

Florian Fieber · Marc-Florian Wendland

Basiswissen



Abnahmetest

Aus- und Weiterbildung zum
ISTQB® Foundation Level
Specialist – Acceptance Testing

 dpunkt.verlag

Inhalt

Cover

Über die Autoren

Titel

Impressum

Vorwort

Geleitwort

Inhaltsübersicht

Inhaltsverzeichnis

1 Einleitung

1.1 Vertrauen, Qualität, Sicherheit!

1.2 ISTQB® Certified Tester – das Zertifizierungsprogramm für Softwaretester

1.3 Nutzen dieses Buches

1.4 Kapitelübersicht

1.5 Fallbeispiel »CA-Cockpit«

2 Grundlagen des Abnahmetests

2.1 Der Abnahmetest im Softwarelebenszyklus

2.1.1 Testen im Softwarelebenszyklus

2.1.2 Zweck und Ziele des Abnahmetests

2.1.3 Abnahmetests in sequenziellen vs. agilen
Softwareentwicklungsmodellen

2.2 Grundlegende Beziehungen

2.2.1 Geschäftsziele, Geschäftsbedarfe und Anforderungen

2.2.2 Anforderungen/User Stories, Abnahmekriterien und Abnahmetests

2.2.3 Die Wichtigkeit der Qualität der Anforderungen

2.3 Businessanalyse und Abnahmetests

- 2.3.1 Beziehungen zwischen Businessanalyse- und Testaktivitäten
- 2.3.2 Zusammenarbeit zwischen Businessanalysten und Testern beim Abnahmetest
- 2.3.3 Wie der Abnahmetest den Entwicklungsprozess vorantreiben kann: ATDD und BDD

3 Abnahmekriterien, Abnahmetests und erfahrungsbasierte Praktiken

- 3.1 Abnahmekriterien erstellen
- 3.2 Abnahmetests entwerfen
 - 3.2.1 Testvorgehensweisen und Testverfahren für den Abnahmetest
 - 3.2.2 Testfälle mit der Gherkin-Sprache erstellen
- 3.3 Erfahrungsbasierte Ansätze für den Abnahmetest
 - 3.3.1 Exploratives Testen
 - 3.3.2 Beta-Tests

4 Nicht funktionale Anforderungen im Abnahmetest

- 4.1 Nicht funktionale Qualitätsmerkmale und Nutzungsqualität
 - 4.1.1 Das ISO-25010-Qualitätsmodell
 - 4.1.2 Das ISO-25010-Modell der Nutzungsqualität
 - 4.1.3 Nicht funktionale Qualitätsmerkmale im Abnahmetest
- 4.2 Gebrauchstauglichkeit und Benutzererlebnis
 - 4.2.1 Gebrauchstauglichkeit
 - 4.2.2 Benutzererlebnis (UX)
 - 4.2.3 Gebrauchstauglichkeitstest
- 4.3 Performanz
 - 4.3.1 Anforderungen und Abnahmekriterien für Performanzabnahmetests
 - 4.3.2 High-Level-Performanzabnahmetests
- 4.4 IT-Sicherheit
 - 4.4.1 Anforderungen und Abnahmekriterien für IT-Sicherheitsabnahmetests
 - 4.4.2 High-Level-IT-Sicherheitsabnahmetests

5 Modellierung von Geschäftsregeln und Geschäftsprozessen

5.1 Grundlagen der Modellierung

5.2 Geschäftsprozess- und Geschäftsregelmodellierung mit BPMN und DMN

5.2.1 Einführung in BPMN

5.2.1.1 Kernbereiche der BPMN

5.2.1.2 Prozesse

5.2.2 Einführung in DMN

5.2.2.1 Verwendungsarten von DMN

5.2.2.2 Kernbereiche von DMN

5.2.2.3 Klassische Entscheidungstabellen

5.2.2.4 DMN-Entscheidungstabellen

5.2.2.5 DMN-Entscheidungstabellen mit BPMN verknüpfen

5.3 Geschäftsprozess-/Geschäftsregelmodelle im Abnahmetest verwenden

5.3.1 Modellbasiertes Testen

5.3.2 Bewährte Testmodellierungspraktiken

5.4 Abnahmetests aus Geschäftsprozess-/Geschäftsregelmodellen ableiten

5.4.1 Ableitung von Abnahmetests aus Geschäftsprozess- und Geschäftsregelmodellen

5.4.1.1 Strukturelle Überdeckungskriterien

5.4.1.2 Weitere Überdeckungskriterien

5.4.1.3 Testentwurf und -realisierung mit BPMN

5.4.2 Ableitung von Abnahmetests aus Entscheidungstabellen

5.4.2.1 Vollständige Entscheidungstabellen

5.4.2.2 Reduzierung vollständiger Entscheidungstabellen mit dem Prüfsummenverfahren

5.4.2.3 Testentwurf und -realisierung mittels Entscheidungstabellen

6 Kollaborativer Abnahmetest

6.1 Zusammenarbeit

6.2 Aktivitäten

6.2.1 Fehleranalyse

6.2.2 Berichterstattung

6.2.3 Qualitätssicherungsaktivitäten für den Abnahmetest

6.3 Werkzeugunterstützung

6.3.1 Werkzeugarten im Abnahmetest

6.3.2 Auswahl und Einführung

Anhang

A Wichtige Hinweise zum Lehrstoff und zur Prüfung zum Certified Tester

A.1 Lernziele des Lehrplans

A.1.1 Lernziele zu Kapitel 2: Grundlagen des Abnahmetests

A.1.2 Lernziele zu Kapitel 3: Abnahmekriterien, Abnahmetests und erfahrungsbasierte Praktiken

A.1.3 Lernziele zu Kapitel 4: Nicht funktionale Anforderungen im Abnahmetest

A.1.4 Lernziele zu Kapitel 5: Modellierung von Geschäftsregeln und Geschäftsprozessen

A.1.5 Lernziele zu Kapitel 6: Kollaborativer Abnahmetest

B Abkürzungsverzeichnis

C Glossar

D Quellenverzeichnis

D.1 Bücher und Artikel

D.2 Lehrpläne

D.3 Normen und Standards

D.4 Web

Index

4 Nicht funktionale Anforderungen im Abnahmetest

Funktionalität ist für die meisten Stakeholder das offensichtlichste Merkmal eines abzunehmenden Produkts. Dass Funktionalität jedoch nicht alles ist, sollte jedem bewusst sein, der schon einmal Frustration bei der Bedienung eines Systems über dessen Benutzungsschnittstelle empfunden oder sich über gestohlene persönliche Daten wie Kreditkarteninformationen und/oder Zugangsdaten erschrocken oder geärgert hat. Die nicht funktionalen Anforderungen eines Produkts können einen großen Einfluss darauf haben, wie ein Benutzer (also der Kunde oder Anwender) die Arbeit mit dem Produkt erlebt, d. h., ob das Produkt erfolgreich ist oder gar vollständig abgelehnt wird. In diesem Kapitel werden die für den Benutzerabnahmetest wichtigsten nicht funktionalen Anforderungen diskutiert und es wird beschrieben, wie der Abnahmetest den Umgang mit den dazugehörigen Abnahmekriterien unterstützt.

4.1 Nicht funktionale Qualitätsmerkmale und Nutzungsqualität

Die Qualität eines Softwaresystems ist definiert als die Erfüllung von Anforderungen. Es gibt jedoch nicht per se *die* Qualität, sprich, Qualität ist kein einzelner Aspekt, sondern beschreibt sich durch eine Menge von einzelnen Qualitätsmerkmalen, die auf unterschiedliche Aspekte eines Softwaresystems abzielen, beispielsweise Performanz oder Gebrauchstauglichkeit. Jedes einzelne Qualitätsmerkmal trägt einen Teil zur gesamten Qualität des Softwaresystems bei, indem für die relevanten Qualitätsmerkmale entsprechende Anforderungen definiert werden, deren Erfüllung zu dem angestrebten Qualitätsniveau beitragen. Anforderungen können jedoch nur erfüllt werden, wenn sie testbar sind, d. h., es eindeutig nachweisbar ist, ob die Anforderungen erfüllt wurden oder nicht. Zu diesem Zweck werden messbare Abnahmekriterien bestimmt, die wiederum die

Grundlage für Abnahmetests bilden. Die Ergebnisse der Abnahmetests liefern Erkenntnisse über der Erfüllungsgrad der überdeckten Anforderungen eines Qualitätsmerkmals. Aus dem Erfüllungsgrad aller Qualitätsmerkmale ergibt sich wiederum ein Abbild der Qualität des Softwaresystems. Abbildung 4-1 illustriert diesen Zusammenhang.

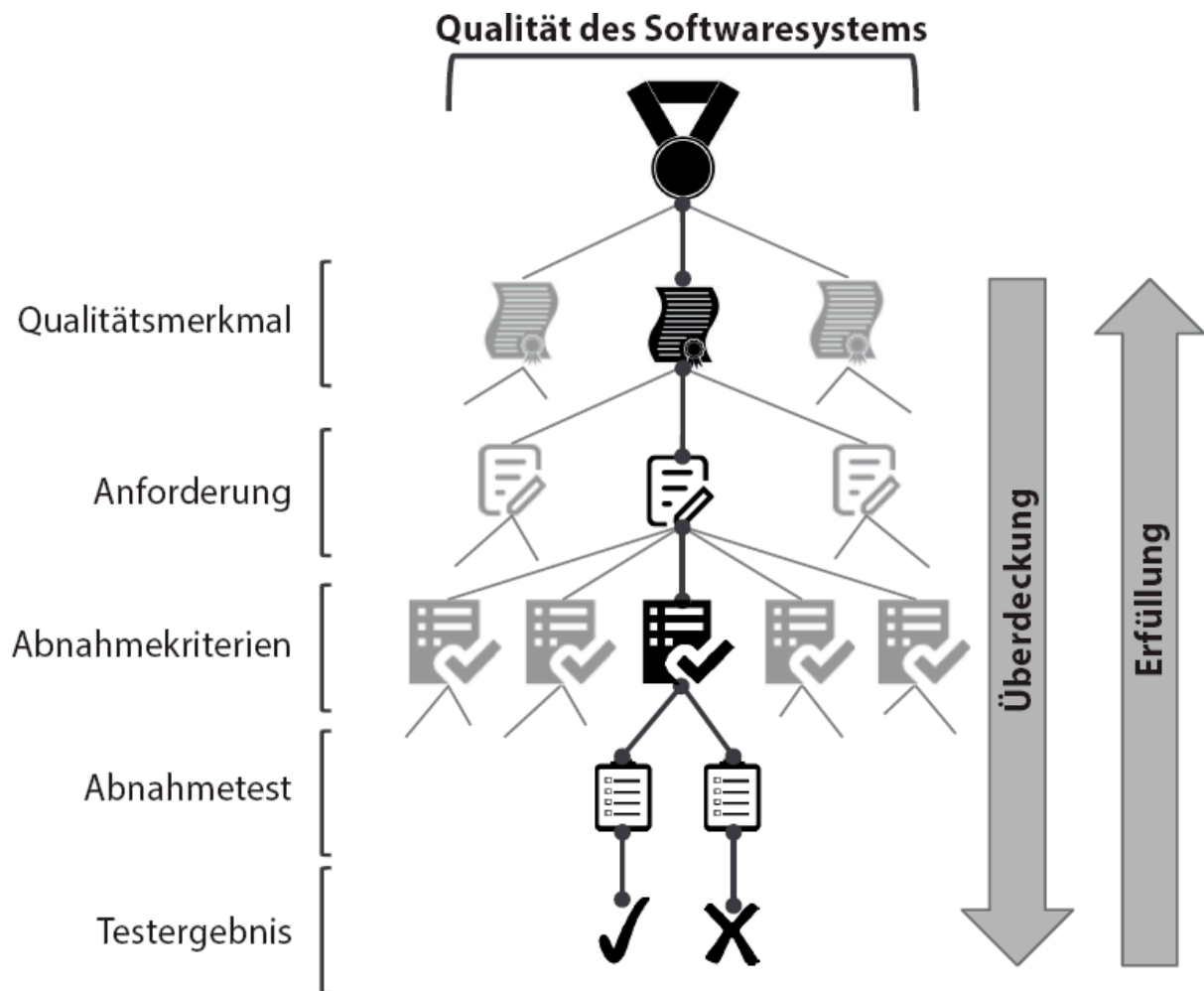


Abb. 4-1

Zusammenhang von Qualität, Qualitätsmerkmalen und Anforderungen

Anforderungen an Softwaresysteme lassen sich grundsätzlich nach funktionalen und nicht funktionalen Eigenschaften kategorisieren. Funktionale Anforderungen spezifizieren, *was* das System leisten soll. Sie sind der Grund, warum Softwaresysteme überhaupt konzipiert und implementiert werden. Von ihrer Umsetzung erhofft man sich einen vorteilhaften Geschäftsnutzen. Sie bilden damit die Voraussetzung dafür, dass das System überhaupt einsetzbar ist. Ihnen gegenüber stehen die nicht funktionalen Anforderungen. Sie spezifizieren, *wie* das System die Leistung erbringen soll. Nicht funktionale Anforderungen beschreiben also Attribute des funktionalen Verhaltens, sprich, mit welcher Qualität das

(Teil-)System seine Funktion erbringen soll (siehe [Spillner & Linz 19], Abschnitt 3.5.1 und 3.5.2).

4.1.1 Das ISO-25010-Qualitätsmodell

Qualitätsmerkmale können in Qualitätsmodellen zusammengefasst werden. Neben firmen- oder branchenspezifischen Softwarequalitätsmodellen existiert mit der ISO 25010 [ISO 25010], dem sogenannten »Systems and software Quality Requirements and Evaluation (SQuaRE)«-Standard, ein Qualitätsmodell, das acht Qualitätsmerkmale samt Untermerkmalen definiert (siehe Abb. 4–2).

ISO 25010 Qualitätsmerkmale			
Funktionale Eignung	Zuverlässigkeit	Performanz	Gebrauchstauglichkeit
<i>funktional</i>	<i>nicht funktional</i>	<i>nicht funktional</i>	<i>nicht funktional</i>
IT-Sicherheit	Wartbarkeit	Portabilität	Kompatibilität
<i>nicht funktional</i>	<i>nicht funktional</i>	<i>nicht funktional</i>	<i>nicht funktional</i>

Abb. 4–2

ISO-25010-Qualitätsmodell

ISO 25010 definiert die Qualitätsmerkmale wie folgt¹:

Funktionale Eignung

Der Grad, zu dem eine Komponente oder ein System Funktionen zur Verfügung stellt, welche unter festgelegten Bedingungen explizit genannte und implizite Bedürfnisse erfüllen. Die funktionale Eignung (oder auch Funktionalität) ist das wohl sichtbarste Merkmal eines Softwaresystems – sowohl für Anwender als auch Hersteller. Funktionale Eignung und Funktionalität sind synonym zu verstehen, der Gewohnheit wegen wird in Diskussionen zumeist der Begriff Funktionalität verwendet. Dies mag auch daran liegen, dass der Vorgänger von ISO 25010, der ISO-Standard 9126, diesen Begriff verwendet hat. Funktionale Eignung trifft den Sachverhalt allerdings besser, da durch das Wort *Eignung* bereits impliziert wird,

dass Funktionalität immer zweckgebunden ist und ungeeignete, aber korrekt implementierte Funktionen dennoch zu einem niedrigen Grad an funktionaler Eignung führen.

Performanz

Der Grad, zu dem eine Komponente oder ein System Zeit, Ressourcen und Kapazität verbraucht, während sie/es seine vorgesehenen Funktionen ausführt. Unter Performanz wird üblicherweise das Zeitverhalten eines Softwaresystems verstanden. Performanz ist ein nicht funktionales Qualitätsmerkmal, das auf Funktionalität aufsetzt. Es wird detaillierter in Abschnitt 4.3 besprochen.

Kompatibilität

Der Grad, zu dem eine Komponente oder ein System Informationen mit anderen Komponenten oder Systemen austauschen kann, und/ oder ihre geforderten Funktionen bei einer gemeinsamen Benutzung einer Hardware- oder Softwareumgebung ausführen kann. Kompatibilität ist eines der beiden Qualitätsmerkmale, die in ISO 25010 neu hinzugekommen sind (das andere ist IT-Sicherheit). Durch die fortwährende Vernetzung und dynamische Konnektivität in Softwaresystemen, insbesondere bei cyberphysischen Systemen und Internet-of-Things-Anwendungen (IoT), wird die Kompatibilität von Systemen, Teilsystemen, Komponenten oder schlichtweg *Things* immer wichtiger.

Gebrauchstauglichkeit

Der Grad, zu dem eine Komponente oder ein System durch bestimmte Benutzer in einem bestimmten Nutzungskontext genutzt werden kann, um festgelegte Ziele effektiv, effizient und zufriedenstellend zu erreichen. Gebrauchstauglichkeit hängt stark mit der funktionalen Eignung zusammen, da Anwender mit einem Softwaresystem interagieren müssen, um die Funktionalität auszuführen. Die Gestaltung einer Mensch-Maschine-Schnittstelle mit hoher Gebrauchstauglichkeit ist ein nicht trivialer, erfolgskritischer Prozess. Gebrauchstauglichkeit ist ein nicht funktionales Qualitätsmerkmal und wird detaillierter in Abschnitt 4.2 besprochen.

Zuverlässigkeit

Der Grad, zu dem eine Komponente oder ein System seine spezifizierten Funktionen unter den festgelegten Bedingungen während einer bestimmten Zeitspanne ausführt. Zuverlässigkeit spielt insbesondere bei sicherheitskritischen Systemen eine tragende Rolle. Zuverlässige Softwaresysteme sind robust gegenüber Fehleingaben und weisen eine hohe Verfügbarkeit auf. Zuverlässigkeit ist ein nicht funktionales Qualitätsmerkmal.

IT-Sicherheit

Der Grad, zu dem eine Komponente oder ein System Informationen und Daten schützt, sodass Personen oder andere Komponenten oder Systeme nur einen solchen Grad an Zugriff erhalten, der ihrer Berechtigungsart und -stufe entspricht. In Zeiten zunehmender Vernetzung (siehe Qualitätsmerkmal *Kompatibilität*) nimmt die IT-Sicherheit mehr und mehr an Bedeutung zu. Das kontinuierliche Fortschreiten der Digitalisierung führt dazu, dass mehr und mehr Geschäftsprozesse und Informationen des alltäglichen Lebens rechnergestützt umgesetzt bzw. gespeichert werden. IT-Sicherheit ist ein nicht funktionales Qualitätsmerkmal und wird detaillierter in Abschnitt 4.4 besprochen.

Wartbarkeit

Der Grad, zu dem eine Komponente oder ein System von den dafür vorgesehenen Personen gewartet werden kann. Die Wartbarkeit eines Softwaresystems ist üblicherweise kein Qualitätsmerkmal, an dem die späteren Anwender großes Interesse haben. Gute Wartbarkeit wird von den Anwendern zumeist kaum wahrgenommen. Schlechte Wartbarkeit macht sich vor allem durch zahlreiche Updates und Patches bemerkbar. Dies kommt daher, dass sich schlecht wartbare Softwaresysteme schlechter analysieren und modifizieren lassen und daher die Wahrscheinlichkeit zur Einbringung neuer Fehler oder unerwünschter Seiteneffekte sehr groß ist. Aus diesem Grund sollte Wartbarkeit insbesondere im Interesse des Herstellers sein, da sie sich auf Weiterentwicklungen und Anpassungen entweder unterstützend oder erschwerend auswirkt.

Portabilität (Übertragbarkeit)

Der Grad, zu dem eine Komponente oder ein System von einer Hardware, Software oder einer anderen Betriebs- oder Nutzungsumgebung auf eine andere übertragen werden kann. Portabilität bzw. Übertragbarkeit ist insbesondere für Hersteller relevant, die eine möglichst große Marktdurchdringung anstreben. Daher ist es nicht verwunderlich, dass in erster Linie Hersteller kommerzieller Produkte, wie beispielsweise Text- oder Videoverarbeitung, aber auch Computerspiele², danach streben, auf möglichst vielen Plattformen (also Betriebs- oder Nutzungsumgebungen) präsent und somit konkurrenzfähig zu sein. Portabilität spielt daher auch bei mobilen Anwendungen und Endgeräten eine gewichtige Rolle.

Nicht alle Qualitätsmerkmale haben die gleiche Wichtigkeit für die Qualität eines Softwaresystems. Vielmehr gilt es zu ermitteln, welches Qualitätsmerkmal für den Erfolg des Softwaresystems wichtiger ist und welches eher vernachlässigt werden

kann. Dies zu ermitteln ist üblicherweise die Aufgabe des Auftraggebers bzw. der Stakeholder. So spielt das Qualitätsmerkmal *Gebrauchstauglichkeit* für eingebettete Systeme gewöhnlich eine geringere Rolle als *Performanz* oder *Zuverlässigkeit*. Ein Softwaresystem, das keinerlei vertrauliche Daten zu schützen hat, muss auf die *IT-Sicherheit* weniger Wert legen als ein Softwaresystem, das vertrauliche oder personenbezogene Daten verarbeitet. Ein Softwaresystem, das nur in einer Zielumgebung eingesetzt werden soll, kann das Qualitätsmerkmal *Portabilität* vernachlässigen. Ein Softwaresystem, für das ein langer Einsatz vorgesehen wird mit dem Bedarf, sich stets an neue Anforderungen anpassen zu müssen, sollte sich durch geringe Wartungsaufwände, also eine gute *Wartbarkeit*, auszeichnen.

