorgaben
- Eskalation an Testmanager
- Termingerechte Bereitstellung der Testdaten in der vereinbarten Menge und Qualität

Kompetenzstufe:

Die Rolle Testdatenmanager kann unterschiedlich ausgestaltet werden:

- Projektebene
- Domänenspezifisch
- Organisationsebene

Qualifikation:

- Erfahrung in IT-Projekten
- Gute Kenntnisse von Standards in Testprozessen (z. B. ISTQB®, TMAP®, ITIL®, ...)
- Gute Kenntnisse von Datenhaltungssystemen
- Vertiefte Kenntnisse in TDM- und integrierten TM-Werkzeugen
- Kenntnisse von Methoden zur Bereitstellung und Erzeugung von Testdaten
- Vertiefte Kenntnisse von Konfigurationsmanagement-Werkzeugen

4.4.2 Testdatenmodellierer

Der Testdatenmodellierer analysiert auf Basis des Testdatenkonzepts die Anforderungen, Testfälle, Systemarchitekturen und Datenmodelle mit dem Ziel, den Testdatenbedarf zu identifizieren. Dabei muss er bereits die regulatorischen Anforderungen mit in Betracht ziehen, soweit sie die bereitzustellenden Testdaten betreffen.

Aufgaben:

- Testfälle und Testdaten-Anforderungen analysieren und prüfen
- Testobjekte, Testumgebungen und Systemumgebungen analysieren
- Testdatenspezifikation erstellen unter Berücksichtigung von regulatorischen Anforderungen (z. B. Datenschutz)
- Ggfs. erstellen von Musterdaten (Datenbeispiele, welche die benötigten Testdaten in ihrer Struktur und ihrem Format veranschaulichen, z. B. in Bezug auf Äquivalenzklassen, Grenzwerte oder Anwendungsfälle)
- Empfehlungen an Testdatenmanager geben
- Im Detail siehe Aktivitäten im Prozess „Testdaten spezifizieren“ im TDM-Prozess

Verantwortlichkeit/Befugnis:

Fachliche Spezifikation der Testdaten (Inhalt, Menge, Qualität, Quelle, Komplexität)

Kompetenzstufe:

- Projektebene

Qualifikationen:

- ISTQB® Qualifikation (Certified Tester Foundation Level)
- Kenntnisse über IT-Architekturen und Datenstrukturen
- Tiefe Kenntnisse in den Datenstrukturen des Testobjektes
- Gute Kenntnisse von Datenhaltungssystemen
- Vertiefte Kenntnisse von Methoden zur Erzeugung und Bereitstellung von Testdaten
- Vertiefte Kenntnisse in TDM- und integrierten TM-Werkzeugen
- Umfassende Kenntnisse regulatorischer Anforderungen (z. B. Datenschutz)

4.4.3 Testdatenrealisierer

Der Testdatenrealisierer ist für die Erzeugung und Bereitstellung der Testdaten zuständig. Dabei erfolgt die Erzeugung innerhalb der festgelegten Testdaten-Strategie und anhand der vom Testdatenmodellierer aufgestellten Testdatenspezifikation. Ihm obliegt die konkrete Durchführung der TDM-Aktivitäten, u. a. Testdaten erzeugen, bereitstellen, vergleichen, dokumentieren und archivieren.

Diese Rolle kann unterschiedliche Schwerpunkte haben, wie z. B.:

- Automatisierung der Erzeugung synthetischer Testdaten
- Bereitstellung von ggf. bearbeiteten Produktionsdaten (Kopieren, Extrahieren, Transformieren, [Anonymisieren, Pseudonymisieren], Laden)

Aufgaben:

Zu den Aufgaben gehören u. a.:

- Bereitstellung von Testdatenkonfigurationen
- Bereitstellung von Testdaten entsprechend der Spezifikation
- Generierung, Anonymisierung, Pseudonymisierung von Daten
- Anpassung und Pflege von Testdaten
- Dokumentieren und Archivieren von Testdaten und der Bereitstellung
- Automatisieren der Bereitstellung von Testdaten
- Einspielen von Testdaten in Testumgebungen
- Bedienung von TDM-Werkzeugen, Versionsverwaltung und Archivierung
- Testdatenbereitstellungsroutinen erstellen
- Testdatenbereitstellungsplan erstellen

Der Testdatenrealisierer führt die drei folgenden Teilprozesse des Testdatenmanagementprozesses aus, siehe Prozessbeschreibungen im TDM-Prozess:

- Testdatenbereitstellung vorbereiten
- Testdaten bereitstellen
- Testdaten archivieren

Verantwortlichkeit/Befugnis:

Technische Umsetzung der Testdatenspezifikation

Kompetenzstufe:

Projektebene

Qualifikationen:

- Tiefe Kenntnisse in den Datenstrukturen des Testobjektes
- Bedienung der Schnittstellen zu Testmanagementsystemen
- Werkzeugkenntnisse
- Programmierkenntnisse
- Vertiefte Kenntnisse in Versionsverwaltungs- und Archivierungssystemen

5. Einbindung in die Organisation – 60 Min.

Schlüsselbegriffe

Testmanagement, Testumgebungsmanagement, Konfigurationsmanagement, Revisionsicherheit

Lernziele

5. Einbindung in die Organisation

- TDM – 5.1 (K2) Die Einbindung des Testdatenmanagements in die Organisation beschreiben.
- TDM – 5.2 (K2) Die Verantwortlichkeiten für die Einbindung des Testdatenmanagements in die Organisation beschreiben können.
- TDM – 5.3 (K2) Die Einbindung des Testdatenmanagements in die Prozesse beschreiben.
- TDM – 5.4 (K2) Die Einbindung des Testdatenmanagements in die IT beschreiben.
- TDM – 5.5 (K2) Die Einführung von TDM in Organisation und IT beschreiben.

5.1 Einbindung in die Organisation

TDM kann nur mit guter Einbindung in die Unternehmens- und Projektorganisation funktionieren. Der TDM-Prozess muss in die Organisation passen, dort eingeführt, regelmäßig angepasst und kommuniziert werden.

In Projekten ist das Testdatenmanagement nur im Zusammenhang mit anderen Prozessen effektiv. Testdatenmanagement soll die existierenden Prozesse im Projektmanagement, in der Softwareentwicklung und im Test unterstützen und verbessern, kann also per Definition nur im größeren Zusammenhang betrachtet werden.

5.2 Verantwortlichkeiten

Die Verantwortlichkeiten im TDM beginnen beim oberen Management, das ein organisations- und projektübergreifendes TDM beauftragt und finanziert. Das Management etabliert einen Testdatenmanager, der für die Erstellung und Umsetzung einer allgemeingültigen Testdatenrichtlinie verantwortlich ist.

In einem konkreten Projekt sind Testdatenmanager, Testdatenmodellierer und Testdatenrealisierer im Rahmen ihrer Rollen (siehe Abschnitt 4.4) für die jeweiligen Ergebnisse ihrer Aktivitäten verantwortlich.

Die konkreten, zum Test benötigten Testdaten müssen geprüft und freigegeben werden. Hierfür sind die Verantwortlichen zu benennen; das können Datenschutzbeauftragte, Betriebsräte, Personalräte und/oder Personen aus dem Fachbereich sein.

5.3 Einbindung in die Prozesse

Testmanagement (TM) ist eng verzahnt mit dem Projektmanagement (PM). Das Testdatenmanagement ergänzt den Testprozess (siehe Abschnitt 4.1) und muss in diese Prozesse der Softwareentwicklung eingebunden werden.

Testdaten werden in unterschiedlichen Ausprägungen und Mengen für verschiedene Testumgebungen benötigt, spielen daher im Testumgebungsmanagement eine große Rolle.

Testdaten werden für spezielle Systemkonfigurationen bereitgestellt und müssen daher auch im Konfigurationsmanagement (KM) berücksichtigt werden. Änderungen in der Konfiguration, zum Beispiel durch umgesetzte Change Requests, sollten sich in den Testdaten widerspiegeln. Die Arbeitsergebnisse des Testdatenmanagements müssen versioniert und in das KM-System übernommen werden.

Dazu gehören die Testdatenspezifikationen, Bereitstellungsrountinen, beschreibende Dokumente und nicht zuletzt die Testdaten selbst.

5.4 Einbindung in die IT

Testdaten werden generiert, pseudonymisiert bzw. anonymisiert, kopiert, geladen, gesichert, zurückgesetzt, validiert und/oder manuell bearbeitet. Für alle diese Tätigkeiten werden Tools eingesetzt. Diese Werkzeuge müssen ausgewählt, angeschafft, installiert, geschult und betrieben werden. Schnittstellen müssen in der Regel für diese Tools angepasst oder realisiert werden. Das bedeutet, dass diese Werkzeuge in die IT-Landschaft eines Unternehmens passen und dort integriert werden müssen.