Wartbarkeit der Testmittel

Testmittel unterliegenden fortwährenden Anpassungen. Sie existieren einzig und alleine zu dem Zweck, ein bestimmtes System in einer bestimmten Version und bestimmten Konfiguration zu testen. Unterliegt dieses System nun Änderungen, ganz gleich ob technischer oder fachlicher Art, so schlagen sich diese sofort und unmittelbar auf die korrespondierenden Testmittel durch. Daher sollten Tester ein großes Bedürfnis entwickeln, dass die entworfenen Testmittel für den Test eines Softwaresystems eine hohe Wartbarkeit aufweisen. Schlechte Wartbarkeit von Testmitteln resultiert zumeist in langen Anpassungszyklen, die zu einem potenziellen Projektrisiko werden können.

Es gibt einige bewährte konstruktive Praktiken, die die Wartbarkeit von Testmittel unterstützen bzw. forcieren. Die Etablierung von bidirektionaler Verfolgbarkeit zwischen den konstruktiven Artefakten und Testmitteln unterstützt die Änderungsauswirkungsanalyse. Durch Einbringung von Abstraktion in automatisierte Testskripte wie es beispielsweise bei der verhaltensgetriebenen Entwicklung oder dem schlüsselwortgetriebenen Testen der Fall ist, wird die Wartbarkeit dieser Testskripte hinsichtlich rein technischer Änderungen verbessert (wie bereits in Abschnitt 2.3.3 erläutert).

Neben den konstruktiven Möglichkeiten gibt es auch bewährte Praktiken auf Prozessebene, die die Wartbarkeit der Testmittel adressieren. Der Test-First-Gedanke geht beispielsweise davon aus, dass die Erstellung oder Anpassung eines Testfalls vor der eigentlichen Entwicklung des zu testenden Codes erfolgt. Dadurch wird die Erstellung bzw. Aktualisierung von Testfällen inhärenter Bestandteil des Prozesses, was zur Folge hat, dass Testfälle stets aktuell und durchführbar sind. Dies trifft u. a. auf die testgetriebene Entwicklung und die abnahmetestgetriebene Entwicklung zu.

Im Abnahmetest müssen sowohl funktionale als auch nicht funktionale Qualitätsmerkmale durch entsprechende Anforderungen und Abnahmekriterien berücksichtigt werden. Die Herausforderung bei nicht funktionalen

Qualitätsmerkmalen ist häufig, dass die Auftraggeber sich nicht darüber im Klaren sind, welche nicht funktionalen Qualitätsmerkmale für das Projekt von Relevanz sind. Wie lassen sich beispielsweise Anforderungen an die Wartbarkeit eines Softwaresystems formulieren? Wie identifiziert und spezifiziert man die Anzahl der gleichzeitig zu unterstützenden Benutzer mit ihren spezifischen operationalen Profilen? Weil die Antwort auf diese Fragen den meisten Auftraggebern schwerfällt, sind nicht funktionale Merkmale im Vergleich zu den funktionalen Merkmalen oftmals unterrepräsentiert und – deutlich schwerwiegender – unterspezifiziert. Häufig existieren sie nur als implizite Annahmen aufseiten des Auftraggebers. Die Software soll *schnell* reagieren (Performanz), bei Fehleingaben *nicht abstürzen* (Zuverlässigkeit) bzw. diese *gar nicht erst zulassen* (Gebrauchstauglichkeit) und die Datenintegrität *auf jeden Fall gewähren* (IT-Sicherheit). Diese vagen Angaben seitens des Auftraggebers führen dazu, dass Entwickler und Tester ihrerseits Annahmen darüber treffen müssen, was der Auftraggeber von dem Softwaresystem genau erwartet. Diese Annahmen können mit den impliziten oder unterspezifizierten Erwartungen der Stakeholder übereinstimmen, müssen es aber nicht (siehe auch Abschnitt 2.2.3). Diese Situation ist umso schwieriger, als nicht funktionale Merkmale von Softwaresystemen in unserer vernetzten und digitalisierten Welt an Bedeutung zunehmen. Zuverlässigkeit, IT-Sicherheit, Kompatibilität sind nur einige wenige nicht funktionale Qualitätsmerkmale, auf denen moderne Dienste und Lösungen aufsetzen. Eine unzureichende Berücksichtigung nicht funktionaler Qualitätsmerkmale kann zu Unzufriedenheit der Kunden bzw. Anwender oder sogar zur Ablehnung des Produkts führen – genauso wie das bei falscher, fehlerhafter oder fehlender Funktionalität der Fall ist. Ferner können Umsatzrückgang, Haftungsrisiken, unerwünschte Berichterstattung über die Organisation in den Medien und damit verbundene Imageschäden die Folge unzureichender Qualität auch von nicht funktionalen Qualitätsmerkmalen sein.

Da die in Abbildung 4-2 aufgeführten Qualitätsmerkmale ein sehr hohes Abstraktionsniveau besitzen und der Vielfalt der verschiedenen Merkmale nicht gerecht werden, unterteilt ISO 25010 [ISO 25010] die Qualitätsmerkmale in spezifischere Teilmerkmale. Diese Teilmerkmale sind nachfolgend dargestellt (siehe Abb. 4-3):

ISO 25010 Qualitätsmerkmale			
Funktionale Eignung	Zuverlässigkeit	Performanz	Gebrauchstauglichkeit
<ul style="list-style-type: none"> • Funktionale Vollständigkeit • Funktionale Korrektheit • Funktionale Angemessenheit 	<ul style="list-style-type: none"> • Reife • Verfügbarkeit • Fehlertoleranz • Wiederherstellbarkeit 	<ul style="list-style-type: none"> • Zeitverhalten • Ressourcennutzung • Kapazität 	<ul style="list-style-type: none"> • Erkennbare Angemessenheit • Erlernbarkeit • Operabilität (Bedienbarkeit) • Benutzerfehlerschutz • Ästhetik der Benutzungsschnittstelle • Zugänglichkeit
IT-Sicherheit	Wartbarkeit	Portabilität	
<ul style="list-style-type: none"> • Vertraulichkeit • Integrität • Nichtabstreitbarkeit • Zurechenbarkeit • Authentizität 	<ul style="list-style-type: none"> • Analysierbarkeit • Modifizierbarkeit • Modularisierbarkeit • Wiederverwendbarkeit • Testbarkeit 	<ul style="list-style-type: none"> • Installierbarkeit • Anpassbarkeit • Austauschbarkeit 	
			Kompatibilität
			<ul style="list-style-type: none"> • Interoperabilität • Koexistenz

Abb. 4-3

Untermerkmale des ISO-25010-Qualitätsmodells

Jedes dieser Teilmerkmale lässt sich hinsichtlich seiner Qualität durch statische und/oder dynamische Tests bewerten – wenn die korrespondierenden Anforderungen testbar sind, beispielsweise durch die Bestimmung von Abnahmekriterien. Einige Merkmale eignen sich insbesondere für statische Tests wie beispielsweise die Teilmerkmale der *Wartbarkeit* oder die *Ästhetik der Benutzungsschnittstelle*. Die meisten der Teilmerkmale lassen sich jedoch durch dynamisches Testen direkt oder indirekt durch weiterführende dynamische Analysen während der Durchführung dynamischer Tests bewerten.

Die Qualitätsmerkmale der ISO 25010 und ihre Teilmerkmale sind hierbei nicht als Selbstzweck zu verstehen. Zum einen helfen die Qualitätsmerkmale dabei, eine gemeinsame Sprache zu sprechen, wenn es um die Definition und Bewertung von Qualität geht. Stakeholder bekommen dadurch ein Verständnis, wie sich Qualität definiert. Zugleich können die relevanten und irrelevanten Qualitätsmerkmale identifiziert, kommuniziert und dokumentiert werden, was insbesondere für eine optimierte Testplanung von Belang ist. Testressourcen sind traditionell knappe Ressourcen. Dadurch, dass man sich bei der Testplanung auf die »richtigen«, d. h. die relevanten Qualitätsmerkmale konzentrieren kann, können die knappen Ressourcen zielführend eingesetzt werden. Die Dokumentation der relevanten nicht funktionalen Qualitätsmerkmale sollte dabei nicht ausschließlich, aber auch

für den Abnahmetest im Testkonzept erfolgen. Daraus ergeben sich dann unter anderem auch weitere zu dokumentierende Informationen wie die Prioritäten der relevanten nicht funktionalen Qualitätsmerkmale sowie die anzuwendenden Teststrategien, -vorgehensweisen und -verfahren. So beeinflusst die Priorisierung der nicht funktionalen Qualitätsmerkmale beispielsweise bereits die Testanalyse. Anforderungen, die sich auf Qualitätsmerkmale oder Teilmerkmale mit geringerer Priorität beziehen, sollten später adressiert werden als solche mit höherer Priorität.

4.1.2 Das ISO-25010-Modell der Nutzungsqualität

Neben dem standardisierten Qualitätsmodell, das die funktionalen und nicht funktionalen Qualitätsmerkmale von Softwaresystemen beschreibt, definiert die Norm ISO 25010 auch ein Modell der Nutzungsqualität. Nutzungsqualität und Gebrauchstauglichkeit sind eng miteinander verwoben, so eng, dass sie in der Norm sogar die gleiche Definition teilen. Wie Gebrauchstauglichkeit, so definiert sich auch Nutzungsqualität als der Grad, zu dem eine Komponente oder ein System durch bestimmte Benutzer in einem bestimmten Nutzungskontext genutzt werden kann, um festgelegte Ziele effektiv, effizient und zufriedenstellend zu erreichen. Stellt sich also die Frage, worin der Unterschied zwischen Gebrauchstauglichkeit und Nutzungsqualität liegt?

Die Norm liefert als Antwort darauf eine zusätzliche Anmerkung zur Definition von Gebrauchstauglichkeit. Dort heißt es:

*Mehr als »nur«
Gebrauchstauglichkeit*

»Gebrauchstauglichkeit kann entweder indirekt als Qualitätsmerkmal der entsprechenden Teilmerkmale spezifiziert, gemessen und evaluiert werden oder direkt durch die Teilbereiche der Nutzungsqualität.« Abbildung 4-4 illustriert die fünf Merkmale, die die Nutzungsqualität in Bezug auf die Ergebnisse der Interaktion des Benutzers mit einem System definiert.

ISO 25010 Nutzungsqualität				
Effektivität	Effizienz	Zufriedenheit	Risikofreiheit	Kontext- abdeckung
• Effektivität	• Effizienz	<ul style="list-style-type: none"> • Nützlichkeit • Vertrauen • Freude • Komfort 	<ul style="list-style-type: none"> • Minderung der wirtschaftlichen Risiken • Minderung der Gesundheits- und Sicherheitsrisiken • Minderung der Umwelt- risiken 	<ul style="list-style-type: none"> • Kontextvollständigkeit • Flexibilität

Abb. 4-4

ISO-25010-Modell der Nutzungsqualität

Diese Merkmale sind wie folgt definiert³:

Effektivität

Der Umfang, in dem richtige und vollständige Ziele erreicht werden. Effektivität ist eng mit der Funktionalität des Systems verknüpft und kann nicht weiter relativiert werden. Entweder ist die Benutzungsschnittstelle effektiv und der Benutzer erreicht sein Ziel oder sie ist es nicht.

Effizienz

Der Grad, zu dem Mittel verwendet werden im Verhältnis zu den erzielten Ergebnissen. Im Gegensatz zur Effektivität ist Effizienz ein relatives Maß und ändert sich fortlaufend aufgrund der Erwartungshaltung der Kunden. Allerdings gibt es auch Benutzungsschnittstellen, die sich seit Jahrzehnten nicht ändern und daher auch heute noch als ausreichend effizient gelten (bspw. die Google-Suchseite). Effizienz ist jedoch ein subjektives Maß und erfordert daher im Gegensatz zur Effektivität einen klaren Schwellenwert, um mess- bzw. vergleichbar zu sein. Diese sind üblicherweise als Abnahmekriterien bei den Nutzungsanforderungen hinterlegt (siehe Abschnitt 4.2.1).

Zufriedenheit

Der Grad, zu dem die Bedürfnisse der Benutzer zufriedengestellt werden, wenn das Produkt oder System in einem definierten Nutzungskontext eingesetzt wird.

Zufriedenheit der Benutzer ist gegeben, wenn der Benutzer während der Interaktionen mit der Benutzungsschnittstelle zur Erreichung seines Ziels keine negativen Wahrnehmungen verspürt hat. Um die Zufriedenheit der Benutzer effektiv bewerten zu können, ist jedoch eine Benutzerbefragung mit einem Fragebogen zwingend notwendig [Geis & Polkehn 18]. Fragebogen ist ein sehr formaler Begriff, ganz so formal wird es üblicherweise nicht gehandhabt. Jedwede Rückfrage mit einer einfachen Skala (bspw. die Sterne bei Amazon oder einfach nur Daumen-hoch/Daumen-runter-Button bzw. Gefällt-mir/Gefällt-mir-nicht-Button) und optionaler Begründung für die Bewertung ist in gewisser Weise ein Fragebogen.

Risikofreiheit

Der Grad, zu dem eine Komponente oder ein System das potenzielle Risiko für den wirtschaftlichen Status, Lebewesen, Gesundheit oder die Umwelt mindert. Die beiden unten genannten Fallbeispiele zu Risikofreiheit zeigen deutlich, dass Nutzungsqualität mehr als nur Gebrauchstauglichkeit ist. Die Sicherheit personenbezogener Daten ist ein essenzielles Merkmal für die Nutzungsqualität und durch die Anforderungsanalyse für die IT-Sicherheit explizit gemacht. Die Bedeutung der Risikofreiheit für die Gebrauchstauglichkeit hingegen wird üblicherweise während der Nutzungsanforderungsanalyse bestimmt und ist durch das Gebrauchstauglichkeitsteilmerkmal Benutzerfehlerschutz reflektiert.

Kontextüberdeckung

Der Grad, zu dem ein Produkt oder System mit Effektivität, Effizienz, Risikofreiheit und Zufriedenheit sowohl in einen definierten Nutzungskontext eingesetzt werden kann als auch in Nutzungskontexten, die über die a priori definierten Kontexte hinausgehen. Der Nutzungskontext von CA-Cockpit wäre ausreichend überdeckt, wenn Kunden mit den gängigen Mobilfunkgeräten, Tablets und PCs in der Lage sind, die Geschäftsprozesse erfolgreich und zufrieden abzuschließen.

Fallbeispiel

Analyse der Nutzungsqualitätsmerkmale

Effektivität

Für den Geschäftsprozess *GP-02b Handyvertrag abschließen (mit Laufzeit)* wäre eine Benutzungsschnittstelle effektiv, wenn der Kunde den für ihn passenden Handyvertrag über CA-Cockpit gekauft hat.

Für den Geschäftsprozess *GP-03 Handyvertrag verwalten* wäre eine Benutzungsschnittstelle effektiv, wenn die Kunden ihr spezifisches Ziel im

Mitgliederbereich erreicht haben.

Effizienz

Für den Geschäftsprozess *GP-02b Handyvertrag abschließen (mit Laufzeit)* wäre eine Benutzungsschnittstelle effizient, wenn der Kunde den für ihn passenden Handyvertrag aus einer Vielzahl möglicher Tarifkonfigurationen mit wenigen Interaktionen identifizieren kann.

Für den Geschäftsprozess *GP-03 Handyvertrag verwalten* wäre eine Benutzungsschnittstelle effizient, wenn die Kunden bei der Verwaltung ihrer laufenden Verträge keine überflüssigen Schritte ausführen müssen.

Zufriedenheit

Für den Geschäftsprozess *GP-02b Handyvertrag abschließen (mit Laufzeit)* wäre ein Kunde zufriedengestellt, wenn er einem Bekannten von der Einfachheit des Bestellprozesses bei *CA-Cockpit* berichtet, sodass dieser Bekannte eventuell seinen nächsten Handytarif über *CA-Cockpit* abschließt.

Für den Geschäftsprozess *GP-03 Handyvertrag verwalten* wäre ein Kunde zufriedengestellt, wenn er bei der Bewertung des Mitgliederbereichs angibt, dass er für die Verwaltung seiner Verträge den Mitgliederbereich der telefonischen Kundenbetreuung durch einen Callcenter-Agenten vorzieht.

Risikofreiheit

Für den Geschäftsprozess *GP-02b Handyvertrag abschließen (mit Laufzeit)* wäre eine Benutzungsschnittstelle risikofrei, wenn der Kunde sichergehen kann, dass der Vertrag erst wirklich dann zustande kommt, wenn er die Allgemeinen Geschäftsbedingungen bestätigt hat und den Vertrag verbindlich abschließt. Das Risiko eines unbeabsichtigten Vertragsabschlusses muss unter allen Umständen minimiert werden.

Für den Geschäftsprozess *GP-03 Handyvertrag verwalten* wäre ein System risikofrei, wenn sensible personenbezogene Daten wie Bankverbindung, Kreditkartennummer oder aber auch die Benutzername-Passwort-Kombination des Kunden vor unberechtigten Zugriffen effektiv geschützt sind.

Nutzungsqualität und Qualitätsmerkmale in verschiedenen Standards

Gebrauchstauglichkeit und Nutzungsqualität teilen sich in der Norm ISO 25010 dieselbe Definition, dennoch sind sie keine synonymen Konzepte, sondern eher komplementär zu betrachten. In der Definition werden die Adjektive *effizient*, *effektiv* und *zufriedenstellend* aufgeführt, diese aber nicht als Teilmerkmal von Gebrauchstauglichkeit, sondern von Nutzungsqualität gelistet. Wie passt dies

zusammen, und auch interessant, wie passt die Nutzungsqualität generell mit den Qualitätsmerkmalen zusammen?

Um diese Fragen beantworten zu können, hilft ein Blick in die Norm ISO 9241-220 [ISO 9241-220], die sich mit Prozessen »... zur Ermöglichung, Durchführung und Bewertung menschenzentrierter Gestaltung für interaktive Systeme in Hersteller- und Betreiberorganisationen« beschäftigt. In dieser Norm wird das Modell der Nutzungsqualität durch die Teilmerkmale *Gebrauchstauglichkeit*, *Barrierefreiheit*, *Benutzererlebnis* und *Vermeidung von Schäden durch die Nutzung* (was synonym zur *Risikofreiheit* des Modells der Nutzungsqualität nach ISO 25010 ist) beschrieben. Gebrauchstauglichkeit ist also ein Teilmerkmal der Nutzungsqualität nach ISO 9241-220 und gleichzeitig ein Qualitätsmerkmal nach ISO 25010.

In dem Buch »Praxiswissen User Requirements« [Geis & Polkehn 18] werden die beiden Standards zu einem gewissen Grad integriert, indem die verbleibenden ISO-25010-Qualitätsmerkmale (also alle außer Gebrauchstauglichkeit) als »Technische Qualität« beschrieben werden. Die Nutzungsqualität beschreibt also die Qualität, die vom Benutzer durch die Interaktion mit dem System aktiv wahrgenommen wird, während die technische Qualität die Grundlagen dafür schafft, dass überhaupt etwas wahrgenommen werden kann. Die Nutzungsqualität beschreibt also die Wahrnehmung des Systems »über der Haube«, während die technische Qualität die Fähigkeiten des Systems »unter der Haube« beschreiben (vgl. Abb. 4–5).

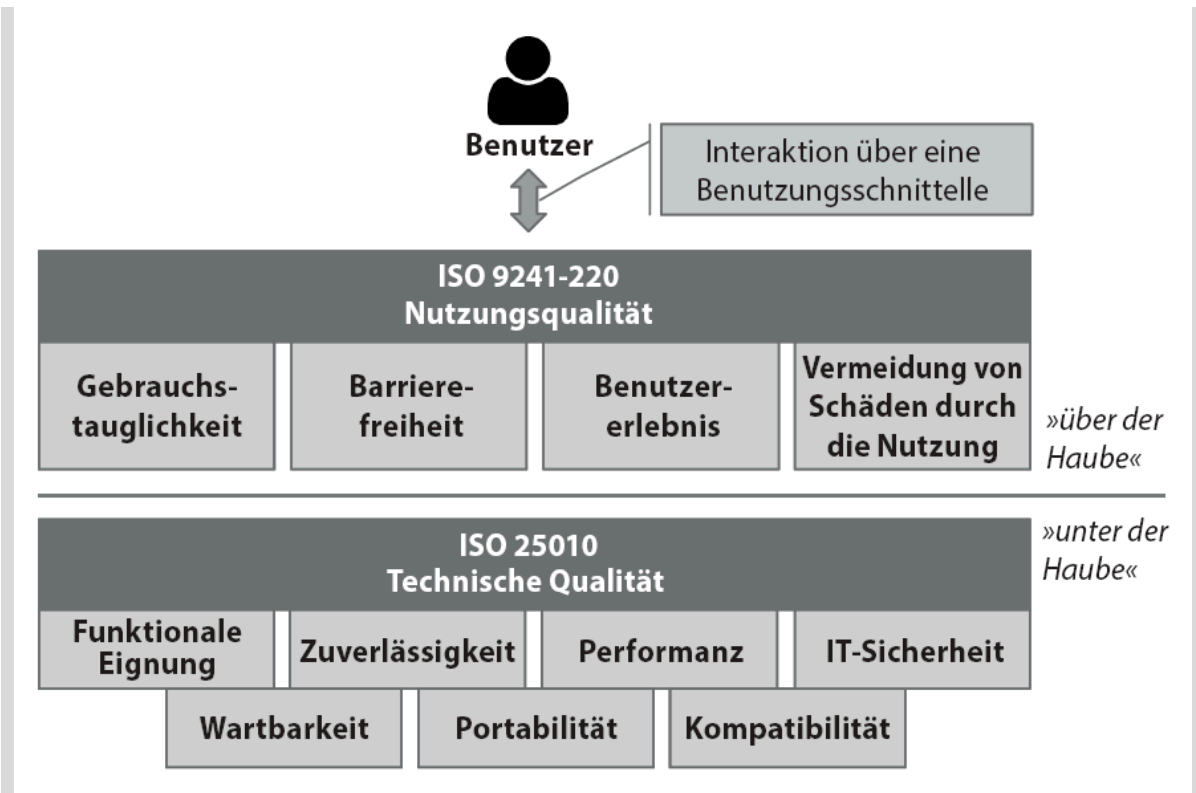


Abb. 4-5

Integration von Nutzungsqualität und technischer Qualität nach Geis [Geis & Polkehn 18]

Aus dieser Zweiteilung der Qualität in die Komponenten Nutzungsqualität und technische Qualität resultiert auch die Tatsache, dass die Qualitätsniveaus beider Komponenten schlussendlich zur Gesamtqualität des Systems beitragen. Ist die technische Qualität eines Systems hoch, die Nutzungsqualität aber gering, ist die Gesamtqualität des Systems eher gering – dies gilt auch andersherum.

Die Anforderungen an den Nutzungskontext werden während der Nutzungskontextanalyse erhoben. Die Überdeckung des Nutzungskontextes von CA-Cockpit

Überdeckung des Nutzungskontextes

kann beispielsweise erfüllt werden, indem bei der Ausgestaltung der Benutzungsschnittstelle des Frontends ein sogenanntes *responsive design* realisiert wird, also eine automatische Anpassung der Anordnung und Darstellung der Benutzungsschnittstelle je nach Displaygröße. Nutzungskontexte setzen sich aus unterschiedlichen Eigenschaften zusammen (Abb. 4-6).

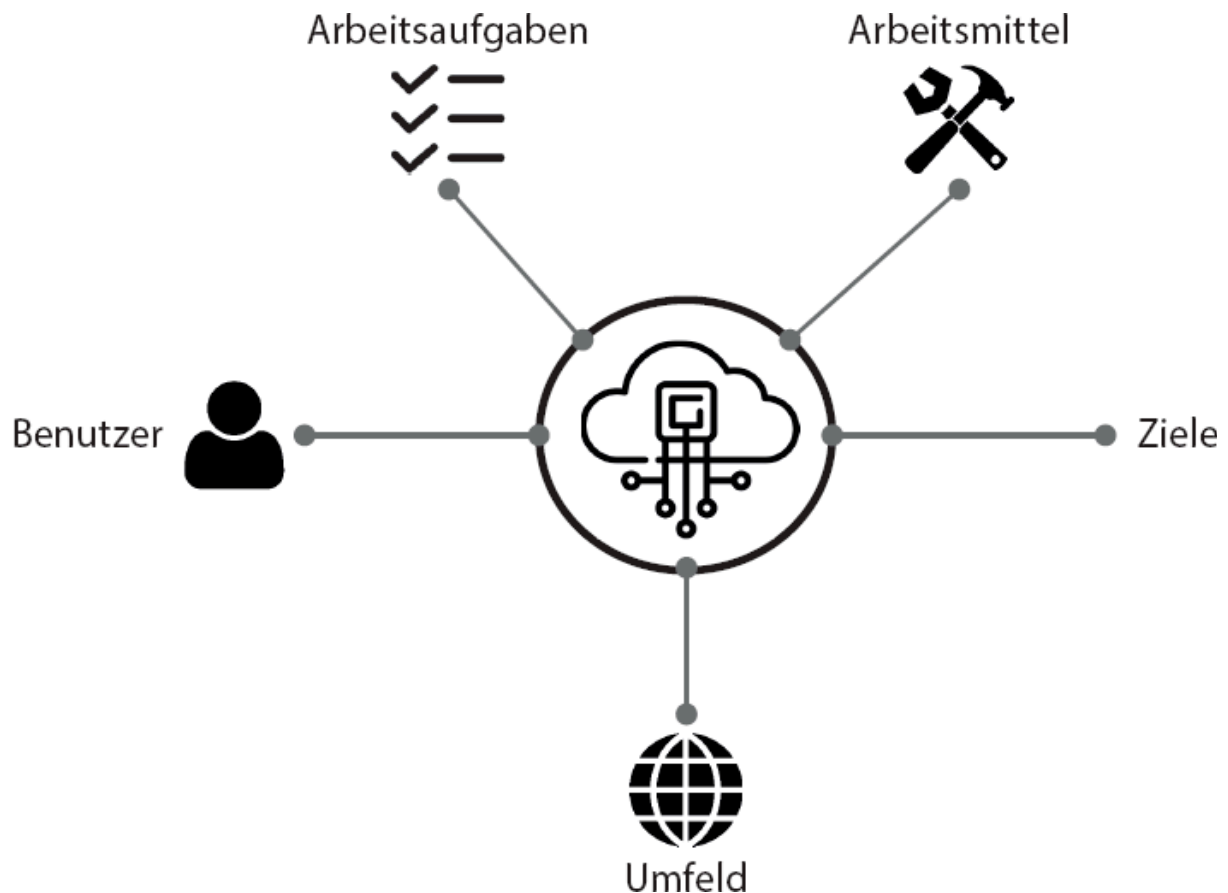


Abb. 4-6
Eigenschaften des Nutzungskontextes

Nutzungskontexte werden nach der Norm ISO 9241⁴ [ISO 9241-11] (von der sich viele der Definitionen und Konzepte in ISO 25010 ableiten) durch die folgenden Eigenschaften abgebildet:

Benutzer

Der Benutzeraspekt unterscheidet zwischen repräsentativen, nicht repräsentativen sowie indirekten Benutzern. Ferner spielen Fertigkeiten, Wissen und Erfahrungen der repräsentativen Benutzer eine Rolle – Erfahrungen mit der Arbeitsaufgabe oder dem Produkt bzw. Softwaresystem. So kann es sein, dass ein Expertensystem durchaus auch für Benutzer vorgesehen wird, die weniger Erfahrungen mit der Arbeitsaufgabe besitzen. Bei solchen Systemen können die Benutzer üblicherweise zwischen verschiedenen vordefinierten und auch benutzerdefinierten Komplexitätsprofilen auswählen. Jedes Komplexitätsprofil ist dann mit einer bestimmten Ausgestaltung der Benutzungsoberfläche verknüpft, sodass einige der eher komplexen Funktionen ausgeblendet werden. Durch diese Flexibilität können die Werkzeuge in verschiedenen Nutzungskontexten eingesetzt werden und erzielen damit in den jeweiligen Kontexten eine hohe Nutzungsqualität. Weitere Charakteristika des

Benutzeraspekts betreffen die persönlichen, psychischen und physischen Merkmale der Benutzer.

Ziele

Die Ziele beschreiben, was die Benutzer durch Anwendung des Produkts oder des Softwaresystems erreichen wollen. Ziele werden durch Arbeitsaufgabe und Teilaufgaben erreicht.

Arbeitsaufgaben

Arbeitsaufgaben werden für die Bestimmung des Nutzungskontextes häufig in Form von Aufgabenmodellen beschrieben. Das Aufgabenmodell zerlegt eine Kernaufgabe in eine Menge von Teilaufgaben, die zur Erledigung der Kernaufgabe beitragen. Eine Kernaufgabe durchläuft dabei immer die folgenden Schritte: Planung, Vorbereitung, Durchführung, Ergebnisbewertung, Ergebnisweitergabe. Der Nutzungskontext variiert bezüglich der Arbeitsaufgaben hinsichtlich der Aufgabenerlegung, Aufgabenhäufigkeit, Aufgabendauer, Handlungsspielräume bei der Erfüllung der Aufgabe, physische und mentale Anforderungen, die die Aufgaben erfordern, der Auswirkungen von Fehlern bei der Aufgabendurchführung, der Aufgabenergebnisse u. v. m.

Arbeitsmittel

Unter Arbeitsmitteln werden jedwede Formen von Hardware, Software und weiterer Materialien zusammengefasst, die dazu beitragen, die Aufgaben zu erfüllen bzw. die für die Benutzung des Produkts oder Softwaresystems erforderlich sind.

Physische und soziale Umgebung

Die physische und soziale Umgebung, in der ein Produkt oder Softwaresystem eingesetzt wird, bestimmt schlussendlich, welche externen Einflussfaktoren auf die Benutzung einwirken. Ein Polarforscher, der mit dicker Kälteschutzausrüstung ein Messgerät bedienen bzw. konfigurieren muss, hat mit anderen Einflussfaktoren des Umfelds zu tun als ein im Sitzen arbeitender Flugsicherungslotse, der die Einflugrouten ankommender Flugzeuge koordinieren muss. Je sicherheitskritischer die Arbeitsaufgabe und risikoreicher das soziale Umfeld, desto höher sollte das Merkmal *Risikofreiheit* der Nutzungsqualität priorisiert werden.

Produkte und Softwaresysteme, die in einem Nutzungskontext eine hohe Nutzungsqualität aufweisen, können für einen anderen Nutzungskontext gänzlich ungeeignet sein. Die Bewertung der Nutzungsqualität ist für den Abnahmetest von besonderer Bedeutung, da im Abnahmetest das abzunehmende Produkt

idealerweise in seiner späteren Produktivumgebung evaluiert wird. Es lässt sich also hoffentlich einfach bewerten, ob das abzunehmende Produkt im Nutzungskontext eine hohe Nutzungsqualität aufweist.

4.1.3 Nicht funktionale Qualitätsmerkmale im Abnahmetest

Wie bereits in Abschnitt 1.2 erläutert liegt der Fokus dieses Buches vor allem auf dem Benutzerabnahmetest, dem vertraglichen/regulatorischen Abnahmetest sowie dem Alpha- und Beta-Test. Aus diesem Grund gibt es Qualitätsmerkmale, die für einige Arten des Abnahmetests wichtiger sind als für andere. So sind die Qualitäts(teil)merkmale Installierbarkeit, Kompatibilität, Austauschbarkeit, Notfallwiederherstellung und Fehlertoleranz, einige Merkmale der Performanz (wie Verhalten bei Überlast, Skalierbarkeit oder Ressourcenverbrauch), Integrität, Softwarereife und Verfügbarkeit eher im betrieblichen Abnahmetest relevant. Die Qualitätsmerkmale Gebrauchstauglichkeit, einige Aspekte der Performanz (wie Lastverhalten) sowie IT-Sicherheit sind dagegen eher im Benutzer- und regulatorischen/vertraglichen Abnahmetest relevant. Nachfolgend werden diese eben aufgezählten, für den Benutzerabnahmetest relevanten nicht funktionalen Qualitätsmerkmale ausführlicher diskutiert.

Im Vorgängerstandard ISO 9126 gab es zu jedem Qualitätsmerkmal auch noch das Teilmerkmal Konformität, also die Einhaltung und Übereinstimmung mit gesetzlichen und regulatorischen Einschränkungen. Dieses wurde jedoch in ISO 25010 entfernt, da Konformität eher eine Eigenschaft der Systemanforderungen ist als ein Aspekt der Qualität. Schlussendlich basieren diese Systemanforderungen auf Qualitätsmerkmalen, sodass Konformität implizit auch für jedes Qualitätsmerkmal gegeben sein muss.

4.2 Gebrauchstauglichkeit und Benutzererlebnis

Gebrauchstauglichkeit und Benutzererlebnis sind zentrale Konzepte, wenn es um den Entwurf von Mensch-Maschine-Schnittstellen geht, die eine hohe Nutzungsqualität aufweisen sollen.

4.2.1 Gebrauchstauglichkeit

Gebrauchstauglichkeit (Usability) ist das Ausmaß, in dem ein Softwareprodukt durch bestimmte Benutzer in einem bestimmten Nutzungskontext genutzt werden kann, um bestimmte Ziele *effektiv*, *effizient* und *zufriedenstellend* zu erreichen [ISTQB Glossar]. Die drei Teilmerkmale Effektivität, Effizienz und Zufriedenheit sind gleichzeitig auch Merkmale des Modells der Nutzungsqualität. Tatsächlich

definiert ISO 25010 Gebrauchstauglichkeit als Untermenge der Nutzungsqualität [ISO 25010].

Wie bei allen Qualitätsmerkmalen, so ergibt sich die Qualität der Gebrauchstauglichkeit durch Erfüllung entsprechender Anforderungen. Diese werden Benutzeranforderungen (nach ISTQB® [ISTQB CTFL-UT]) oder Nutzungsanforderungen (nach UXQB [UXQB CPUX-F]) genannt. Sie repräsentieren Anforderungen an die Nutzung eines Produkts, die die Grundlage für die Gestaltung und die Usability-Evaluierung der Mensch-Maschine-Schnittstelle eines interaktiven Systems bildet (nach UXQB).

Mangelnde Gebrauchstauglichkeit kann zu Frustration, Verweigerung der Verwendung der Software und in den kritischsten Fällen zu Verletzungen oder zum Tod des Benutzers führen. Die Mensch-Maschine-Schnittstelle von industriellen Schneidemaschinen ist beispielsweise darauf ausgelegt, den Benutzern dieser Schneidemaschinen die maximale Risikofreiheit während der Nutzung angedeihen zu lassen. Dies erfolgt zumeist durch verschiedene konstruktive bzw. funktionale Sicherheitsmechanismen wie einrastende Schutzvorrichtungen oder mehrere Knöpfe, die gedrückt und gehalten werden müssen, damit der Schneidprozess überhaupt erst ausgelöst werden kann. Damit soll unter allen Umständen verhindert werden, dass die Benutzer während der Benutzung des Systems verletzt werden.

Die Teilmerkmale der Gebrauchstauglichkeit nach ISO 25010 (siehe Abb. 4–3) bestimmen die einzelnen Aspekte, die bei der Interaktion eines Benutzers bzw. einer Benutzergruppe mit dem Produkt zur Qualität beitragen:

Teilmerkmale der
Gebrauchstauglichkeit

Erkennbare Angemessenheit

Der Grad, zu dem Benutzer erkennen können, ob eine Komponente oder ein System für ihre Bedürfnisse angemessen ist. Eine hohe erkennbare Angemessenheit fördert beim Benutzer ein zuversichtliches Gefühl hinsichtlich der Interaktion mit dem Produkt. Erkennbare Angemessenheit geht daher stark einher mit dem funktionalen Teilmerkmal *Funktionale Angemessenheit* (definiert als: *Der Grad, zu dem Benutzer erkennen können, ob eine Komponente oder ein System für ihre Bedürfnisse angemessen ist*). Während also die funktionale Angemessenheit bewertet, ob das Produkt hinsichtlich der zu erwarteten Aufgaben die entsprechenden Funktionen anbietet, bezieht sich die erkennbare Angemessenheit darauf, wie dem Benutzer die Funktionalität des Produkts dargeboten wird bzw. ob der Benutzer erkennt, ob er seine Aufgaben mit dem Produkt erledigen kann. Produkte können demnach ein hohes Maß an funktionaler Angemessenheit innehaben, aber nur ein Mindestmaß an

erkennbarer Angemessenheit (was wahrscheinlich zum Scheitern bzw. zur Ablehnung des Produkts führt).

Erlernbarkeit

Der Grad, zu dem eine Komponente oder ein System durch bestimmte Anwender genutzt werden kann, um bestimmte Ziele des Lernens zufriedenstellend und risikofrei in einem spezifizierten Nutzungskontext zu erreichen. Erlernbarkeit zielt darauf ab, den Benutzer bei der Interaktion mit dem Produkt über die Benutzungsschnittstelle anzuleiten, damit der Benutzer sein Wissen über die Benutzung des Produkts stetig erweitern kann. Dazu zählen mitunter so einfache Dinge wie Tooltips, Hilfsdialoge oder Tutorials.

Operabilität: *Der Grad, zu dem eine Komponente oder ein System Eigenschaften aufweist, die einen einfachen Betrieb und eine einfache Steuerung ermöglichen.*

Die Operabilität erweitert die erkennbare Angemessenheit dahingehend, dass sie dem Benutzer eine effiziente und effektive Nutzung ermöglicht. Während die erkennbare Angemessenheit in erster Linie Vorfremde oder Skepsis bei den Benutzern auslöst, werden diese Empfindungen durch die Operabilität entweder verstärkt oder abgeschwächt. Operabilität ist ebenfalls eng mit der Funktionalität eines Produkts verknüpft. Dabei ist es nicht immer notwendig, dass tatsächlich auch jedwedes kleinste Detail der Funktionalität dargeboten wird. Eine Überfrachtung der Mensch-Maschine-Schnittstelle mit Funktionalität schränkt mitunter die Operabilität ein und führt zum Scheitern des Produkts. Als Beispiel eines im Vergleich zu aktuellen Smartphones gescheiterten Produkts war die Verbreitung und Akzeptanz damaliger Handhelds als digitale Assistenten. Diese Geräte boten bereits ein hohes Maß an Funktionalität an, die man seinerzeit auch von PCs und Laptops kannte. Benutzer konnten schon damals mit Tabellenkalkulationsprogrammen, Textverarbeitungsprogrammen etc. mit einem Handheld und einem Eingabestift arbeiten, Kalendereinträge vornehmen, ja sogar Spiele konnte auf diesen Geräten gespielt werden. Den Durchbruch zum Massenmarkt schafften diese Vorgänger unserer heutigen Smartphones jedoch nicht. Sie haben schlichtweg zu wenig Benutzer erreicht, bzw. die Arbeit mit diesen Geräten hat einfach keinen Spaß gemacht. Es fühlte sich umständlich und irgendwie nicht richtig an, Teile seiner Arbeit auf diesen Handhelds zu erledigen.

Benutzerfehlerschutz

Der Grad, zu dem eine Komponente oder ein System den Benutzer davor schützt, Fehler zu begehen. Benutzerfehlerschutz kann eine essenzielle, wenn nicht die essenzielle Rolle spielen, wenn Produkte oder Softwaresysteme den regulatorischen oder vertraglichen Abnahmetest nicht bestehen. Die Wichtigkeit,

fehlerhafte Eingaben des Benutzers zu erkennen und ggf. gar nicht erst zuzulassen, variiert mit der Kritikalität des Nutzungskontextes. Vereinfachend kann man sagen, dass der Benutzerfehlerschutz an Stellenwert gewinnt, je kritischer die Auswirkungen möglicher Fehleingaben durch die Benutzer sind. Aus diesem Grund verwundert es auch nicht, dass Mensch-Maschine-Schnittstellen in diesen Nutzungskontexten, beispielsweise im Cockpit eines Flugzeugs oder in Zügen, hinsichtlich ihrer Operabilität vergleichsweise einfach gehalten sind, wenngleich sie für die meisten Benutzergruppen sehr komplex wirken. Die Vielzahl von einzelnen Knöpfen und Schaltern ohne Display oder weiteren Eingabemöglichkeiten sind ein wesentlicher Bestandteil, um Benutzerfehleringaben vorzubeugen. So steht jeder Schalter für eine bestimmte Form der (technisch gesehen bereits parametrisierten) Eingabe des Benutzers. Dieser effektive und äußerst effiziente Schutz vor Fehleingaben kommt mit einer wenig ästhetischen Gestaltung daher, diese hat jedoch gegenüber dem Fehlerschutz eine nachrangige Rolle. Aber auch für weniger sicherheitskritische Produkte kann der Benutzerfehlerschutz zu Frustration oder Freude während der Benutzung des Produkts führen. Benutzerfehlerschutz ist im ISO-25010-Modell der Nutzungsqualität (siehe Abb. 4–4) durch das Merkmal *Risikofreiheit* reflektiert.

Ästhetik der Benutzungsschnittstelle

Der Grad, zu dem eine Benutzungsschnittstelle eine angenehme und zufriedenstellende Interaktion für den Benutzer ermöglicht. Die Ästhetik einer Benutzerschnittstelle kann bis zu einem gewissen Grad eine schlechte Erlernbarkeit oder Operabilität kaschieren. Menschen, denen Design wichtiger ist als Funktionalität, fühlen sich von ästhetischen Aspekten deutlich mehr angezogen als Menschen, denen es eher auf eine einfache Erreichbarkeit der angemessenen Funktionalität ankommt. Die Wichtigkeit der Ästhetik der Mensch-Maschine-Schnittstelle wurde durch das Aufkommen mobiler Geräte wie Smartphones und Tablets neu geordnet. Das Erscheinungsbild einer App entscheidet heutzutage oftmals viel schneller über Akzeptanz oder Ablehnung als die dargebotene Funktionalität.