Ein weiterer Aspekt für die IT eines Unternehmens ist die Verwaltung der Testdatenbestände und die Bewirtschaftung der Testumgebung. Die Testdaten müssen erzeugt und gesichert werden, was mit sehr großem Kapazitätsbedarf einhergehen kann. Der Speicherplatzbedarf muss für Gegenwart und Zukunft berechnet und bereitgestellt werden.

Die Erstellung von Kopien und die Pseudonymisierung bzw. Anonymisierung großer Testdatenbestände können nur mit Einplanung und Mitarbeit der betreffenden Stellen in der IT erfolgen.

5.5 Einführung des TDM in die Organisation

Der Testdatenmanagementprozess muss in der Organisation eingeführt werden. Dazu sollten projektunabhängig die generellen Ziele und Anforderungen an das Testdatenmanagement in einer Testdatenrichtlinie festgelegt werden. Dann sollte mit dem TDM ganz konkret in einem oder in zwei Projekten begonnen werden. Anhand der erarbeiteten Arbeitsergebnisse (siehe Abschnitt 4.3) und der gesammelten Erkenntnisse können die Anforderungen in die allgemeingültige Testdatenrichtlinie einfließen.

Die Testdatenrichtlinie, Prozessbeschreibungen, Dokumentenvorlagen und Werkzeugeinsätze müssen durch die Verantwortlichen (siehe Abschnitt 5.2) permanent überprüft und verbessert werden, mit dem Ziel ein übergreifendes Testdatenmanagement zu schaffen, das die Entwicklungs- und Testteams effektiv, revisionssicher und regelkonform bei der Testdatenerstellung und somit beim Testen unterstützt.

Dabei umfasst die Revisionssicherheit weniger die technischen Aspekte, sondern eher den organisatorischen Prozess zur Archivierung. Dazu zählen unter anderem die Dauer der Archivierung, der Ablauf zur Archivierung und die Vollständigkeit der Daten.

Auch die Nachvollziehbarkeit, die Dokumentation des Verfahrens und die Zugriffskontrolle sowie der Schutz vor rückwirkender Veränderung der Daten spielen eine wichtige Rolle.

Die Anforderungen an die Revisionssicherheit ergeben sich aus dem HGB der Abgabenordnung (AO) und detailliert aus den Grundsätzen zur ordnungsgemäßen Führung und Aufbewahrung von Büchern, Aufzeichnungen und Unterlagen in elektronischer Form sowie zum Datenzugriff GoBD.

6. Werkzeuge – 45 Min.

Schlüsselbegriffe

Data Profiling Werkzeug, Testentwurfswerkzeug, Testdateneditor und -generator, ETL-Werkzeug, Testausführungswerkzeug

Lernziele

6.1 Typen von Werkzeugen

TDM - 6.1 (K2) Testdatenwerkzeugtypen nach ihrem Zweck und Einsatz im Testdatenmanagementprozess beschreiben.

6.2 Werkzeuge im TDM-Prozess

TDM - 6.2.1 (K2) Aufzeigen von Einsatzszenarien für TDM-spezifische Werkzeugtypen.

TDM - 6.2.2 (K2) Aufzeigen von Einsatzszenarien zur Wiederverwendung vorhandener Werkzeugtypen für TDM.

6.3 Besonderheiten im Zusammenhang mit Testautomatisierung

TDM - 6.3 (K2) Besonderheiten des Testdatenmanagements im Zusammenhang mit der Testautomatisierung begründen können.

6.1 Typen von Werkzeugen

Die im Syllabus zum ISTQB® Foundation Level [CTFL] beschriebenen Grundlagen zu Typen von Testwerkzeugen, deren effektive Anwendung und deren Einführung in eine Organisation sind ebenso relevant für einen Einsatz von Werkzeugen im Testdatenmanagement. Darüber hinaus gibt es spezifische Anforderungen im Testdatenmanagement, die durch weitere Typen von Werkzeugen abgedeckt werden können.

Am Markt sind einerseits spezialisierte Werkzeuge verfügbar, die das Suchen, Finden und Validieren von Daten unterstützen, die eine Datenmigration, -extraktion und -manipulation ermöglichen, oder Daten anonymisieren bzw. synthetisch erzeugen können.

Andererseits bieten Testmanagement-Werkzeuge Unterstützung in einigen Bereichen des Testdatenmanagements, wie beispielsweise der Verwaltung von Testdatenpools und der Dokumentation von Bereitstellungsprozessen.

Wenn Rahmenbedingungen gegen einen Kauf von Werkzeugen sprechen oder spezifische Anforderungen durch diese nicht abgedeckt werden, bietet sich ggfs. eine eigene Entwicklung von Skripten und Datentabellen an.