Zugänglichkeit

Der Grad, zu dem eine Komponente oder ein System von Menschen mit den unterschiedlichsten Eigenschaften und Fähigkeiten gebraucht werden kann, um ein gegebenes Ziel in einem gegebenen Nutzungskontext zu erreichen. Dem Teilmerkmal Zugänglichkeit⁵ kommt in dieser Liste eine Sonderstellung zu. Zugänglichkeit ist ein essenzielles Teilmerkmal, da hier die vorweg definierten Teilmerkmale ebenfalls subsumiert sind. Zugänglichkeit besagt, dass Menschen

mit besonderen Bedürfnissen einen angemessenen Zugang zu dem Produkt erhalten sollen. Dies können Menschen mit körperlichen Beeinträchtigungen sein (mit verminderter Seh-/Hör-/Sprechfähigkeit), aber auch Benutzer, die beispielsweise das Produkt mithilfe von externen Steuergeräten benutzen müssen.

Ein Beispiel aus der Welt des modellbasierten Testens

Testgeneratoren, wie sie vor gut zehn Jahren vermehrt durch die Popularität des modellbasierten Testens auf den Markt strömten, operieren üblicherweise auf Verhaltensbeschreibungen des zu testenden Systems. Ein Testgenerator ist also auf die Erstellung eines syntaktisch korrekten Modells angewiesen. Die meisten Testgeneratoren boten (und bieten) jedoch keine dedizierte Modellierungsumgebung an. Aus Marktsicht ist dies absolut nachvollziehbar, kann man doch so argumentieren, dass der Testgenerator auf den Modellen verschiedener Modellierungswerkzeuge arbeiten kann, solange diese nur syntaktisch korrekt im Sinne des Testgenerators⁶ sind. Dies hat jedoch zur Folge gehabt, dass die Testmodellierer ohne zielführende Unterstützung seitens des Modellierungswerkzeugs die für den Testgenerator syntaktisch korrekten Modelle erzeugen mussten. Dies wiederum führte dazu, dass die Anzahl der Benutzerfehler üblicherweise sehr hoch war, ehe man sich daran gewöhnt hatte, welche Konstrukte der Modellierungssprache man wie verwenden konnte. Ein Fakt, der leider nicht dazu beigetragen hat, dass sich modellbasiertes Testen trotz seiner Potenziale großflächig in der Praxis durchgesetzt hat.

4.2.2 Benutzererlebnis (UX)

Benutzererlebnis (UX – User eXperience) erweitert den Begriff der Gebrauchstauglichkeit um ästhetische und emotionale Faktoren wie eine ansprechende, wünschenswerte Gestaltung, Aspekte der Vertrauensbildung oder die Zufriedenheit bei der Verwendung (z. B. Vergnügen, Komfort). Es beschreibt die »Wahrnehmungen und Reaktionen einer Person, die aus der tatsächlichen und/oder der erwarteten Benutzung eines Softwareproduktes resultieren« [ISTQB Glossar]. Gebrauchstauglichkeit hingegen konzentriert sich vorrangig auf das subjektive Empfinden des Benutzers während der Interaktion mit dem Produkt bzw. dem Softwaresystem (siehe Abb. 4–7).

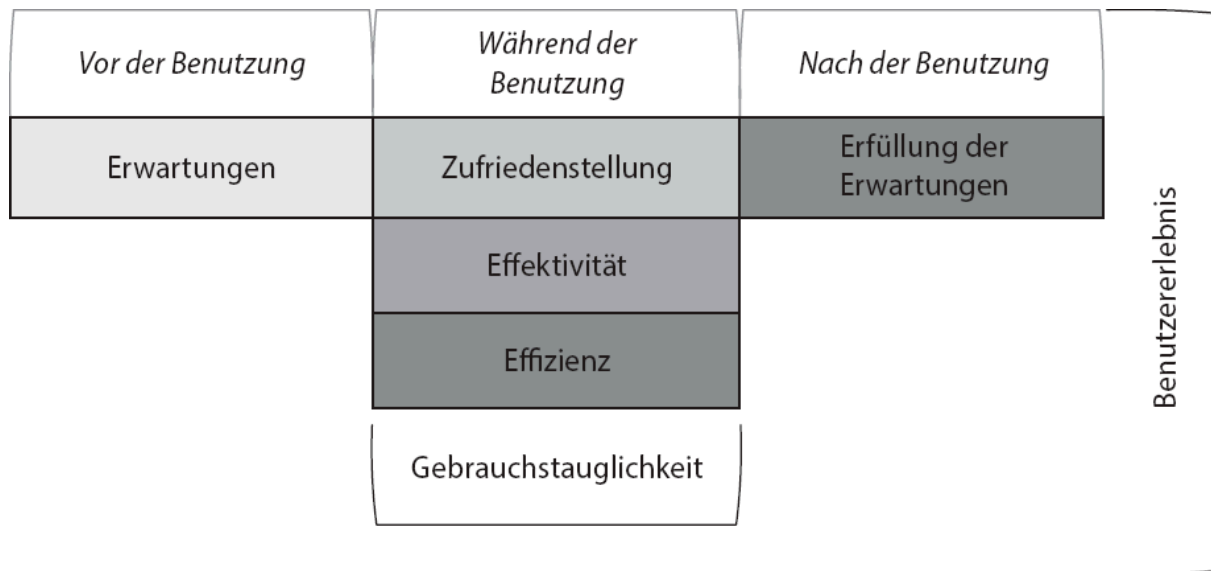


Abb. 4-7

Beziehung zwischen Gebrauchstauglichkeit und Benutzererlebnis

Benutzererlebnis charakterisiert sich durch die folgenden Teilmerkmale [ISTQB CTFL-UT]:

- Emotionen
- Überzeugungen
- Vorlieben
- Wahrnehmungen
- Physiologische und psychologische Reaktionen
- Verhalten und Erfolge

Zudem wird das Benutzererlebnis durch folgende Aspekte beeinflusst:

- Markenimage (d. h. das Vertrauen des Nutzers in den Hersteller)
- Gestaltung (d. h. Erscheinungsbild des Softwareprodukts, einschließlich Verpackung und Dokumentation)
- Funktionalität
- Leistungsmerkmale des Softwareprodukts
- Interaktives Verhalten
- Hilfe/Unterstützung durch das Softwareprodukt, einschließlich Hilfesystem, Systemsupport und Trainingsunterstützung
- Erlernbarkeit
- Individueller Nutzerstatus (physiologisch/psychologisch), der sich aus gesammelten Erfahrungen, Einstellungen, Fähigkeiten, Persönlichkeit, Bildung

und Intelligenz ergibt
Nutzungskontext

Der Nutzungskontext, in dem das System verwendet wird, hat einen starken Einfluss auf das Benutzererlebnis, da sich dieses aufgrund einer Reihe von Faktoren sehr stark unterscheiden kann. Hierzu gehört z. B. der Standort (z. B. der Benutzer sitzt am Schreibtisch, fährt Auto oder wandert), das Wetter (z. B. Sonne, Regen, Kälte), der Gesundheitszustand des Benutzers (z. B. Müdigkeit, Alter), die Umgebung (z. B. Stress, Lärm).

Fallbeispiel

Benutzererlebnis vor und nach der Benutzung

Eine Neuerung von *CA-Cockpit* ist die Möglichkeit, dass Kunden die Details ihres Handyvertrags selbst einsehen, verändern, sprich verwalten können. Faktoren, die das Benutzererlebnis noch vor der Benutzung wahrscheinlich positiv beeinflussen, jedoch nichts mit der Gebrauchstauglichkeit zu tun haben:

Die Möglichkeit, Anpassungen am Tarif ohne das Callcenter vornehmen zu können.

Die Möglichkeit, Rechnungen auch nach Monaten noch einsehen und ausdrucken zu können.

Faktoren, die das Benutzererlebnis nach der Benutzung wahrscheinlich negativ beeinflussen, jedoch nichts mit der Gebrauchstauglichkeit zu tun haben:

Wenn sich herausgestellt hat, dass tarifspezifische Anpassungen nicht in dem erwarteten Umfang möglich sind bzw. die erneute Kommunikation mit dem Callcenter notwendig wird.

Nutzungsanforderungen ergeben sich durch die spezifischen Erfordernisse, die für die Erreichung der jeweiligen Ziele im Nutzungskontext gegeben sind.

*Nutzungsanalyse und
Nutzungsanforderungen*

Erfordernisse sind ein Zwischenarbeitsprodukt, das helfen soll, die vorliegenden Ausprägungen der Nutzungskontextaspekte in entsprechende Nutzungsanforderungen zu transformieren. Ein Erfordernis ist dabei losgelöst von der technischen Lösung und entspricht daher im Grunde Anforderungen, die in einem Lastenheft aufgeführt sind. Die Nutzungsanforderungen werden aus den Erfordernissen abgeleitet, bringen diese mit anderen relevanten Anforderungsarten (bspw. der funktionalen und IT-Sicherheit) zusammen und

schaffen die Grundlage für die Evaluierung des Benutzererlebnisses und der Gebrauchstauglichkeit.

Um sowohl Erfordernisse als auch Nutzungsanforderungen zu identifizieren, können die vier Säulen der Anforderungsanalyse für das Benutzererlebnis und der Gebrauchstauglichkeit eingesetzt werden (Abb. 4–8).

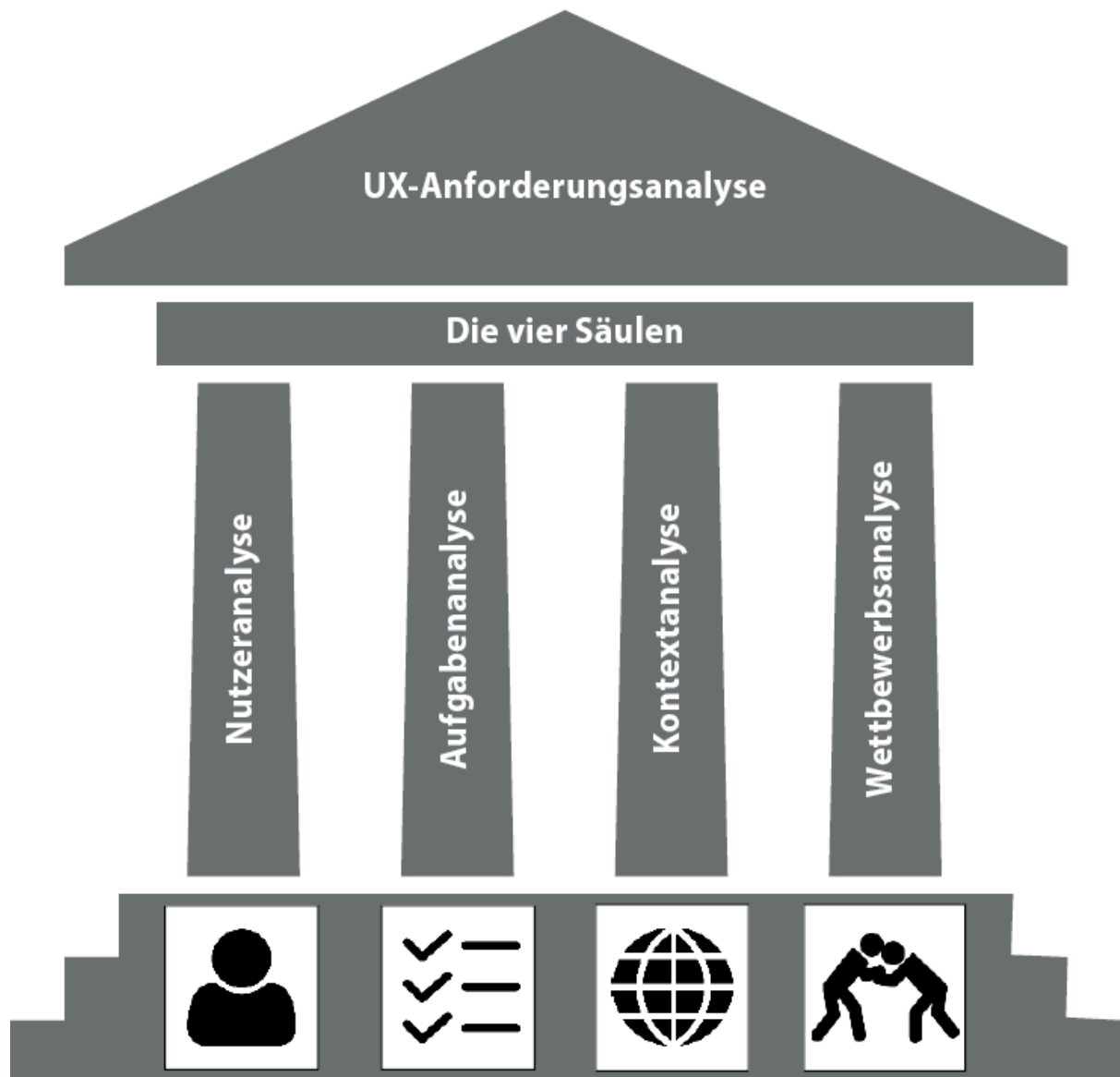


Abb. 4–8

Die vier Säulen der UX-Anforderungsanalyse

Diese vier Analysearten ähneln stark den Aspekten des Nutzungskontextes und führen auch zu ähnlichen Erkenntnissen. Auch wenn die Tester üblicherweise an der Analyse und der Spezifikation der Nutzungsanforderungen nicht oder nur in der Rolle eines Reviewers teilnehmen, hilft es Testern für die Spezifikation von Gebrauchstauglichkeits- und Benutzererlebnistest zu verstehen, auf welche Weise bestimmte Erfordernisse erhoben wurden:

Nutzeranalyse

Benutzer werden nach physischen und intellektuellen Merkmalen, technischen Fähigkeiten, betrieblichen Kenntnissen, sozioökonomischem und kulturellem Hintergrund kategorisiert. Businessanalysten können zu diesem Zweck auch auf Benutzergruppenprofile oder Personas zurückgreifen und sie als Modelle von Benutzern verwenden. Ein Benutzergruppenprofil ist eine verallgemeinerte Beschreibung einer Sammlung von Benutzern mit den gleichen oder ähnlichen persönlichen Eigenschaften und Nutzungskontext in Bezug auf die Benutzung des Produkts oder Softwaresystems. Ein anderes Wort für Benutzergruppenprofil ist Zielgruppe. Personas sind Beschreibungen fiktiver, aber realistischer Personen, die stellvertretend für eine bestimmte Benutzergruppe sind.

Aufgabenanalyse:

Benutzer haben in den meisten Fällen konkrete Aufgaben im Sinn, wenn sie ein Produkt oder Softwaresystem benutzen. In der Aufgabenanalyse geht es darum, diese konkreten Aufgaben zu identifizieren und zu formalisieren (z. B. anhand von Anwendungsfällen oder User Stories). Jede Aufgabenanalyse beruht darauf, eine Kernaufgabe in ihre einzelnen Teilaufgaben zu zerlegen, sodass Erfordernisse abgeleitet werden können, um ein optimiertes System oder Produkt zu entwerfen. Aufgabenanalyse zerfällt üblicherweise in zwei Teilbereiche:

- Die aktionsgetriebene Aufgabenanalyse adressiert die erforderlichen Handlungen der Benutzer, um die Aufgaben zu erfüllen (z. B. manuelle Tätigkeiten, Bewegung oder Objektmanipulation).
- Die kognitionsgetriebene Aufgabenanalyse untersucht hingegen die intellektuellen und mentalen Prozesse, die Benutzer bei der Erfüllung der Aufgaben durchlaufen. Diese beinhalten wichtige kognitive Aspekte der Entscheidungsfindung, des Problemlösens, der Aufmerksamkeit und des Gedächtnisses.

Kontextanalyse

Der Nutzungskontext hat, wie bereits erwähnt, einen maßgeblichen Einfluss auf die Nutzungsqualität, die Gebrauchstauglichkeit und das Benutzererlebnis. Äußere Bedingungen (z. B. Licht, Temperatur, Bewegung, Feuchtigkeit oder Staub), physische Bedingungen (z. B. Sitzen, Stehen, Liegen, Bewegen, Freihändigkeit) oder »psychologische« Bedingungen (z. B. Stresslevel, Motivation oder der Unterschied zwischen privatem und professionellem Gebrauch) werden dabei in Erfordernisse zerlegt und in Nutzungsanforderungen überführt, um die nachfolgenden Entwurfsschritte in eine bestimmte Richtung zu lenken. Geräte,

Plattformen und Formfaktoren (gerätespezifische Anzeige) werden ebenfalls als Teil des Nutzungskontextes betrachtet.

Wettbewerbsanalyse

Die Analyse eines bereits auf dem Markt angebotenen Konkurrenzprodukts eignet sich immer dann als kostengünstige Möglichkeit, Nutzungsanforderungen zu erheben, solange kein völlig neuartiges, disruptives Design angestrebt wird. Dies können erfolgreiche Lösungen aus ähnlichen oder sogar ganz anderen Branchen sein. Ziel ist es hier, zu verstehen, was Benutzer eines bestimmten Nutzungskontextes erwarten, was man ggf. besser machen kann und was man unbedingt vermeiden sollte.

Aufgrund allgemeiner menschlicher Einschränkungen und Voreingenommenheit (z. B. kognitive oder perzeptive Voreingenommenheit, Sehbehinderung, Unerfahrenheit) [ISTQB CTFL-UT] können einige Benutzer mit spezifischeren und manchmal herausfordernden Schwierigkeiten bei der Verwendung von Software oder Produkten konfrontiert sein, die als Teil der Geschäftslösung vorgesehen sind. Businessanalysten und Tester sollten beurteilen, ob Produkte oder Dienste für alle Benutzer zugänglich sind, indem sie diese Einschränkungen beim Entwurf von Abnahmekriterien und Testfällen berücksichtigen.

Fallbeispiel

Nutzungskontextanalyse des Geschäftsprozesses *GP-02b Handyvertrag abschließen (mit Laufzeit)*

Nutzeranalyse

Die Analyse der zu erwartenden Benutzer für den Geschäftsprozess *Handyvertrag abschließen (mit Laufzeit)* ergab drei repräsentative bzw. zu erwartende Benutzergruppen. Die Analyse wurde gemeinsam vom Businessanalysten und dem Marketing durchgeführt. Diese Benutzergruppen sind nachfolgend kurz umrissen.

Benutzergruppe 1

Benutzer dieser Gruppe sind Bestandskunden mit Prepaid-Tarifen, die diese in einen Handytarif mit Vertrag umwandeln möchten. Der Großteil dieser Benutzergruppe ist zwischen 20 und 30 Jahren alt, hat umfassende Erfahrungen mit Online-Shops und Online-Bestellvorgängen. Als Bestandskunden kennen diese Benutzer zudem auch die Geschäftsprozesse des Unternehmens, die üblichen Kommunikationsmöglichkeiten sowie das CI des Unternehmens.

Benutzergruppe 2

Benutzer dieser Benutzergruppe sind Neukunden, die bereits über Erfahrungen im Umgang mit Onlineshops und -Bestellvorgängen haben. Die Feinheiten der angebotenen Tarife sind jedoch überwiegend nicht bekannt. Diese Benutzergruppe zeichnet sich dadurch aus, dass sie viele verschiedene Angebote auf dem Markt vergleichen und dadurch bestrebt sind, möglichst einfach die angebotenen Tarife eines Unternehmens miteinander vergleichen zu können. Diese Benutzergruppe legt Wert auf eine möglichst große Flexibilität bei der Zusammenstellung der Tarifdetails, während die Laufzeit des Vertrages keine große Rolle bei diesen Überlegungen spielt. Die Altersspanne dieser Benutzergruppe liegt zwischen 30 und 50 Jahren.

Benutzergruppe 3

Benutzer dieser Benutzergruppe sind potenzielle Neukunden, die bislang kaum oder wenig Erfahrungen mit Onlineprozessen gemacht haben und Online-Bestellvorgängen grundsätzlich skeptisch gegenüberstehen. Die Altersspanne dieser Benutzergruppe ist 60+. Diese Benutzergruppe legt Wert auf möglichst übersichtlich vorkonfigurierte Tarife ohne besondere Details oder Konfigurationsmöglichkeiten.

Aufgabenanalyse

Das Ziel der Nutzung ist der Abschluss eines Handytarifs mit Vertrag zur vollsten Zufriedenheit der Benutzer. Um dieses Ziel zu erreichen, hat der Benutzer die Aufgabe, den Bestellprozess in Interaktion mit dem System durchzuführen und vertragsrechtlich (sprich erfolgreich) abzuschließen. Diese Aufgabe lässt sich in eine Menge von Teilaufgaben zerlegen, die sich ihrerseits wiederum in Teilaufgaben zerlegen ließen. Aus Gründen der Übersichtlichkeit wurde die weitere Dekomposition von Teilaufgaben nur für die Teilaufgabe »Finanzielle Details klären« illustriert (siehe Abb. 4–9).

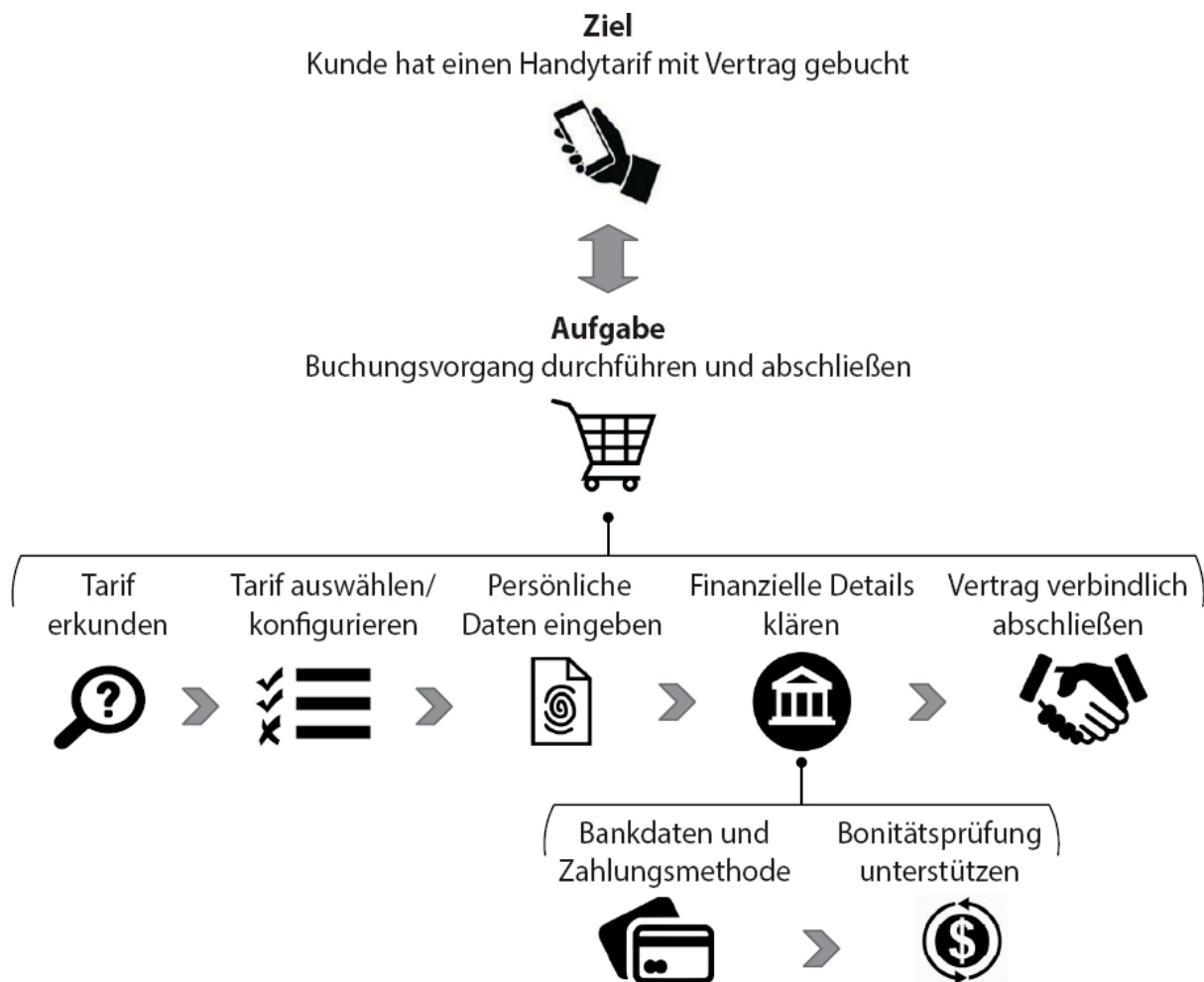


Abb. 4-9

Aufgabenzerlegung des Geschäftsprozesses GP-02b Handytarif abschließen (mit Laufzeit)

Kontextanalyse

Die Kontextanalyse wurde überwiegend vom Marketing durchgeführt. Die relevanten Aspekte der Kontextanalyse für den Geschäftsprozess sind die Arbeitsmittel, die zur Durchführung der Aufgaben und Teilaufgaben benötigt werden, sowie das Umfeld, in dem die Aufgaben und Teilaufgaben üblicherweise durchgeführt werden. Dazu wurde unter anderem auch ausgewertet, mit welchen Gerätetypen Kunden und Besucher das Webportal bislang besucht haben.

Arbeitsmittel

Die Auswertung der Nutzungsdaten aus dem Webportal hat ergeben, dass zu 80% mobile Endgeräte eingesetzt werden, um Tarife zu vergleichen und den Bestellvorgang durchzuführen. Weitere 15% entfallen auf PCs, die restlichen 5% sind Tablets.

Umfeld

Die Verweildauer der Benutzer/Besucher auf dem Webportal und die durchschnittlichen Zeiten beim Vergleich der Tarife und Durchführung eines Bestellprozesses weisen darauf hin, dass kein zeitlicher Drang bzw. Hektik bei den Benutzern/Besuchern vorliegt.

Wettbewerbsanalyse

Die Wettbewerbsanalyse wurde gemeinsam vom Marketing und dem Businessanalysten durchgeführt. Es wird nicht nach einem disruptiven Ansatz für die Gebrauchstauglichkeit gestrebt. Die durch Rückmeldungen von Bestandskunden bewertete Nutzungsqualität des Geschäftsprozesses *Handyvertrag abschließen* ist bereits hoch, sodass die Gebrauchstauglichkeitsaspekte des bestehenden Geschäftsprozesses *Handyvertrag abschließen (ohne Laufzeit)* für den neuen Geschäftsprozess *Handyvertrag abschließen (mit Laufzeit)* übernommen werden sollen.

Vereinzelte Rückmeldungen der Bestandskunden enthielten Beschwerden oder Empfehlungen, zwischen den verschiedenen Teilaufgaben des Buchungsprozesses einfacher navigieren zu können. Es wird daher empfohlen, die verschiedenen Teilaufgaben auf jeweils verschiedene Masken aufzuteilen und eine einfache Navigation zwischen den Masken zu ermöglichen. Vor dem Abschluss des Bestellprozesses sollen zudem sämtliche relevante Informationen abschnittsweise dem Benutzer präsentiert werden und ihm die Möglichkeit gegeben werden, über einen direkten Link zu der entsprechenden Maske der Teilaufgabe zu springen.

Erfordernisse und Nutzungsanforderungen

Die Ergebnisse der Nutzungskontextanalyse werden weiterhin verwendet, um konkrete Erfordernisse der tarifsuchenden Benutzer (kurz: Tarifsuchende) abzuleiten, aus denen sich im Anschluss konkrete Nutzungsanforderungen und Abnahmekriterien ergeben.

Erfordernisse

Erfordernis E-01

Ein Tarifsuchender muss so einfach wie möglich durch den Such- und Bestellprozess geleitet werden, um einen Handytarifvertrag abzuschließen.

Erfordernis E-02

Ein Tarifsuchender muss jederzeit zwischen unterschiedlichen Tarifoptionen wechseln können, um die jeweiligen Auswirkungen auf die Berechnung der Vertragssumme bei einem Tarifabschluss bewerten zu können.

Erfordernis E-03

Ein Tarifsuchender muss die Möglichkeit haben, verschieden konfigurierte Handytarife miteinander zu vergleichen, um stets den jeweils passendsten Handytarif für sich auswählen zu können.

Anforderung	
ID	A-38
Titel	Einfachheit des Bestellprozesses
Beschreibung	Das System muss dem Tarifsuchenden das Ausfüllen der relevanten Informationen der Teilaufgaben Persönliche Daten, Tarifkonfiguration, Bankverbindung und Bestellzusammenfassung sowie rechtliche Bestätigungen während des Bestellprozesses so einfach wie möglich machen.
Art	UX
Erfordernis	E-01
Abnahmekriterien	A-38-01 Ein Tarifsuchender der Benutzergruppe 1 muss den Bestellprozess bei einer zuvor bekannten Tarifkonfiguration innerhalb von 5 Minuten erfolgreich abschließen.
	A-38-02 Ein Tarifsuchender der Benutzergruppe 2 muss den Bestellprozess bei einer zuvor bekannten Tarifkonfiguration innerhalb von 10 Minuten erfolgreich abschließen.
	A-38-03 Ein Tarifsuchender der Benutzergruppe 3 muss den Bestellprozess bei einer zuvor bekannten Tarifkonfiguration durch Verwendung der Hilfsfunktionen innerhalb von 10 Minuten erfolgreich abschließen.
	A-38-04 Die einzugebenden Informationen für die verschiedenen Teilaufgaben des Bestellprozesses sollen auf separaten Bildschirmmasken dem Benutzer präsentiert werden und eine nahtlose Navigation zwischen den Benutzermasken ermöglichen.

	A-38-05	Die Zufriedenheitsbewertung der Tarifsuchenden soll nach Vertragsabschluss erfragt werden und mindestens durch eine Bewertungsnote 4 (auf einer 5-stelligen Bewertungsskala) nachweisbar sein.
--	----------------	--

Abnahmekriterien der Nutzungsanforderungen werden auch *quantitative Nutzungsanforderungen* genannt, während die funktionalen Beschreibungen der Aufgaben als *qualitative Nutzungsanforderungen* bezeichnet werden.

4.2.3 Gebrauchstauglichkeitstest

Anhand der erhobenen und spezifizierten Nutzungsanforderungen kann die Gebrauchstauglichkeit und das Benutzererlebnis evaluiert werden. Dabei konzentriert sich der Gebrauchstauglichkeitstest auf die Nutzungsqualitätsmerkmale Effizienz, Effektivität und Zufriedenheit.

Die Ziele der Gebrauchstauglichkeitsevaluierung sind:

- Beurteilen, ob die Nutzungsanforderungen erfüllt wurden
- Aufdecken von Gebrauchstauglichkeitsproblemen, damit diese behoben werden können
- Messen der Gebrauchstauglichkeit eines Produkts oder Softwaresystems

Die eingesetzten Verfahren zur Evaluierung der Gebrauchstauglichkeit werden in formative und summative Evaluierungen unterschieden. Formative (oder explorative) Evaluierung wird durchgeführt, um Schwierigkeiten bzw. Probleme mit der Gebrauchstauglichkeit zu verstehen und den Entwurf von Mensch-Maschine-Schnittstellen zu formen. Formative Bewertungen werden zumeist frühzeitig im Entwicklungsprozess anhand von Entwurfsdokumenten und Prototypen durchgeführt. Dazu kann es auch notwendig sein, den Nutzungskontext physisch in einem Testlabor nachzubauen, um zu sehen, ob die Operabilität effizient möglich ist. Labore und physische Prototypen werden vor allem in sicherheitskritischen, eingebetteten oder cyber-physischen Systemen eingesetzt.

Summative Evaluierung wird zu einem späteren Zeitpunkt im Entwicklungsprozess durchgeführt mit dem Ziel, die Implementierung der Mensch-Maschine-Schnittstelle gegen die Nutzungsanforderungen zu prüfen. Die summative Evaluierung ist quantitativ, d. h., der Schwerpunkt liegt auf der Messung von Effektivität, Effizienz oder Zufriedenheit des Benutzers bei der Interaktion mit dem Produkt oder Softwaresystem.

Die Evaluierung der Gebrauchstauglichkeit besteht aus drei Verfahren, die für unterschiedliche Ziele eingesetzt werden können [ISTQB CTFL-UT]:

Benutzerbefragungen

Eine Benutzerbefragung ist eine Usability-Evaluierung, bei der eine repräsentative Auswahl von Benutzern gebeten wird, ihre subjektive Bewertung in einem Fragebogen zu berichten, der ihre Erfahrungen bei der Interaktion mit einem Produkt oder Softwaresystem dokumentiert. Entsprechende Fragebogen können entweder standardmäßige oder proprietäre Fragebogen sein, es wird jedoch dringend empfohlen, einen Standardfragebogen zu verwenden, wenn dies möglich ist. Beispiele von Standardfragebogen sind System Usability Scale (SUS), Software Usability Measurement Inventory (SUMI) und Website Analysis and Measurement Inventory (WAMMI). Benutzerbefragungen eignen sich vor allem, um die Zufriedenheit der Benutzer nach der Benutzung des Produkts oder Softwaresystems zu bewerten.

Reviews auf Gebrauchstauglichkeit

Gebrauchstauglichkeitsreviews basieren häufig auf Benutzeranforderungen, geltenden Richtlinien für die Benutzungsschnittstelle, Gebrauchstauglichkeitsstandards und/oder frühere Erfahrungen von Nutzern mit Gebrauchstauglichkeitsproblemen. Bei einem Gebrauchstauglichkeitsreview werden häufig Heuristiken verwendet. [Rubin & Chisnell 08] [Nielsen 94]. Eine Gebrauchstauglichkeitsheuristik ist eine Liste mit »Daumenregeln«, die helfen, die Gebrauchstauglichkeit von Produkten und Softwaresystemen zu bewerten. Eine Heuristik hat den Zweck, dem Gutachter eine verlässliche und nützliche Hilfestellung für die Evaluierung der Gebrauchstauglichkeit eines Softwareprodukts zu geben. Reviews können sowohl formativ als auch summativ eingesetzt werden. Generell sind Reviews einfacher durchzuführen, wenn man die Mensch-Maschine-Schnittstelle sehen kann. Ein Review sollte daher durch Prototyping oder Wireframing (oder eine andere Form der Visualisierung) unterstützt werden. Dazu zählen auch Mensch-Maschine-Schnittstellen, die mit Stift und Papier skizziert wurden, wie es beispielsweise in einer agilen Vorgehensweise häufig für die Sprint-Planung der Fall ist.

Usability-Tests

Gebrauchstauglichkeitstests eignen sich zur Bewertung der Effektivität, aber vor allem zur Überprüfung, ob alle Anforderungen an die Effizienz des Produkts oder Softwaresystems erfüllt wurden. Gebrauchstauglichkeitstests, insbesondere im Abnahmetest, sollten unter Bedingungen durchgeführt werden, die den Bedingungen der antizipierten Produktivumgebung soweit es geht entsprechen.

Im Gegensatz zu den ersten beiden Verfahren sind Gebrauchstauglichkeitstests dynamische Tests.

Das beste Ergebnis bei der Evaluierung der Gebrauchstauglichkeit wird erzielt, wenn die drei Evaluierungsverfahren kombiniert werden und sowohl formativ als auch summativ durchgeführt werden.

Die Verfahren können zu unterschiedlichen Zeitpunkten im Entwicklungsprozess eingesetzt werden. Gebrauchstauglichkeitsreviews lassen sich beispielsweise schon früh im Projekt durchführen, um möglichst viele Befunde frühzeitig zu identifizieren, bevor es in den vergleichsweise aufwendigen dynamischen Gebrauchstauglichkeitstest geht. In diesem Sinne steht die Evaluierung der Gebrauchstauglichkeit der allgemeinen Empfehlung, möglichst frühzeitig mit dem Test eines Systems zu beginnen, in nichts nach.

Folgende Methoden werden im Abnahmetest häufig für die Evaluierung der Gebrauchstauglichkeit eingesetzt:

Checklistenbasierte Evaluierungen

Benutzer bewerten das zu testende System oder Produkt anhand von Checklisten [Rubin & Chisnell 08], um ihre Erfahrungen zu bewerten, zu vergleichen und zu qualifizieren. Zu diesen Checklisten zählen auch Heuristiken, beispielsweise die anerkannten Heuristiken von Jakob Nielsen [Nielsen 94]. Eine Kategorie der Nielsen-Heuristiken ist beispielsweise die Empfehlung, dass keine Fehlercodes als Fehlermeldung verwendet werden sollen, sondern natürlichsprachige Fehlermeldungen. Checklisten und Heuristiken können sowohl in Gebrauchstauglichkeitsreviews als auch für Gebrauchstauglichkeitstests verwendet werden.

Expertenreviews der Gebrauchstauglichkeit

Gebrauchstauglichkeitsexperten bewerten die Gebrauchstauglichkeit des Systems oder Produkts anhand vordefinierter Kriterien oder Checklisten auf der Grundlage von Gebrauchstauglichkeitsheuristiken, um die Stärken und Schwächen einer Mensch-Maschine-Schnittstelle zu identifizieren. Expertenreviews werden, wie der Name suggeriert, von Gebrauchstauglichkeitsexperten durchgeführt. Ein Expertenreview der Gebrauchstauglichkeit ist eine konkrete Reviewart der Gebrauchstauglichkeitsreviews.

Walkthrough und Think-Aloud-Techniken

Benutzer erkunden das Produkt oder das Softwaresystem und teilen dabei ihre Handlungen, Eindrücke und Erwartungen laut mit [ISTQB CTFL-UT], während sie

eine Gebrauchstauglichkeits-Testaufgabe durchführen. Diese Methode wird insbesondere bei Gebrauchstauglichkeitstests eingesetzt. Der Vorteil dieses Verfahrens ist, dass auf diese Weise Verständnis bei den Gebrauchstauglichkeitsexperten geschaffen wird, wie Benutzer mit dem Produkt interagieren, und mehr über Erwartungen oder Schwierigkeiten erfahren wird.

Biometrische Evaluierungen

Das Benutzerverhalten wird mit speziellen biometrischen Geräten (z. B. zur Aufzeichnung von Augenbewegungen, Aufzeichnung von Mausbewegungen) überwacht, um zu verstehen, wie der Benutzer mit einer Webseite oder einer anderen Mensch-Maschine-Schnittstelle interagiert, was seine Aufmerksamkeit erregt oder was für Benutzer mehr oder weniger sichtbar ist.

Protokolldateianalyse

Eine nachträgliche Analyse wird durchgeführt, um die Interaktion der Benutzer mit dem System zu überprüfen, um Bereiche für mögliche Verbesserungen zu ermitteln und um zu verifizieren, ob die tatsächliche Nutzung mit dem beabsichtigten Profil bzw. der Nutzung korreliert.

Das Thema Gebrauchstauglichkeit und Benutzererlebnis wird von einer ganzen Reihe an Gremien und Lehrplänen behandelt. Neben dem ISTQB®-Lehrplan Usability Testing befasst sich auch das UXQB (International Usability and User Experience Qualification Board [URL: UXQB]) mit einer Ausbildung im Bereich der Gebrauchstauglichkeit und des Benutzererlebnisses.

4.3 Performanz

Die Performanz eines Produkts oder Softwaresystems beschreibt die Reaktionsfähigkeit und Stabilität eines Systems unter bestimmten Nutzungsbedingungen. Sie ist ein wesentlicher Aspekt, um Benutzern ein »gutes Erlebnis« zu bieten, wenn sie für ihre Aufgaben eine Vielzahl mobiler und nicht mobiler Plattformen nutzen [ISTQB CTFL-PT]. Das »gute Erlebnis« deutet bereits darauf hin, dass Performanz eng mit der Gebrauchstauglichkeit zusammenhängt, denn Benutzer können ihre Aufgaben im jeweiligen Nutzungskontext nur dann effizient und zufriedenstellend erledigen, wenn das Zeitverhalten und die Verarbeitungsgeschwindigkeit des Produkts angemessen sind. Dabei ist die Erwartungshaltung der Benutzer in vielen Nutzungskontexten bereits so hoch, dass schlechte Performanz bei Anwendungen kaum mehr toleriert wird.

Performanz ist ein nicht funktionales Qualitätsmerkmal der ISO 25010 mit den folgenden Teilmerkmalen (siehe Abb. 4–3):

Zeitverhalten

Der Grad, zu dem eine Komponente oder ein System ihre bzw. seine erforderlichen Funktionen innerhalb der erforderlichen Antwortzeiten, Verarbeitungszeiten und Durchsatzraten erfüllen kann. Das Teilmerkmal Zeitverhalten ist unweigerlich mit dem Begriff *Performanz* verwoben. Für Benutzer ist das Zeitverhalten ein maßgebliches Merkmal während der Interaktion mit einem System. Das Zeitverhalten umfasst gemeinhin die Zeit, die es benötigt, bis eine Reaktion des Systems auf ein auslösendes Ereignis – eine Benutzeranfrage oder eine Zustandsänderung der Umgebung – erfolgt. Gutes Zeitverhalten wird üblicherweise honoriert und wertgeschätzt. Schlechtes Zeitverhalten führt rasch zu Ablehnung eines Systems. In den Zeiten von schnellem Internet immer und überall wird schlechtes Zeitverhalten durch die Benutzer nur wenig toleriert.

Ressourcennutzung

Der Grad, bis zu dem die Ressourcen gemäß den in den Anforderungen definierten Mengen und Arten genutzt werden können, wenn eine Komponente oder ein System ihre bzw. seine Funktionen ausführt. Ressourcennutzung ist ein Teilmerkmal, das von den Benutzern eines Systems immer nur dann bemerkt wird, wenn die Ressourcennutzung nicht optimal ist. Lange Wartezeiten, verlangsamte Verarbeitungsgeschwindigkeit oder abstürzende Server infolge von ungeplant hoher Anzahl an Anfragen an ein System sind die typischen Auswirkungen von unzureichender Ressourcennutzung.