Im Folgenden sind einige Definitionen von Werkzeugtypen aus dem Syllabus ISTQB®-CTFL übernommen, diese sind als Zitate gekennzeichnet.

6.2 Werkzeuge im TDM-Prozess

Testdatenwerkzeuge können eine oder mehrere Aktivitäten des Testdatenmanagement-Prozesses unterstützen.

6.2.1 Planung und Steuerung

Für die Planung und Steuerung des Testdatenmanagements bieten sich die Werkzeuge des Testmanagements an, da die Aufgaben in engem Bezug zueinanderstehen und in weiten Teilen deckungsgleich sind.

6.2.2 Testdaten spezifizieren

Bei der Spezifizierung von Testdaten können **Data Profiling Werkzeuge** nützlich sein. Sie bieten Unterstützung bei der Ermittlung von Datenstrukturen, Inhalten und Beziehungen/Abhängigkeiten.

„**Data Quality**“- und „**Data Explorer**“-Werkzeuge zielen ebenfalls darauf ab, Transparenz über die vorhandenen Datenstrukturen und den vorhandenen Datenbestand zu gewinnen.

„Testentwurfswerkzeuge

Diese Werkzeuge werden verwendet, um Testeingaben, ausführbare Tests und/oder Testorakel aus den Anforderungen, der graphischen Benutzungsschnittstelle (GUI), dem Entwurfsmodell (Zustands-, Daten- oder Objektmodell) oder aus dem Code zu generieren.“

In vielen Fällen ist es möglich, die Testdaten mit diesen Werkzeugen in einer für die Erzeugung und Bereitstellung erforderlichen Detaillierung zu beschreiben.

Mit **Konfigurationsmanagementwerkzeugen** können weitere Beschreibungen (Metadaten) der Testdaten hinterlegt werden. Beispielsweise eine Klassifizierung nach Datenschutzzklassen, Abhängigkeiten zu anderen Testdaten oder Informationen zur Temporalität (Behandlung des Zeitbezugs).

Einige **Testdateneditoren und -generatoren** unterstützen diese Funktionalitäten ebenfalls.

6.2.3 Testdatenerzeugung konzipieren

Wesentliche Anforderungen bei der Konzeption der Testdatenerzeugung sind:

- Erstellen des Testdatenbereitstellungsplans (Zusammenstellung von gemeinsam bereitzustellenden Testdaten, bei Anonymisierung/Pseudonymisierung: Definition des Anonymisierungsprozesses sowie der Lade- und Entladeprozesse)
- Erzeugen von konzeptionellen Bereitstellungsroutinen (Transferrountinen, Synthetische Erzeugungsroutinen, Anonymisierungs-/Pseudonymisierungsregeln)
- Erzeugen eines Testdatenpools
- Automatisierung der Bereitstellung (Ablaufplanung, Cron-Jobs, usw.)
- Dokumentation der notwendigen Bereitstellungsschritte (Automatisch oder Manuell)
- Versionierung der erzeugten Mittel (Routinen, Testdaten und Dokumentationen)

Testdateneditoren und -generatoren und **ETL-Werkzeuge** (Extrahieren, Transformieren, Laden) unterstützen die Umsetzung dieser Anforderungen.

„Mit Hilfe von Testdateneditoren und -generatoren können aus Datenbanken, Dateien oder Datenströmen zunächst Testdaten ermittelt und dann sämtliche für einen Test benötigte Testdaten bearbeitet werden, um den Datenschutz durch Anonymisierung sicherzustellen.“

Für einige Teilbereiche der Testdatenerzeugung sind verschiedene spezialisierte Werkzeuge am Markt verfügbar:

- Werkzeuge für das Auslesen, Manipulieren und Laden von Datenbeständen, oder Teilen davon (Migrationswerkzeuge, „Data Subsetting“, ETL-Werkzeuge).
- Werkzeuge für das Erkennen von kritischen Feldern/Spalten im Sinne des Datenschutzes, oder anderer regulatorischer Anforderungen („Data Quality“- und „Data Explorer“, Maskierungswerkzeuge).
- Werkzeuge für eine regelbasierte Manipulation von Daten (Maskierungs- und Anonymisierungswerkzeuge).