Kapazität

Der Grad, zu dem die Höchstgrenzen eines Komponenten- oder Systemparameters Anforderungen erfüllen. Die Kapazität eines Systems hat vor allem Bedeutung für das operative Personal eines Systems. Die Bestimmung der Kapazität ist wichtig, um herauszufinden, wo die ressourcenbedingten Grenzen des Systems liegen. Zum einen dient diese Information zur Erstellung von Lastprofilen, die beim Stresstest eingesetzt werden, zum anderen können konstruktive Maßnahmen installiert werden, die das operative Team alarmieren, wenn die Ressourcen kurz davor sind, die Kapazitätsgrenzen im Produktivbetrieb zu überschreiten, oder sie bereits überschritten haben. Auf diese Weise soll schlechter Ressourcennutzung eines Systems vorgebeugt werden.

Zeitverhalten als Ideengeber neuer Dienste

Herausragendes Zeitverhalten ist die Grundlage für immer neue und spektakuläre Dienste. Autonomes oder teilautonomes Fahren ist nur deshalb möglich, weil die Systeme in der Lage sind, kontinuierliche Ströme an

Informationsdaten umgehend auszuwerten und auf den Informationsgehalt *schnell* zu reagieren. Dies wird vor allem durch Machine Learning, künstliche Intelligenz und Mehrprozessorrechner ermöglicht. Durch immer schnellere Kommunikationsprotokolle, Durchsatzraten und Antwortzeiten (Stichwort: 5G und in ferner Zukunft 6G) dringt die Digitalisierung sukzessive in viele industrielle Bereiche und Bereiche des öffentlichen Lebens vor. Das Auslesen und Integrieren von Maschinendaten zum Zwecke optimierter Wartungs- oder Fertigungsprozesse ist nur ein Beispiel. Durch immer größere Datenmengen, die prinzipiell verarbeitet werden müssen, sind neben dem mittlerweile gut verstandenen Cloud Computing Forschungsbereiche wie das sogenannte *Edge Computing* [URL: Edge Computing] entstanden, um auch bei großen zu durchsuchenden Datenmengen ein schnelles Zeitverhalten bei vernetzten Systemen zu gewährleisten.

Der Begriff *Performanztest* steht repräsentativ für eine Menge von verschiedenen Testarten, wobei sich jede Testart auf eine der drei Teilmerkmale konzentriert

Die verschiedenen Arten des Performanztests

[ISTQB CTFL-PT]. Tests, die sich auf die Bewertung des Zeitverhaltens eines Produkts oder Softwaresystems konzentrieren, sind vermutlich die am häufigsten durchgeführte Art der Performanztests. Das Zeitverhalten ist insbesondere in Echtzeitsystemen, also Systemen mit fest definierten Antwortzeiten, von enormer Bedeutung. Allerdings müssen auch bei Nicht-Echtzeitsystemen vertraglich festgelegte Antwortzeiten, die häufig in Form von Service Level Agreements (SLAs) vorliegen, eingehalten werden. Diese werden üblicherweise in Form von Performanzanforderungen und -abnahmekriterien spezifiziert. Ziel eines Performanztests ist es, sicherzustellen, dass das Zeitverhalten konform zu den entsprechenden Anforderungen des Produkts oder Softwaresystems ist. Im Folgenden sind einige der üblichen Performanztestarten knapp umrissen.

Der Lasttest ist eine Ausprägung des Performanztests, der üblicherweise gemeint wird,

Lasttest

wenn ganz allgemein von Performanztests gesprochen wird. Der Lasttest zielt darauf ab, das Zeitverhalten des zu testenden Systems in den Grenzen der zu erwartenden Last zu prüfen. Die zu erwartende Last wird durch Lastprofile verschiedener realistischer Benutzer erzeugt. Ein Lasttest wird daher häufig auch implizit durchgeführt, wenn beim Benutzerabnahmetest die abzunehmenden Geschäftsprozesse oder Anwendungsfälle, also Ende-zu-Ende-Szenarien, validiert werden.

Lastspitzentest

Der Lastspitzentest ist eine besondere Form des Lasttests, der darauf abzielt, den Umgang des zu testenden Systems mit plötzlich auftretender, aber noch spezifikationskonformer Spitzenlast zu prüfen. Da schlagartig eine Vielzahl von

Ressourcen allokiert werden müssen, kann es sein, dass es zu Problemen kommt. Insbesondere die Fähigkeit des Systems, nach der bewältigten Spitzenlast wieder in einen stabilen Zustand zurückzukehren, ist hierbei von Interesse.

Der Stresstest überprüft das Zeitverhalten des zu testenden Systems, wenn die tatsächlich anfallende Last außerhalb der spezifizierten, erlaubten Grenzen liegt. Stresstests werden insbesondere auch mit dynamischen Analysen kombiniert, um den Umgang des zu testenden Systems mit knapper werdenden Systemressourcen zu überwachen. Stresstests werden üblicherweise vor oder während des betrieblichen Abnahmetests durchgeführt.

Stresstest

Das Ziel des Dauertests ist es, sicherzustellen, dass auch bei einer langfristigen Nutzung des zu testenden Systems in üblichen Nutzungskontexten (also konform zu der Spezifikation des Systems) keine Probleme mit den Ressourcenkapazitäten auftreten. Dauertests werden gemeinhin nicht im Benutzerabnahmetest durchgeführt, sind aber als Teil der vereinbarten Service Level Agreements und zur Bestimmung der Verfügbarkeit von großer Bedeutung. Diese Art des Performanztests wird überwiegend vom operativen Personal durchgeführt. In seiner Natur ist der Dauertest sehr eng mit Zuverlässigkeitstests verwandt. Eine Ausprägung für einen langlaufenden Zuverlässigkeitstest ist beispielsweise der Dauertest.

Dauertest

Der Skalierbarkeitstest ist in seiner Natur ein informativer Test, da er auslotet, bis zu welchen Leistungsgrenzen das zu testende System skalieren kann, bevor es zu ersten Fehlverhalten oder problematischen Situationen kommt. Der Skalierbarkeitstest hat dadurch auch Ähnlichkeiten mit dem Stresstest, da bewusst die spezifizierten Lastgrenzen ausgereizt und überschritten werden. Durch Einsatz dynamischer Analysewerkzeuge lässt sich beispielsweise auch herausfinden, welche Ressource als Erstes einen Flaschenhals darstellt. Zudem lassen sich auf diese Weise auch technische Schwellenwerte definieren, bei deren Annäherung oder Überschreitung Warnungen an das operative Personal gemeldet werden. Der Skalierbarkeitstest ist üblicherweise kein Test, der während des Benutzerabnahmetests durchgeführt wird, sondern im betrieblichen Abnahmetest oder auch als kontinuierlicher Wartungstest, während das System produktiv ist.

Skalierbarkeitstest

Der Kapazitätstest ähnelt in der Art und Weise dem Skalierbarkeitstest, allerdings hat der Kapazitätstest eine kritisierende Funktion. Beim Kapazitätstest wird eruiert, wie viel Last das zu testende System bewältigen kann unter Berücksichtigung der vereinbarten Leistungsgrenzen. Es wird also die Menge der zu erzeugenden Last bewertet, die notwendig ist, um das System an seine Leistungsgrenzen zu bringen, und nicht das Verhalten des zu testenden Systems, wenn die Leistungsgrenze erreicht oder

Kapazitätstest

überschritten wird. Der Kapazitätstest lässt sich auch hervorragend als Vorbereitung für einen Stresstest durchführen, da auf diese Weise schnell ersichtlich ist, welche Last der Stresstest erzeugen muss, um die vereinbarten Leistungsgrenzen zu überschreiten. Der Kapazitätstest wird in der Regel auch nicht im Benutzerabnahmetest durchgeführt, sondern eher im betrieblichen Abnahmetest.

Gerade die letzten drei genannten Performanztestarten – Dauertest, Skalierbarkeits- und Kapazitätstest – gehen immer auch einher mit einer kritischen Bewertung oder zumindest informativen Überwachung der Ressourcennutzung eines Systems. Die Ressourcennutzung wird vor allem durch dynamische Analysewerkzeuge überprüft. Sogenannte Monitoring-Werkzeuge liefern ein Abbild der technischen Ressourcennutzung der zugrunde liegenden Hardware, um mögliche Fehlerzustände zu identifizieren, die sich mit funktionalen Tests oder statischer Analyse nur schwer nachvollziehen ließen. Eine treffende Kategorie von Fehlerursachen, die durch dynamische Analyse der Ressourcennutzung aufgedeckt werden können, sind Speicherlecks, die kontinuierlich zu Fehlerwirkungen des Produkts führen können. Häufig werden dynamische Analysen bei langlaufenden Zuverlässigkeitstests eingesetzt, um sicherzustellen, dass derartige Fehlerursachen ausgeschlossen werden können. Allerdings spielt die Ressourcennutzung auch bei Stress- und Lasttests eine große Rolle, da sichergestellt werden soll, dass das Produkt oder Softwaresystem bei fehlenden oder zumindest einem Mangel an Ressourcen nicht abstürzt, sondern kontinuierlich in seinem Zeitverhalten zurückgeht, bis sich die Ressourcenverfügbarkeit verbessert hat. Die dynamische Analyse eines Stresstests kann Aufschluss darüber geben, wann und warum ein System abgestürzt ist.

*Performanztests und
dynamische Analysen*

Generell sollten Performanztests auf allen Teststufen berücksichtigt werden. Je nachdem, was gemessen werden muss, können verschiedene Testarten durchgeführt werden. Dazu gehören Lasttests, Stresstests und Dauer-/Stabilitätstests. Performanztests werden auf den verschiedenen Teststufen gewöhnlich zu verschiedenen Zwecken eingesetzt:

Teststufen

Komponententest

Im Komponententest geht es überwiegend darum, die Verarbeitungsgeschwindigkeit einzelner Komponenten zu bewerten und Flaschenhälse in der Ressourcenverwendung innerhalb dieser Komponenten zu identifizieren. Wenn die Verarbeitungsgeschwindigkeit einer einzelnen Komponente für sich genommen nicht optimal erscheint, lohnt es sich ggf., die Performanz der Komponente zu optimieren. Kommen zu einem späteren Zeitpunkt, beispielsweise im Komponentenintegrationstest, weitere

Komponenten in den Verarbeitungsprozess hinzu, kann sich durch die Aufsummierung der einzelnen Verarbeitungszeiten durchaus ein relevantes Performanzproblem ergeben.

Komponentenintegrationstest

Je nach Integrationsstrategie werden die Komponenten integriert und so zu immer größeren Teilsystemen zusammengesetzt. Der Performanztest auf dieser Stufe konzentriert sich darauf, zu überwachen, ob durch die Integration einzelner Komponenten Performanzprobleme vorhanden sind oder wahrscheinlich vorhanden sein werden. Die Aussagekraft eines Performanztests auf Integrationsteststufe ist mit Vorsicht zu genießen, da noch nicht alle Komponenten integriert sind. Dies ist erst beim Systemtest der Fall. Es gibt daher keine Gewähr, dass ein auf den ersten Blick sehr gutes Zeitverhalten eines teilintegrierten Systems durch Hinzunahme einer weiteren Komponente gleich gut bleibt. Die Durchführung von Performanztests auf Komponententestebene verringert das Risiko eines plötzlich schlechten Ergebnisses jedoch ganz erheblich.

Systemtest

Beim System erfolgt die Bewertung verschiedener Performanzaspekte durch Verwendung realistischer bzw. zu erwartender Ende-zu-Ende-Szenarien. Anwendungsfälle bieten eine gute Grundlage, um die Performanztests zu entwerfen und durchzuführen. Das primäre Ziel des Performanztests auf Systemteststufe ist es, Fehler bezüglich der Performanzanforderungen in den Ende-zu-Ende-Szenarien vor der anstehenden Abnahme des Produkts durch den Auftraggeber zu identifizieren.

Systemintegrationstest

Der Systemintegrationstest testet das System im Zusammenspiel mit Drittsystemen. Auch hier bilden Anwendungsfälle und vor allem Anwendungsfalldiagramme eine solide Grundlage, da sie den Systemkontext und die Schnittstellen der relevanten Umgebungssysteme mit dem Produkt oder Softwaresystem spezifizieren.

Abnahmetest

Bei den verschiedenen Varianten des Abnahmetests geht es primär stets darum, das Vertrauen der Benutzer, Kunden und des operativen/betreibenden Personals in die ordnungsgemäße Performanz des Systems aufzubauen und um eine Feinabstimmung des Systems im Betrieb unter realen Bedingungen

durchzuführen. Wie bereits erwähnt ist es nicht mehr das primäre Ziel, Performanzfehler aufzudecken.

Die meisten Ausprägungen der Performanztests werden während des betrieblichen Abnahmetests bzw. von den operativen Teams durchgeführt. An der Entwicklung von Abnahmekriterien und der zugehörigen Testfälle sollten jedoch auch Businessanalysten und Tester mitwirken. Abnahmekriterien für Performanzanforderungen sollten genauso wie für funktionale Anforderungen objektive Maße bereitstellen, anhand derer man eindeutig bewerten kann, ob das Performanzabnahmekriterium erfüllt wurde oder nicht. Damit wird eine subjektive Performanzbewertung während der Ausführung der Abnahmetests vermieden.

Neben den betrieblichen Performanzabnahmetests, zu denen insbesondere Stresstests, Dauertests und Kapazitätstest gehören, werden im Benutzerabnahmetest insbesondere Lasttests (oftmals implizit) im Zuge der Abnahme durchgeführt. Das Ziel besteht dabei aber nicht primär darin, Performanzfehler zu finden, sondern Vertrauen der Benutzer in die ordnungsgemäße Performanz des abzunehmenden Produkts zu generieren.

4.3.1 Anforderungen und Abnahmekriterien für Performanzabnahmetests

Performanztests sollen in erster Linie die Reaktionsfähigkeit (zeitliches Antwortverhalten) und den Durchsatz eines Produkts oder Softwaresystems unter bestimmter Arbeitslast ermitteln. Performanztests werden auch von anderen Testarten (bspw. Zuverlässigkeitstests und IT-Sicherheitstests) eingesetzt. Dauertests helfen bei der Bewertung der Zuverlässigkeit eines Systems, Stresstests werden für Denial-of-Service-Angriffe (DoS) eingesetzt, die häufig bei IT-Sicherheitstests durchgeführt werden, um die Verfügbarkeit und somit die Integrität eines Systems oder einzelner Komponenten zu beschädigen.

Dabei ist für die Benutzer primär die Antwortzeit des Produkts relevant, der Durchsatz (also die Anzahl der Transaktionen eines bestimmten Typs, die innerhalb einer bestimmten Zeiteinheit verarbeitet werden können) ist ein Geschäftsanliegen der Businessanalysten und die Ressourcennutzung ein technisches Anliegen der operativen/betreibenden Teams. Diese drei Perspektiven (Benutzerperspektive, Geschäftsperspektive, Betriebsperspektive) sollten bei der Entwicklung von Abnahmekriterien und Abnahmetests für Performanzanforderungen berücksichtigt werden:

Aus Sicht des Benutzers spiegelt die wahrgenommene Antwortzeit die tatsächliche Erfahrung des Benutzers mit dem System wider. Im Benutzerabnahmetest werden inakzeptable Antwortzeiten eher implizit durch die Abnahme der zugrunde liegenden Geschäftsprozesse und Anwendungsfälle

gefunden. Das Antwortzeitverhalten des abzunehmenden Produkts wirkt sich zudem bedeutend auf das Benutzererlebnis aus. Beispielsweise kann es passieren, dass Benutzer eine Website verlassen, wenn die Antwortzeit mehr als zehn Sekunden beträgt⁷.

Aus geschäftlicher Sicht sind die Anzahl der gleichzeitigen Benutzer, die Art der durchgeführten Szenarien oder Transaktionen und die erwarteten Antwortzeiten Faktoren, die berücksichtigt werden müssen. Eine höhere Anzahl gleichzeitiger Benutzer, die ressourcenintensive Transaktionen ausführen, führt zu längeren Antwortzeiten. Auch andere Faktoren können die Antwortzeit je nach Ort, Uhrzeit oder Zeitzone beeinflussen.

Aus technischer Sicht sind die verfügbaren Systemressourcen (z. B. Netzwerkbandbreite, Prozessorauslastung, Speicherkapazität) und Systemarchitektur (z. B. Lastverteilung zwischen Servern, Verwendung von Daten-Caching) Faktoren, die die Performanz beeinflussen. Beispielsweise weisen webbasierte Systeme mit begrenzter Netzwerkbandbreite tendenziell eine geringere Performanz auf, insbesondere, wenn sie hoher Last ausgesetzt sind, die durch eine große Anzahl von Benutzern verursacht wird, die Aufgaben ausführen, die erheblichen Netzwerkverkehr erzeugen.

Fallbeispiel

CA-Cockpit Performanzanforderungen und -abnahmekriterien

Für CA-Cockpit sind unter anderem die folgenden Performanzanforderungen und -abnahmekriterien spezifiziert:

Geschäftsprozess *GP-02b Handyvertrag abschließen (mit Laufzeit)*

Anforderung	
ID	A-57
Titel	Durchschnittliches Antwortzeitverhalten Buchungsaktivität
Beschreibung	Das System muss dem Kunden während einer Buchungsaktivität innerhalb von drei Sekunden antworten.
Art	Performanz – Zeitverhalten

Abnahmekriterien	A-57-01	Von 1000 gleichzeitig stattfindenden Buchungsaktivitäten darf die Antwortzeit von 80 % aller Anfragen nicht länger als eine Sekunde betragen.
	A-57-02	Bei 3000 gleichzeitig stattfindenden Buchungsaktivitäten darf die Antwortzeit von 90 % aller Anfragen nicht länger als zwei Sekunden betragen.
	A-57-03	Bei 5000 gleichzeitig stattfindenden Buchungsaktivitäten darf die Antwortzeit von 95 % aller Anfragen nicht länger als drei Sekunden betragen.

Geschäftsprozess *GP-03 Handyvertrag verwalten*

Anforderung		
ID	A-62	
Titel	Durchschnittliches Antwortzeitverhalten Kundenbereich	
Beschreibung	Das System muss dem Kunden bei Anfragen im Kundenbereich innerhalb von sieben Sekunden antworten.	
Art	Performanz – Zeitverhalten	
Abnahmekriterien	A-62-01	Bei 10.000 gleichzeitig stattfindenden Anfragen aus dem Kundenbereich darf die Antwortzeit von 75% aller Anfragen nicht länger als sieben Sekunden betragen.
	A-62-02	Bei mehr als 10.000 gleichzeitiger Anfragen werden Anfragen per Prioritätenliste von dem System verarbeitet. Für alle Anfragen der Priorität 1 und 2 gelten weiterhin die Antwortzeiten von A-62-01. Alle anderen Anfragen werden bis zur Abarbeitung der hoch priorisierten Anfragen zurückgestellt.

Das Abnahmekriterium A-62-02 geht davon aus, dass eine existierende Spezifikation bezüglich der Prioritäten von Anfragen im Kundenbereich existiert.

4.3.2 High-Level-Performanzabnahmetests

Ähnlich wie andere Testarten, so verifizieren auch Performanzabnahmetests, dass die Performanz des Systems die ursprünglich spezifizierten Benutzerbedürfnisse und Abnahmekriterien erfüllt. Darüber hinaus kann die Bewertung der funktionalen Eignung, der Gebrauchstauglichkeit und anderer Qualitätsmerkmale unter Lastbedingungen, z. B. während der Durchführung eines Performanztests, lastspezifische Probleme aufdecken, die sich auf diese Qualitätsmerkmale auswirken [ISTQB CTFL-PT].

In einem typischen Performanztest wird die für eine bestimmte Ausprägung eines Performanztests benötigte Last mit spezifischen Werkzeugen generiert. Während die Realitätsnähe der simulierten Last auf den unteren Teststufen noch nicht zwingend gegeben sein muss, ist es zu empfehlen, dass die erzeugte Last in den verschiedenen Abnahmetests die tatsächlichen Bedingungen mit echten Benutzern und realistischen Interaktionen so genau wie möglich nachahmt. Diese realistischen Interaktionen werden durch sogenannte operationale Profile beschrieben.

Ein operationales Profil eines Benutzers (auch Nutzungsprofil genannt) spezifiziert das realistische bzw. zu erwartende oder analysierte (bspw. aus Logdateien) Verhalten des Benutzers im Umgang mit dem System. Dazu zählen zum einen der Zugriff auf verschiedene Funktionen des Produkts, zum anderen die Daten, die von bestimmten Benutzergruppen oder Personas mit dem System ausgetauscht werden. Es kann auch mehrere verschiedene Nutzungsprofile für die gleiche Benutzergruppe geben, die das Verhalten dieser Gruppe zu unterschiedlichen Tageszeiten oder zu besonderen Anlässen widerspiegelt. So könnte bei einem Nachrichtenportal bei Tagesbeginn eine hohe Nutzungslast anfallen, während diese bis zum Abend geringer ist und abends wieder ansteigt. Nutzungsprofile lassen sich auch adäquat aus Anwendungsfällen gewinnen, da diese ja die zu erwartenden Interaktionen der Akteure mit dem Produkt spezifizieren. Handelt es sich bei den Akteuren um menschliche Benutzer, dann sollten auch Bedenkzeiten dieser Benutzer in dem Nutzungsprofil reflektiert sein. Bedenkzeiten sind immer dann sinnvoll, wenn der Benutzer während der Interaktionen Entscheidungen treffen muss. Wird einem Kunden beispielsweise die Option geboten, für einen Onlinekauf eine kostenpflichtige Versicherung abzuschließen, dann wird dies üblicherweise zu einer Verzögerung bedingt durch die Entscheidungsfindung des Benutzers führen.

Nutzungsprofile, Lastprofile und Last

Aus Nutzungsprofilen werden dann mit einem entsprechenden Werkzeug sogenannte Lastprofile erzeugt, also technische Abbilder der zuvor spezifizierten Nutzungsprofile zum Zwecke der Lasterzeugung. Lastprofile können dabei mehrere Nutzungsprofile

Ableiten von Lastprofilen

unterschiedlicher Benutzergruppen oder Personas kombinieren, um ein möglichst realistisches Abbild der zu erwartenden oder üblicherweise anfallenden Last zu simulieren (da es zu erwarten ist, dass nicht immer nur eine Benutzergruppe mit dem Produkt zu jedem Zeitpunkt interagiert).

Lastgeneratoren erzeugen die spezifizierten Lastprofile und führen diese dann in Form von Testfällen aus. Performanz ist ein nicht funktionales Qualitätsmerkmal, das auf Funktionalität aufsetzt. Daher sind funktionale Testfälle häufig eine gute Grundlage für die Durchführung von Performanztests.

Erzeugung effektiver Last

Um die Performanz eines Produkts oder Softwaresystems gegen die objektiv spezifizierten Abnahmekriterien messen zu können, ist es unerlässlich, das Antwortzeitverhalten oder den Ressourcenverbrauch zu überwachen und aufzuzeichnen. Zu diesem Zweck werden dynamische Analysewerkzeuge eingesetzt, im Falle von Performanztests sogenannte Test- oder Performanzmonitore. Diese Werkzeuge überwachen bestimmte Schlüsselemente wie Durchsatz oder Antwortzeitverhalten des Produkts oder Softwaresystems während eines Performanztests. Die Antwortzeiten dieser Schlüsselemente des zu testenden Systems oder Komponenten (z. B. Webserver, Applikationsserver, Datenbanken) werden nach der Messung mit den vordefinierten Performanzabnahmekriterien verglichen. Es können auch andere Schlüsselemente überwacht, aufgezeichnet und verglichen werden, beispielsweise die Speichernutzung, Systemeingaben und -ausgaben, die Prozessorauslastung und der Zugriff auf Sicherheitsgeräte, jeweils abhängig davon, welches Schlüsselement der (erwartete) Flaschenhals ist bzw. getestet werden soll.

Monitoring

Basierend auf der Analyse der Ergebnisse können bestimmte Elemente in der Architektur (Hardware und Software) geändert werden (z. B. die Bereitstellung zusätzlicher Serverkapazität). Der Zyklus des Testens, Analysierens und Verbesserns sollte wiederholt werden, bis die angestrebte Performanz erreicht ist [ISTQB CTFL-AcT].

Der ISTQB[®]-Lehrplan Performance Testing [ISTQB CTFL-PT] behandelt vertieft und umfangreich das Thema Performanztests.

4.4 IT-Sicherheit

In einer zunehmend digitalisierten und vernetzten Welt, in der sich internetfähige Geräte scheinbar nach Belieben miteinander verbinden können, in der autonom fahrende Autos pausenlos ihre Umgebung analysieren und mit entsprechenden Entitäten kommunizieren, in der quasi immer und überall Zugang zu Netzwerken,

Cloud-Speichern und Informationen besteht, ist die Absicherung von Informationssystemen gegen unerlaubte, oftmals böswillige Zugriffe zu einem zentralen Qualitätsmerkmal geworden. Sicherheit muss für den Betrieb eines alltagsgebräuchlichen IT-Systems gewährleistet werden. Im Kontext von IT-Systemen gibt es zwei Ausprägungen von *Sicherheit*, die unterschiedliche, aber komplementäre Eigenschaften beschreiben:

Funktionale Sicherheit (safety)

Sicherheit vs. Sicherheit

Von funktionaler Sicherheit spricht man, wenn elektronische oder elektrische Schutzsysteme implementiert werden, damit ein softwareintensives System seine Funktionalität zuverlässig erbringen kann. Durch die Schutzsysteme wird die relevante Umgebung des softwareintensiven Systems vor möglichen durch fehlerhaften Betrieb ausgelösten Schäden geschützt. Die relevante Umgebung kann dabei sehr unterschiedliche Ausmaße annehmen. So kann etwa nur der unmittelbare Benutzer des Systems betroffen sein, beispielsweise ein Industriearbeiter, der an einer Automatisierungszelle arbeitet; eine kleine Gruppe von Personen, beispielsweise Passagiere in einem Flugzeug oder Zug; aber auch die gesamte (teil)globale Umwelt, beispielsweise durch eine Explosion eines Kernkraftwerks. Die Tragweite möglicher durch ein fehlerhaftes softwareintensives System verursachter Schäden skaliert je nach Einsatzgebiet des Systems.

IT-Sicherheit (security)

Von IT-Sicherheit spricht man, wenn ein softwareintensives System vor zufälligen oder bewusst boshaften Handlungen der Akteure der relevanten Umgebung des Systems geschützt werden soll. Ziel ist es, die Informationen, Dateien, Architekturen, kurz Assets, die für ein Unternehmen oder eine Privatperson einen Wert besitzen, vor der relevanten Umwelt zu schützen. Je wertvoller die zu schützenden Assets eines Unternehmens oder einer privaten Person sind, desto umfassender sollten auch die Schutzmechanismen für die IT-Sicherheit sein.

Die Begriffe beschreiben also unterschiedliche Wirkungsrichtungen fehlerhafter bzw. bössartiger Auslöser (vgl. Abb. 4–10). Dies führt zu der Erkenntnis, dass funktionale Sicherheit nur durch IT-Sicherheit zu gewährleisten ist. Ein gutes Beispiel, das diesen Zusammenhang beschreibt, ist die erfolgreiche Ausnutzung einer IT-Sicherheitslücke eines Jeep Cherokee, durch die die Hacker Zugriff auf die Kontrollsysteme des fahrenden Wagens und somit auf kritische Funktionen wie Getriebesteuerung, Bremssysteme etc. erhielten [URL: Jeep]. Durch diese IT-Sicherheitslücke und einen ausnutzenden IT-Sicherheitsangriff hätte der relevanten Umgebung – in diesem Fall den Insassen des Wagens sowie den zu der

Zeit des Angriffs erreichbaren Entitäten (andere Autos, Zweiradfahrer, Fußgänger) – ein kritischer Schaden zugefügt werden können. Eine verbeulte Motorhaube oder Blechschäden wären sicherlich noch die geringsten Schadensausmaße.

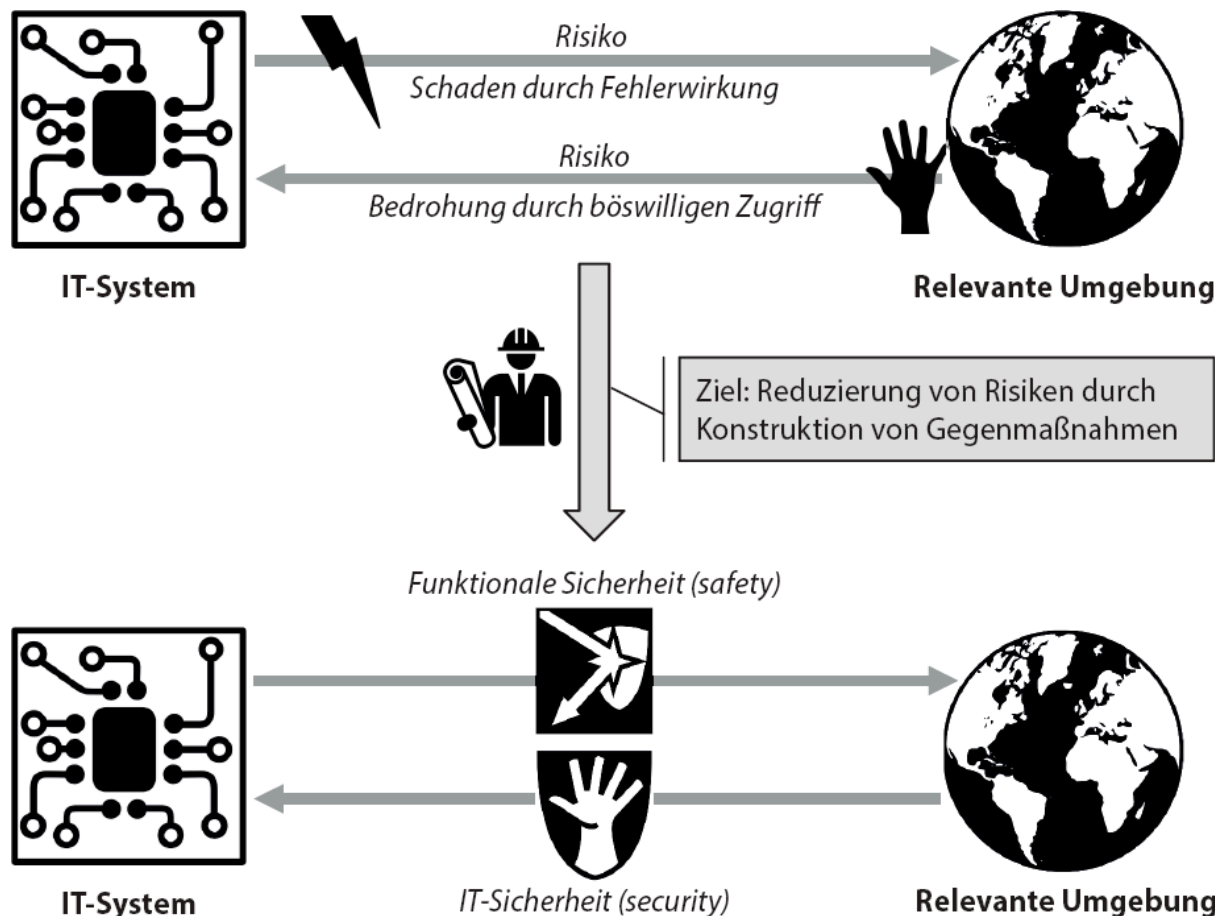


Abb. 4-10
Wechselwirkung von funktionaler und IT-Sicherheit

Diese Beziehung von IT-Sicherheit und funktionaler Sicherheit lässt sich zwar andersherum nicht auf ähnliche Weise argumentieren, dennoch gibt es hier zumindest Überschneidungen, was die Vermeidung der entsprechenden Sicherheitsrisiken betrifft. Funktionale Sicherheit wird insbesondere durch sogenannte funktionale Gegenmaßnahmen sichergestellt. Das heißt, eine besondere Funktion (das sogenannte Schutzsystem) des softwareintensiven Systems ist dafür zuständig, einen möglichen Schaden durch Ausführung der Funktionalität des softwareintensiven Systems vorzubeugen. So sind automatische medizinische Infusionspumpen beispielsweise mit Sensoren ausgerüstet, die einen Alarm anschlagen, wenn Luft im Schlauch oder der Kanüle festgestellt wird. Für funktionale Gegenmaßnahmen werden gewöhnlich Anforderungen formuliert und umgesetzt. Diese werden vereinzelt auch Safety-Anforderungen genannt, sind aber im Prinzip nichts anderes als funktionale Anforderungen. Als solche lassen sie sich hinsichtlich Umsetzung natürlich

verifizieren und validieren. Gegenmaßnahmen gibt es auch bei der IT-Sicherheit, beispielsweise Mehrfaktor-Authentifizierung [URL: Multi-Faktor-Authentisierung], Verschlüsselung [ISTQB Glossar], Vier-Augen-Prinzip [URL: Vier-Augen-Prinzip] oder die Vermeidung von Account Harvesting (Benutzerkonto-Einfangen [ISTQB Glossar]).

IT-Sicherheit gehört zu den nicht funktionalen Qualitätsmerkmalen der ISO 25010 mit den folgenden Teilmerkmalen:

Teilmerkmale von IT-Sicherheit

Vertraulichkeit

Der Grad, zu dem eine Komponente oder ein System sicherstellt, dass Daten nur für diejenigen zugänglich sind, die über eine Zugangsberechtigung verfügen. Vertraulichkeit hängt stark mit dem Teilmerkmal Authentizität und der Autorisierung zusammen.

Integrität

Der Grad, zu dem eine Komponente oder ein System nur autorisierten Zugriff und Änderung einer Komponente, eines Systems oder von Daten zulässt. Dabei sollte jede Änderung durch eine entsprechend autorisierte Person nicht abstreitbar protokolliert werden.

Nichtabstreitbarkeit

Der Grad, zu dem Aktionen oder Ereignisse nachweislich stattgefunden haben, sodass die Aktionen oder Ereignisse später nicht abgestritten werden können. Nichtabstreitbarkeit verhindert ein unzulässiges Abstreiten durchgeführter Aktionen, beispielsweise durch die Verwendung elektronischer Signaturen, wie es beim Abschluss elektronischer Anträge oder Verträge häufig der Fall ist.

Zurechenbarkeit

Der Grad, zu dem Aktionen eines Akteurs eindeutig zu diesem verfolgt werden können. Zurechenbarkeit steht damit in engem Verhältnis zu Nichtabstreitbarkeit und Authentizität.

Authentizität

Der Grad, zu dem die Identität eines Subjekts oder einer Ressource als die behauptete nachgewiesen werden kann. Es wird gewöhnlich zwischen Authentizität, Authentisierung (Echtheit) und Authentifizierung unterschieden. Authentisierung beschreibt die Art und Weise, wie eine Identifikation nachweislich erbracht wird. Authentifizierung beschreibt die Überprüfung, ob

eine Authentisierung glaubhaft ist. Auch wenn Authentizität und Autorisierung häufig zusammen aufgeführt wird, so handelt es sich doch um unterschiedliche Aspekte. Erst wird ein Subjekt authentifiziert, dann autorisiert. Die Autorisierung vertraut also darauf, dass die Authentifizierung erfolgreich abgeschlossen wurde. Was insbesondere bei gestohlenen Kreditkarten, Zugangspasswörtern oder sonstigen Authentifizierungsmethoden problematisch ist. Allerdings muss an einer bestimmten, sinnvollen Stelle bei der Verwendung eines IT-Systems darauf vertraut werden, dass ein Subjekt auch dasjenige ist, das es vorgibt zu sein.

Das Management der Informationssicherheit sowie allgemeine IT-Sicherheitsanforderungen sollten Teil einer umfassenden IT-Sicherheitsrichtlinie eines Unternehmens sein (weitere Informationen sind im »Sicherheitstester«-Lehrplan [ISTQB CTAL-ST] und im ISO-Standard 27005 [ISO 27005] zu finden).

IT-Sicherheit bezieht sich dabei nicht nur auf rein technische Aspekte eines zu schützenden IT-Systems, sondern umfasst daneben noch geschäftliche und regulatorische IT-Sicherheitsrichtlinien, die etabliert, umgesetzt und gelebt werden müssen, um ein IT-System ausreichend vor Sicherheitsangriffen zu schützen. In Anlehnung an ISO 27034-1 [ISO 27034-11] illustriert Abbildung 4-11 ein Rahmenwerk für IT-Sicherheit (adaptiert von [Simon et al. 19]).

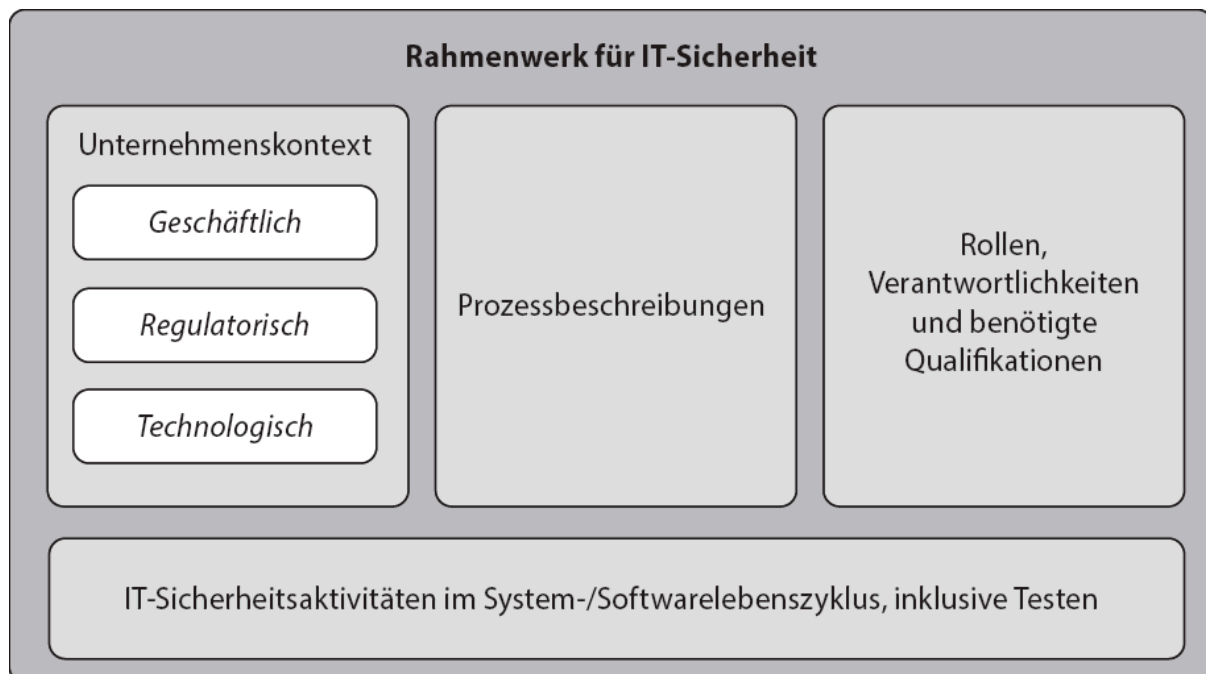


Abb. 4-11

IT-Sicherheitsrahmenwerk nach ISO 27034

Es wird deutlich, dass IT-Sicherheit – im Gegensatz zu allen anderen Qualitätsmerkmalen der ISO 25010 – ein Alleinstellungsmerkmal besitzt, nämlich, dass IT-Sicherheit nicht nur durch Überdeckung rein technischer Aspekte des

Produkts zu garantieren ist, sondern auch organisatorische Gegebenheiten (Prozesse, Personen) mittels entsprechender IT-Sicherheitsrichtlinien umfasst. Businessanalysten und Tester sollten die IT-Sicherheitsrichtlinien für Empfehlungen und Orientierungshilfen sowie als Grundlage für das Management von IT-Sicherheitsrisiken in ihren Projekten verwenden. IT-Sicherheitsrichtlinien wenden sich in erster Linie an die Personen, die mit IT-Systemen einer Organisation interagieren, und stellen Verhaltensregeln im Umgang mit der IT-Infrastruktur des Unternehmens auf. Durch IT-Sicherheitsrichtlinien werden Prozesse bestimmt, die die entsprechenden Richtlinien implementieren und die dann auch gelebt werden müssen.

Social Engineering

Ein spannender Teilbereich der IT-Sicherheit ist das sogenannte Social Engineering. Darunter versteht man die Umgehung von IT-Sicherheitsrichtlinien, indem die (allergrößte) Sicherheitsschwachstelle *Mensch* von anderen Menschen ausgenutzt wird, um an sensible Informationen zu gelangen. Das BSI definiert Social Engineering ([URL: BSI Glossar], Stichpunkt Social Engineering) wie folgt:

»Social Engineering ist eine Methode, um unberechtigten Zugang zu Informationen oder IT-Systemen durch soziale Handlungen zu erlangen. Beim Social Engineering werden menschliche Eigenschaften wie z. B. Hilfsbereitschaft, Vertrauen, Angst oder Respekt vor Autorität ausgenutzt. Dadurch können Mitarbeiter so manipuliert werden, dass sie unzulässig handeln.« Ein typisches Beispiel für Social Engineering sind Phishing-Mails oder ein (boshafter bzw. scheinbarer) Administrator, der wegen eines Systemfehlers bei einem Kunden anruft, da er zur Fehlerbehebung noch das Passwort des Benutzers benötigt.