6.2.4 Testdaten bereitstellen

Wesentliche Anforderungen im Bereitstellungsprozess sind:

- Ausführen der Testdatenbereitstellungsroutinen gemäß Testdatenbereitstellungsplan (Bedienung von Werkzeugen, Implementierung von Regeln, etc.)
- Verifikation und Qualitätssicherung der Testdatenbereitstellung
- Nachverfolgung der Testdaten pflegen (Verwendung, Reservierung, Zeitpunkte, Testfälle, Testergebnisse)
- Bericht der Testdatenbereitstellung und Freigabe der Testdaten (Testdatenbereitstellungsbericht)
- Interaktion mit den Testausführungswerkzeugen
- Interaktion mit Vergleichswerkzeugen

In der Praxis sind die Erzeugung und Bereitstellung oft in einem Schritt zusammengefasst und die genannten Werkzeugarten unterstützen beide Aktivitäten.

6.2.5 Testdatenabschlussbericht erstellen

Da der Testdatenmanagement-Prozess als Sub-Prozess des Testmanagementprozesses eng mit diesem verzahnt ist, sollten die Werkzeuge eine Konsolidierung der Berichte (Testdatenbereitstellungs- und Abschlussbericht) im Testmanagement unterstützen.

6.2.6 Testdaten archivieren

Für die Versionierung und Archivierung der Artefakte bieten sich die Konfigurationsmanagementwerkzeuge an, die üblicherweise im Projekt verwendet werden.

6.3 Besondere Betrachtung von Testdaten für die Testautomatisierung

Eine der wesentlichen Anforderungen an automatisierte Testläufe stellt deren Stabilität dar. Sie finden häufig unbeaufsichtigt und/oder nachts statt. Ein erfolgreich etabliertes Testdatenmanagement trägt wesentlich zur Stabilität der Testläufe bei, indem es dafür Sorge trägt, dass die passenden Testdaten in den benötigten Ausprägungen zum Zeitpunkt der Testausführung bereitstehen.

Die einzelnen Aktivitäten im Testdatenmanagement sollten in den Prozessen der Testautomatisierung aufgeführt und mit den Test- und sonstigen Aktivitäten verknüpft sein.

Die Skripte zur Testdatenerzeugung und -bereitstellung müssen zeitlich und inhaltlich mit den Skripten der Testautomatisierung abgestimmt sein.

Um den Aufwand im Zusammenhang mit Testdaten und deren Bereitstellung für die Automatisierung in den Testumgebungen möglichst gering zu halten, sollten folgende Punkte beachtet werden:

- Pflege der Testdaten mit dem Ziel der möglichen Wiederverwendung,
- Effizienz beim Extrahieren, Transformieren und Laden der Daten, z. B. in komplexen Testumgebungen,
- Kenntnisse über die Vollständigkeit und Korrektheit der Daten und
- Löschung nicht mehr benötigter Datenbestände.

7. Glossar

In diesem Kapitel werden nur solche Begriffe definiert, die nicht im aktuellen ISTQB® Glossar enthalten sind.

Begriff	Definition
Anonymisierung	Das Verändern der Originaldaten derart, dass die Einzelangaben über persönliche oder sachliche Verhältnisse nicht oder nur mit einem unverhältnismäßig hohen Aufwand einer natürlichen Person zu geordnet werden können. [BDSG]
Anonymisierungsgrad	Die Wahrscheinlichkeit, dass ein vermeintlicher Angreifer die produktive Information aufgrund der ihm zugänglichen anonymisierten Informationen errät.
Auditierbarkeit	Die Prüfbarkeit eines Prozesses im Rahmen eines Audits.
Data Profiling Werkzeug	Ein Software-Werkzeug zur Analyse der Struktur, des Inhalts und der Beziehungen von Datenbeständen, zur Durchführung von Aufbereitungsarbeiten und zur Untersuchung der Datenqualität.
Datenalterung	Manipulation von zeitrelevanten Feldern z. B. durch SQL, um die Konsistenz der Daten zu erhalten, wenn diese von der Produktion entkoppelt vorliegen.
Datenmaskierung	Das Verändern von Originaldaten durch zufällige Zeichen oder Daten so, dass die Originaldaten nicht wiedererkennbar sind.
Datenminimierung	Der Grundsatz der Datenminimierung des Art. 5 DSGVO entspricht teilweise dem früheren Grundsatz der Datenvermeidung des alten Bundesdatenschutzgesetzes (§ 3a BDSG a.F.). Der Datenschutzgrundsatz der Datenminimierung besagt zusammengefasst, dass Daten nur in dem Umfang und für die Dauer erhoben werden dürfen, wie sie auch zur Erreichung des jeweiligen Zweckes gebraucht werden.

Begriff	Definition
Datensicherheit	<p>Im Unterschied zum Datenschutz befasst sich die Datensicherheit mit dem Schutz von Daten, unabhängig davon, ob diese einen Personenbezug aufweisen oder nicht. Unter den Begriff Datensicherheit fallen daher grundsätzlich auch Daten, die keinen Personenbezug haben.</p> <p>Die Datensicherheit ist im Kontext des Datenschutzes gemäß § 9 BDSG (inkl. Anlage) durch Umsetzung geeigneter technischer und organisatorischer Maßnahmen zu gewährleisten.</p> <p>Die Datensicherheit wird als Bestandteil der Informationssicherheit angesehen, welche umfassender ist.</p>
Datensparsamkeit	<p>"Die Erhebung, Verarbeitung und Nutzung personenbezogener Daten und die Auswahl und Gestaltung von Datenverarbeitungssystemen sind an dem Ziel auszurichten, so wenig personenbezogene Daten wie möglich zu erheben, zu verarbeiten oder zu nutzen. Insbesondere sind personenbezogene Daten zu anonymisieren oder zu pseudonymisieren, soweit dies nach dem Verwendungszweck möglich ist und keinen im Verhältnis zu dem angestrebten Schutzzweck unverhältnismäßigen Aufwand erfordert." Zitat: §3a Bundesdatenschutzgesetz zur "Datenvermeidung und Datensparsamkeit"</p>
ETL-Werkzeug	<p>Ein Werkzeug, mit dessen Hilfe Daten aus mehreren gegebenenfalls unterschiedlich strukturierten Quellen extrahiert, anhand von Regeln in die Struktur einer Zieldatenbank transformiert und in eine Zieldatenbank geladen werden. ETL ist das Akronym von extract, transform und load.</p>
IT-Governance	<p>Die Verantwortung des Vorstands und der Geschäftsführung für die Sicherstellung, dass die IT des Unternehmens die Unternehmensstrategie und die Unternehmensziele unterstützt, wahrgenommen durch die Leitung, die Organisationsstruktur und die Prozesse. [ISACA]</p>

Begriff	Definition
Metadaten	Daten, die Informationen über Merkmale anderer Daten enthalten, aber nicht diese Daten selbst.
personenbezogene Daten	Einzelangaben über persönliche oder sachliche Verhältnisse einer bestimmten oder bestimmbarer natürlichen Person (Betroffener). [BDSG]
Perturbation	Siehe Shuffling
Produktivdaten	Daten, die aus dem produktiven Einsatz eines Systems oder einer Komponente stammen.
pseudonymisierte Daten	Daten die durch Pseudonymisierung von Produktivdaten erzeugt wurden. Dabei gibt es mindestens eine Zuordnungsregel, die einem Originalwert ein Pseudonym zuordnet.
Pseudonymisierung	Im §3 Abs. 6a des BDSG ist die Pseudonymisierung „das Ersetzen des Namens und anderer Identifikationsmerkmale durch einen Kennzeichner zu dem Zweck, die Bestimmung des Betroffenen auszuschließen oder wesentlich zu erschweren.“ [BDSG]
Revisionssicherheit	Die Nachvollziehbarkeit der Führung und Aufbewahrung von geschäftlichen Büchern, Aufzeichnungen und Unterlagen im Unternehmen. Siehe auch: Bundesfinanzministerium – Grundsätze zur ordnungsmäßigen Führung und Aufbewahrung von Büchern, Aufzeichnungen und Unterlagen in elektronischer Form sowie zum Datenzugriff (GoBD)
Shuffling	Shuffling oder auch Perturbation bedeutet „Durcheinanderwirbeln“. Mit Shuffling werden gleiche Datensätze durcheinandergewirbelt, d. h. die enthaltenen Daten werden zwischen den Datensätzen nach einem vorgegebenen Schlüssel getauscht. [TDM Basiswissen]
synthetische Daten	Testdaten eines Systems oder einer Komponente, die keinen direkten Bezug zu

Begriff	Definition
	Produktivdaten haben.
Testdaten	Werte von Variablen bzw. Objekte von Klassen eines Systems oder einer Komponente, welche mit konkreten Testfällen in Verbindung stehen.
Testdatenabschlussbericht	Ein Dokument, das die erfolgten Aktivitäten des Testdatenmanagements und deren Ergebnisse in einem konkreten Projekt zusammenfasst.
Testdatenbereitstellungsbericht	Ein Dokument, das die bereitgestellten Testdaten identifiziert und ihre Freigabe für das Testen zusammenfasst.
Testdatenbereitstellungsplan	Ein Plan der festlegt, wann welche Testdatenbereitstellungsroutinen von wem ausgeführt werden müssen.
Testdatenbereitstellungsroutine	Eine Komponente aus Software und manuellen Prozessen sowie deren Dokumentation, die ausgeführt werden müssen, um konkrete Testdaten für die Durchführung der Testfälle zu erzeugen und um Testdaten und Testumgebungen in einen definierten Zustand zu versetzen.
Testdatenbewertungsbericht	Ein Dokument, das die Bewertung der Qualität der bereit gestellten Testdaten zusammenfasst.
Testdatengewinnungsart	Klassifikation der Gewinnung der Testdaten für ein System oder eine Komponente in synthetische Testdaten, Produktivdaten, anonymisierte Daten und pseudonymisierte Daten.
Testdatenkoeffizient	Der Anteil an in einer Testsuite genutzten Testdaten am Gesamtvolumen der Testdaten eines Systems oder einer Komponente.
Testdatenkonzept	Ein Dokument, das die Prozesse des Testdatenmanagements, die Schnittstellen zu anderen Prozessen, die Ressourcen und die Zeitplanung der Aktivitäten in einem konkreten Projekt beschreibt.

Begriff	Definition
Testdatenmanager	Eine Person, die für das Management der Testdaten verantwortlich ist, um die im Testprozess benötigten Testdaten zeitgerecht, kosteneffizient und in der geforderten Qualität bereit zu stellen.
Testdatenmodellierer	Eine fachlich und methodisch sachkundige Person, die Testdatenanforderungen analysiert und Testdaten spezifiziert.
Testdatenrealisierer	Eine IT-technisch qualifizierte Person, die konkrete Testdaten erzeugt bzw. bereitstellt.
Testdatenrichtlinie	Ein Dokument, das auf hohem Abstraktionsniveau die Prinzipien, den Ansatz und die wichtigsten Ziele des Testdatenmanagements zusammenfasst, mit organisationsweiter oder projektübergreifender Gültigkeit. Entspricht dem Testdatenkonzept auf einem umfassenderen, abstrakteren Level.
Testdatenspezifikation	Ein Dokument, das Ein- oder Ausgabewerte für ein Testobjekt formal beschreibt, sowie die Herkunft der Testdaten, deren Qualität und Menge vorgibt.
Testdatenverwendungsart	Klassifikation der Verwendung von Testdaten für eine Testsuite in Eingabedaten, Zustandsdaten, Ausgabedaten und Soll-Daten.
Testumgebungsmanagement	Dynamischer Testprozess der Einrichtung und Wartung einer geforderten Testumgebung [ISO/IEC/IEEE 29119]
Zeitreisen	Zeitlich gerraftes Ausführen von (fachlichen) Anwendungen, um die Konsistenz der Daten zu erhalten.

8. Literaturhinweise

8.1 Normen und Standards

Referenz	Titel
[ISACA]	ISACA (r) Glossary of Terms English-German, 3rd Ed., 2015, www.isaca.org
[DSGVO]	Datenschutz-Grundverordnung https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=CELEX%3A32016R0679&qid=1614950868670 sowie https://dsgvo-gesetz.de
[GoBD]	Grundsätze zur ordnungsmäßigen Führung und Aufbewahrung von Büchern, Aufzeichnungen und Unterlagen in elektronischer Form sowie zum Datenzugriff (GoBD) https://www.bundesfinanzministerium.de/Content/DE/Downloads/BMF_Schreiben/Weitere_Steuerthemen/Abgabenordnung/2019-11-28-GoBD.html
[ISO 9000]	DIN EN ISO 9000:2015-11 Qualitätsmanagementsysteme – Grundlagen und Begriffe (ISO 9000:2015); Deutsche und Englische Fassung EN ISO 9000:2015. Beuth Verlag, 2015.
[ISO27001]	DIN EN ISO/IEC 27001:2017: Informationstechnik – Sicherheitsverfahren – Informationssicherheitsmanagementsysteme – Anforderungen. Beuth Verlag GmbH, 2017.
[ISO/IEC/IEEE 29119-3]	ISO/IEC/IEEE 29119-3: Software and systems engineering – Software testing – Part 3: Test documentation, 2013.

8.2 ISTQB®-Dokumente

Referenz	Titel
	Lehrplan ISTQB Certified Tester Foundation Level 2018 V3.1D https://www.german-testing-board.info/lehrplaene/istqbr-certified-tester-schema/lehrplaene/
[GTB-GLOSSAR]	https://www.german-testing-board.info/lehrplaene/istqbr-certified-tester-schema/glossar/

8.3 Fachliteratur

Referenz	Titel
[TDM Basiswissen]	Franz, Tremmel, Kruse: Basiswissen Testdatenmanagement, Aus- und Weiterbildung zum Test Data Specialist Certified Tester Foundation Level nach GTB®, 2018, dpunkt.verlag, ISBN 978-3-86490-558-2
[BDSG]	Gola, Schomerus: BDSG Bundesdatenschutzgesetz, Verlag: C.H. Beck, 10. Auflage, ISBN 978 3 406 598340
[SpLi10]	Andreas Spillner, Tilo Linz; Basiswissen Softwaretest: Aus- und Weiterbildung zum Certified Tester - Foundation Level nach ISTQB®-Standard; dpunkt Verlag, 4. Auflage, 2010, ISBN 3898646242
[GFB]	Testdatenmanagement - Aufgaben im Testprozess, hg. von der GFB Softwareentwicklungsgesellschaft mbH, Oberursel 2011, http://www.gfb-consulting.de/wp-content/uploads/2015/10/WP_Testdatenmanagement_1.pdf

8.4 Internetquellen

Referenz	URL
[ASQF]	ASQF Arbeitsgruppe Testdatenmanagement https://www.asqf.de/arbeitsgruppen/arbeitsgruppe-testdatenmanagement/#1487749832500-07d53ffa-857c

9. Index

abstrakter Testfall 19, 20
Anonymisierung 15, 19, 23, 26, 27, 28, 29, 33, 40, 50, 54, 58, 61
Anonymisierungsgrad 19, 23, 27, 61
Auditierbarkeit 30, 34, 61
Data Profiling Werkzeug 56, 61
Datenalterung 25, 29, 61
Datenmaskierung 19, 23, 61
Datenminimierung 18, 19, 29, 32, 33, 61
Datenschutz 13, 14, 16, 17, 23, 26, 31, 32, 33, 49, 58, 62, 66
Datensicherheit 13, 17, 18, 30, 33, 62
Datensparsamkeit 62
ETL-Werkzeug 56, 62
IT-Governance 30, 35, 62
Konfigurationsmanagement 48, 52, 53
Konformität 2, 16, 30, 32, 34, 39
konkreter Testfall 19, 20
Metadaten 19, 28, 58, 63
personenbezogene Daten 16, 23, 24, 30, 32, 33, 62, 63
Perturbation 27, 28, 63
Produktivdaten 19, 22, 23, 24, 26, 27, 28, 29, 33, 63, 64
pseudonymisierte Daten 19, 23, 63, 64
Pseudonymisierung 19, 23, 27, 50, 54, 58, 63
Revisionssicherheit 16, 52, 54, 55, 63
Shuffling 28, 63
synthetische Daten 18, 19, 63
Testausführungswerkzeug 56

Testdaten 9, 10, 11, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 26, 27, 28, 29, 30, 33, 34, 37, 39, 40, 41, 42, 44, 45, 46, 47, 48, 49, 50, 53, 54, 57, 58, 59, 60, 63, 64, 65

Testdatenabschlussbericht 36, 37, 42, 46, 48, 59, 64

Testdatenart 19

Testdatenbereitstellungsbericht 36, 42, 46, 59, 64

Testdatenbereitstellungsplan 36, 41, 42, 46, 50, 59, 64

Testdatenbereitstellungsroutine 36, 64

Testdatenbewertungsbericht 36, 43, 46, 47, 64

Testdateneditor und -generator 56

Testdatengewinnungsart 19, 64

Testdatenkoeffizient 19, 29, 64

Testdatenkonzept 36, 39, 41, 43, 44, 45, 48, 64, 65

Testdatenmanagement 8, 10, 11, 12, 13, 14, 15, 17, 19, 20, 23, 33, 34, 35, 36, 37, 42, 47, 53, 54, 57, 59, 60, 67

Testdatenmanager 36, 43, 44, 46, 47, 48, 49, 53, 65

Testdatenmodellierer 36, 44, 45, 46, 47, 49, 50, 53, 65

Testdatenrealisierer 36, 44, 45, 46, 47, 50, 53, 65

Testdatenrichtlinie 36, 39, 43, 48, 53, 54, 65

Testdatenspezifikation 36, 40, 41, 45, 49, 50, 51, 65

Testdatenverwendungsart 19, 65

Testentwurfswerkzeug 56

Testinfrastruktur 13, 17, 34

Testmanagement 43, 46, 52, 53, 57, 59

Testprozess 8, 11, 14, 15, 16, 28, 36, 37, 40, 43, 44, 47, 53, 65, 67

Testumgebung 13, 14, 15, 25, 54, 65

Testumgebungsmanagement 52, 53, 65

Zeitreisen 41, 45, 65