Zusammengenommen deckt IT-Sicherheit ein gewaltiges Spektrum potenziell relevanter Aspekte für Organisationen und IT-Systeme ab. Dies macht IT-

IT-Sicherheit als komplexes Thema

Sicherheit zu einem komplexen und vielschichtigen Qualitätsmerkmal, für dessen Gütenachweis eine ganze Reihe von Methoden, Werkzeugen und vor allem Wissen benötigt wird. Zudem ist die Sicherstellung der IT-Sicherheit ein niemals endender Prozess. Ein Produkt, das heute produktiv geht und vergleichsweise guten Schutz vor bekannten Angriffsmustern bietet, kann morgen durch neue Angriffsmuster bereits veraltet sein und als unsicher gelten. Es ist unerlässlich, dass IT-Systeme sich ständig an das Gefahrenpotenzial anpassen und neu identifizierte Schwachstellen so schnell wie möglich schließen.

In allen Phasen der Businessanalyse, des Requirements Engineering und der zugehörigen Abnahmetests sollten daher IT-Sicherheitsanforderungen und deren Abnahmekriterien wie folgt berücksichtigt werden:

Die Informationssicherheit sollte Teil des Risikomanagements und der Erhebung und Analyse nicht funktionaler Anforderungen sein. Der Wert von schützenswerten Informationen (Assets) im getesteten System oder in einem bestimmten Geschäftsprozess sollte bewertet werden, gefolgt von einer Bewertung und Priorisierung der IT-Sicherheitsrisiken.

Für die identifizierten IT-Sicherheitsanforderungen sollten messbare Abnahmekriterien definiert werden. Sie adressieren gemeinhin eine große Bandbreite technischer Aspekte wie Authentifizierungs-, Autorisierungs- und Zurechenbarkeitsverfahren, Bereinigung von Eingabedaten, Verwendung von Kryptografie sowie Datenschutzbeschränkungen.

High-Level-IT-Sicherheitstests sollten basierend auf den IT-Sicherheitsanforderungen und -abnahmekriterien entworfen werden. Diese dynamischen IT-Sicherheitstests definieren den Kontext bzw. die Umgebung des Tests, die auszuführenden Schritte sowie die erwarteten Ergebnisse.

Einige IT-Sicherheitstests werden unter Zuhilfenahme anderer Testarten umgesetzt. So ist der Stresstest beispielsweise hilfreich dabei, bestimmte IT-Sicherheitsangriffe zu simulieren, etwa eine Denial-of-Service-Attacke (DoS), also das gezielte Überfrachten von Servern mit Anfragen mit dem Ziel, den Server ressourcentechnisch zu überlasten und zum Absturz zu bringen.

Fallbeispiel

IT-Sicherheitsanforderungen für den Geschäftsprozess *GP-03 Handyvertrag verwalten*

Aus den IT-Sicherheitsrichtlinien von *ComAccept AG* sowie den rechtlichen und regulatorischen Anforderungen an die Datensicherheit personenbezogener und personenbezogener Daten ergeben sich unter anderem die nachfolgenden IT-Sicherheitsanforderungen:

Anforderung	
ID	A-85
Titel	Sicherheits-PIN-Abfrage
Beschreibung	Das System muss den Kunden zwingen, eine separate Sicherheits-PIN einzugeben, um sich zu authentifizieren, wenn ungewöhnliche Verhaltensmuster des Kunden identifiziert wurden.

Art	Die Sicherheits-PIN wird an die mit dem Kundenaccount verbundene Mobilfunknummer per SMS übermittelt.
------------	---

Anforderung		
ID	A-86	
Titel	Inaktivitätsabmeldung	
Beschreibung	Das System muss den Kunden aus seinem Zugangsbereich abmelden, wenn der Kunde innerhalb des Inaktivitätszeitraums von fünf Minuten keine Anfrage an das System übermittelt hat.	
Art	IT-Sicherheit	
Abnahmekriterien	A-86-01	Die Zeit des Inaktivitätszeitraums soll bei jeder Anfrage des Kunden an das System wieder auf fünf Minuten zurückgesetzt werden.
	A-86-02	Die verbleibende Zeit des Inaktivitätszeitraums soll dem Kunden angezeigt werden.
	A-86-03	Nach Ablauf des Inaktivitätszeitraums soll der Kunde automatisch vom System abgemeldet werden und die Meldung INAKTIV_ABMELDUNG ⁸ angezeigt bekommen.
Abnahmekriterien (Fortsetzung)	A-86-04	Jeder Versuch, zu einer URL zu springen (per Rücksprungfunktion des Browsers oder per manueller Eingabe der URL), führt dazu, dass der Kunde auf die Login-Seite des Kundenportals umgeleitet wird.

Anforderung	
ID	A-87
Titel	Vermeidung von Account Harvesting

Beschreibung	Das System darf unter keinen Umständen den Grund eines fehlgeschlagenen Anmeldeversuchs dem Kunden mitteilen.	
Art	IT-Sicherheit	
Abnahmekriterien	A-87-01	Wird die Anmeldung mit einem unbekanntem Benutzernamen durchgeführt, muss das System dem Kunden die Meldung UNGÜLTIGE_ANMELDUNG anzeigen.
	A-87-02	Wird die Anmeldung mit einem ungültigen Passwort zu einem existierenden Benutzernamen durchgeführt, muss das System dem Kunden die Meldung UNGÜLTIGE_ANMELDUNG anzeigen.

Anforderung		
ID	A-88	
Titel	Authentifikation per Authentifizierungs-PIN telefonischer Kundenbetreuung	
Beschreibung	Bei telefonischer Betreuung eines Kunden durch einen Kundenmitarbeiter ist durch den Kundenbetreuer die Authentifizierungs-PIN des Kunden einzuholen.	
Art	IT-Sicherheit	
Abnahmekriterien	A-88-01	Der Kundenbetreuer muss den Kunden zu Beginn des Gesprächs nach seiner Authentifizierungs-PIN fragen.
	A-88-02	Der Kundenbetreuer darf keinerlei weitere Informationen über die Vertragsdetails des Kunden preisgeben, wenn die telefonische Authentifizierung mittels der Authentifizierungs-PIN nicht erfolgreich war.

Die letzte IT-Sicherheitsanforderung zeigt, dass nicht alle IT-Sicherheitstests zwingend technisch durchzuführen sind. Mittels klassischer Social Engineering-Methoden ließe sich gut überprüfen, ob die Kundenbetreuer sich an die IT-Sicherheitsrichtlinien der telefonischen Beratung halten.

4.4.1 Anforderungen und Abnahmekriterien für IT-Sicherheitsabnahmetests

Sicherheitsrelevante Abnahmekriterien beschreiben entsprechende Maßnahmen im Kontext eines IT-Systems wie Authentifizierung, Zugriffsschutz und Rechteverwaltung, die verwendeten kryptografischen Verfahren und Algorithmen inklusive der geforderten Schlüsselstärken und -längen oder auch das Vorhandensein einer durchgängigen und konsistenten Fehlerbehandlung sowie Logging- und Audit-Schnittstellen [Simon et al. 19].

Abnahmekriterien für IT-Sicherheitsanforderungen können sich dabei auf globale sicherheitsrelevante Maßnahmen wie den eben beschriebenen beziehen oder auf konkrete Szenarien bzw. konkrete Funktionen. So erfasst der DFB die Spielberichte aller Profi- und Amateurspiele seit einigen Jahren komplett digital. Der Schiedsrichter ist dafür zuständig, den Spielbericht zu verfassen und diesen mit seiner Zugangskennung zu bestätigen (Zurechenbarkeit, Nichtabstreitbarkeit). Nachfolgend müssen beide Trainer ebenfalls mit ihren Zugangskennungen den Spielbericht bestätigen. Erst wenn alle involvierten Parteien den Spielbericht bestätigt haben, gilt er für den DFB als abgenommen. Das Prinzip, das dieser Funktion zugrunde liegt, ist das zuvor schon erwähnte »Vier-Augen-Prinzip«, das die Bestätigung einer Aussage durch einen (oder in diesem Fall sogar zwei) andere autorisierte Personen verlangt.

4.4.2 High-Level-IT-Sicherheitsabnahmetests

Die zuvor definierten Abnahmekriterien dienen als Grundlage für den Entwurf von IT-Sicherheitstests. Insofern unterscheiden sich IT-Sicherheitstests nicht wesentlich von anderen Testarten. Es gibt vier allgemeine Ziele von IT-Sicherheitstests [ISTQB CTAL-ST]:

- Risiken (z. B. Schwachstellen in einem Produkt bzw. IT-System) zu identifizieren
- Existierende Gegen- bzw. Schutzmaßnahmen zu bewerten
- Einen Nachweis zu liefern, dass die IT-Sicherheitsrichtlinien des Unternehmens mit der gebotenen Sorgfalt umgesetzt worden sind und auch gelebt werden
- Vertrauen bei Kunden und Anwendern in die eigene Organisation zu schaffen

Bezogen auf den IT-Sicherheitsabnahmetest eines Produkts oder Softwaresystems sind vor allem das zweite und das letzte Ziel dieser Auflistung relevant. Das erste Ziel wird dabei implizit durch die Testanalyse umgesetzt. Das dritte Ziel wird überwiegend in Form von Audits eruiert und bezieht sich vor allem auf die Komponente *Mensch*.

Im Abnahmetest geht es auch beim IT-Sicherheitstest, wie auch bei den zuvor besprochenen Qualitätsmerkmalen, nicht mehr primär um das systematische Aufdecken von IT-Sicherheitslücken. Dennoch sollten im regulatorischen und Benutzerabnahmetest insbesondere die funktionalen IT-Sicherheitsanforderungen überprüft werden sowie die zutreffenden IT-Sicherheitsrichtlinien der zu testenden Geschäftsprozesse. Dazu gehört im Zweifelsfall auch das Testen der IT-Sicherheitsprozesse gegen Social-Engineering-Angriffe.

Die Schwierigkeit bei IT-Sicherheitstests ist, dass der Bereich IT-Sicherheit sich über verschiedene Ebenen spannt. Neben den organisatorischen IT-Sicherheitsrichtlinien, die im Grunde nicht viel mit den dynamischen IT-Sicherheitstests der funktionalen Gegenmaßnahmen eines Produkts oder Softwaresystems gemein haben, hat auch eine Vielzahl rein technischer Maßnahmen ein hohes Gewicht, um IT-Sicherheitsschwachstellen abzusichern wie Firewalls, Intrusion Detection Systems (IDS), Verschlüsselung, Virens Scanner, Logfile-Analyse, um Angriffsmuster frühzeitig zu erkennen. Darüber hinaus befasst sich IT-Sicherheit nicht immer nur damit, Schwachstellen abzusichern, sondern es geht auch darum, die wertvollen Informationen des Unternehmens abzusichern, sodass im Falle eines Exploits (der Ausnutzung einer Schwachstelle durch einen Sicherheitsangriff) ein Angreifer keinen Nutzen daraus ziehen kann. Darunter fallen beispielsweise die Konzepte der Datenmaskierung. Auch die Konformität mit der entsprechenden Gesetzeslage (Stichwort: EU-DSGVO) bezüglich personenbeziehbarer und personenbezogener Daten gehört zu dem umfassenden Gebiet der IT-Sicherheit.

IT-Sicherheit auf unterschiedlichen Ebenen

Dadurch ergibt sich eine derart große Bandbreite an notwendigem Wissen, um IT-Systeme vor einem böswilligen Zugriff durch die Umwelt zu schützen, dass es nahezu unmöglich ist, alle relevanten Aspekte der IT-Sicherheit in einer Person zu vereinen. Zudem reichen sich statische und dynamische Sicherheitstests oftmals die Klinke in die Hand. Durch statische Analyse werden potenzielle Schwachstellen identifiziert, dynamische Tests versuchen diese Schwachstellen mit den entsprechenden Verfahren auszunutzen.

Umfangreiche Expertise oft gefordert

Neben der fachlichen Vielseitigkeit hinsichtlich der IT-Sicherheit gibt es darüber hinaus eine breite Palette von technischen Verfahren, die für die Simulation unterschiedlicher Gefährdungen eingesetzt werden können. Dazu gehören, wie schon erwähnt, Performanztestwerkzeuge, mit denen DoS-Attacken simuliert werden können. Penetrationstests sind Tests, die IT-Sicherheitsangriffe nachahmen mit dem Ziel, in das zu schützende System einzudringen, es zu penetrieren. Penetrationstests zielen in der Regel darauf ab, ein bereits produktives IT-System auf bestimmte IT-Sicherheitsaspekte zu prüfen, um ggf.

Maßnahmen zur Systemhärtung abzuleiten. Penetrationstests werden häufig von externen Dienstleistern erbracht.

Die Zunahme der fachlichen und technischen Komplexität führt dazu, dass einige der anstehenden IT-Sicherheitsabnahmetests durchaus von Abnahmetestern durchgeführt werden können – dazu zählen insbesondere Tests, die die Wirksamkeit von IT-sicherheitsrelevanten funktionalen Gegenmaßnahmen betreffen. Andere Ausprägungen der IT-Sicherheitstests erfordern aber sehr wahrscheinlich Nischenwissen und werden daher eher von spezialisierten IT-Sicherheitstestern durchgeführt.

Der »Sicherheitstester«-Lehrplan [ISTQB CTAL-ST] vertieft das Thema IT-Sicherheitstests. Das zugehörige Lehrbuch »Basiswissen Sicherheitstest« [Simon et al. 19] dringt noch deutlich tiefer in die Materie ein als der Lehrplan und wartet sowohl mit fundierter Theorie, wissenschaftlichen Ausblicken als auch einschlägigen Praxisbeispielen auf.

Index

A

Abnahme 47

Abnahmekriterium 26, 29

Erstellung 28–29, 47

Gute Praktiken 50

Identifikation und Spezifikation 47

Qualitätssicherung 50

Abnahmetest 9, 28, 45, 105

Alpha- und Beta-Test 12

Benutzerabnahmetest 11

betrieblicher Abnahmetest 12

Schwerpunkt 9

Testentwurf 51

Teststrategien/Testvorgehensweisen 53

Umfang 11

vertraglicher und regulatorischer Abnahmetest 12

vertrauensbildende Maßnahme 9

Zweck 13

Abnahmetestfall 45–46

Abnahmetestgetriebene Entwicklung 44, 56

Abweichung

Bewertung 191

Fehleranalyse 191

Agiles Projektmanagementwerkzeug 202

Aktivitäten 190

Akzeptanzkriterium 27
Anforderung 21, 25
 eindeutig dokumentiert 30
 Geschäftsanforderung 21
 interpretieren 30
 Produktanforderung 22
 Qualität 30–32
 Qualitätskriterien 31
Anforderungsbasiertes Testen 53
Anforderungsdefinition 28
Anforderungsmanagement 24
Anforderungsmanagementwerkzeug 201
Anforderungsspezifikation 21–22, 25, 48, 150, 198–199
Anwender 14
Ästhetik der Benutzungsschnittstelle 87
ATDD 44
Aufgabenanalyse 92
Auftraggeber 13
Auftragnehmer 13
Auswirkungsanalyse 195
Authentizität 113

B

Barrierefreiheit 87
BDD 45
Benutzerbefragungen 97
Benutzererlebnis 88
Benutzerfehlerschutz 86
Biometrische Evaluierungen 99
BPMN 124
Business Case 18
Businessanalyse
 Lösungsbewertung und -optimierung 34
 Management 34

Requirements Engineering 34

sequenzielle 27

Strategiedefinition 34

Businessanalysten 14

C

Checklistenbasierte Evaluierungen 99

D

Daily Scrum 195

Dauertest 102

Decision Model and Notation 134

Definition of Ready 28, 32

DMN 134

DMN-Entscheidungstabelle 141, 143

Dynamische Analyse 103

E

Entscheidungsregel 136, 170

aktionsgleich 138

Entscheidungstabelle 136

eingabebegrenzt 137

eingabeerweitert 139

reduziert 138

vollständig 137, 171

Epic 24

Erfahrungsbasiertes Testen 54

Erkennbare Angemessenheit 85

Erlernbarkeit 86

Expertenreviews der Gebrauchstauglichkeit 99

F

Falsch negatives Ergebnis 30

Falsch positives Ergebnis 191

Feature 24

Fehleranalyse 191–192

Fehlerbehebungskosten 10
Fehlerbericht 191–192
Fehlereindämmung 11
Fehlermanagementwerkzeug 203
Fehlerzustand 29
Formative Evaluierung 97
Freigabe 47, 196
 keine 197
 mit Einschränkungen 196
Frühes Testen 11
Funktionale Eignung 71
Funktionale Sicherheit 110

G

Gebrauchstauglichkeit 72, 84
Gebrauchstauglichkeitsevaluierung 97
Gebrauchstauglichkeitstest 97
Geschäftsauswirkung 191
Geschäftsbedarf 15, 18
Geschäftslösung 15, 18
Geschäftsproblem 15
Geschäftsprozess 15, 124
Geschäftsprozessbasiertes Testen 54
Geschäftsprozessmanagementwerkzeug 202
Geschäftsprozessmodell 144, 157
Geschäftsprozesssicherung 123
Geschäftsregelmodell 144, 157
Geschäftsziele 15
Gherkin 56, 58
Given-When-Then 56

H

Hauptaktivitäten der Businessanalyse 33
Hauptaktivitäten des Testens 35

I

Inbetriebnahme 12
Informationssicherheit 115
Integrationstest 8
Integrität 112
INVEST 32
ISO 25010 71, 75, 77
IT-Sicherheit 73, 110–111
IT-Sicherheitsabnahmetest 118

K

Kapazität 101
Kapazitätstest 103
Klassische Entscheidungstabelle 136
Kollaboration 47
Kompatibilität 72
Komponentenintegrationstest 104
Komponententest 8, 104
Kontextanalyse 92
Kontinuierliche Integration 45
Kritikalität 191
Kunde 13

L

Lastgenerator 109
Lastprofil 109
Lastspitzentest 102
Lasttest 102

M

Modell 122
Modellbasiertes Testen 145
Modellbasiertes Testwerkzeug 203
Modellgetriebene Softwareentwicklung 121
Modellierung 121
Modellierungspraktik 150
Modellierungssprache 123–124

Monitoring 109

N

Nichtabstreitbarkeit 112

Nutzeranalyse 91

Nutzungskontext 81

- Arbeitsaufgaben 83

- Arbeitsmittel 83

- Benutzer 82

- Umgebung 83

- Ziele 82

Nutzungsprofil 108

Nutzungsqualität 77

- Effektivität 78

- Kontextüberdeckung 79

- Risikofreiheit 79

- Zufriedenheit 78

O

Operabilität 86

P

Performanz 72, 100

Performanzabnahmetest 106, 108

Performanztest 102

Perspektivisches Lesen 48

Portabilität 73

Product Backlog 25

Product Owner 14

Protokolldateianalyse 100

Prüfsummenverfahren 138, 173

Q

Qualitätsmerkmal 71, 75

- funktional 71

- nicht funktional 71, 84

Qualitätsmodell 71
Qualitätssicherung der Testmittel 197
 Review der Abnahmekriterien 197
 Review der Abnahmetestfälle 197
 Review der Testberichte 198
 Verfolgbarkeit 197

R

Ressourcennutzung 101
Reviews auf Gebrauchstauglichkeit 98
Risikobasiertes Testen 53
Rolle
 Businessanalyst 39
 Tester 39
Rot-Grün-Refactoring 44

S

Schlüsselwortgetriebenes Testen 58
Schnittstellenart 56
Shift-Left 29, 43
Skalierbarkeitstest 103
Social Engineering 114
Softwareentwicklungslebenszyklus 8
Softwareentwicklungsmodell 2, 9, 12–13, 28
 agil 9, 12, 25, 27, 43, 193, 195
 sequenziell 9, 12, 25–26, 41, 43
Stakeholder 15
Statischer Test 11
Stresstest 102
Summative Evaluierung 97
Systemintegrationstest 105
Systemtest 8, 105

T

Testabschluss 37
Testabschlussbericht 195

Testanalyse 26, 36
Testautomatisierungswerkzeug 203
Testbarkeit 26, 29
Testbasis 26
Testbedingung 26
Testdurchführung 29, 37
Testentwurf 26, 36
Testfall 26
 abstrakt 159
 konkret 159
Test-First 43
Testfortschrittsbericht 193
Testgetriebene Entwicklung 44
Testmanagementwerkzeug 202
Testmetrik 194
Testmodell 151
Testobjekt 29
Testplanung 36
Testprozess 35
Testrealisierung 36
Teststeuerung 36
Teststrategie 51
 analytisch 51
 angeleitet 52
 methodisch 52
 modellbasiert 52
 prozesskonform (oder standardkonform) 52
 reaktiv 52
 regressionsvermeidend 52
Teststufe 8, 10, 35
Testüberwachung 36
Testverfahren 53
 Blackbox 53
 erfahrungsbasiert 53

Whitebox 53

Testvorgehensweise 53

Themen 24

Think-Aloud-Techniken 99

U

Usability 84

Usability-Tests 98

User eXperience 88

User Story 12, 24

UX-Anforderungsanalyse 91

V

Verfolgbarkeit 54, 151

Verhaltensgetriebene Entwicklung 45, 56

Vertrauen 10

Vertraulichkeit 112

V-Modell 7

- entwicklungsbegleitend 42

- sequenziell 41

- XT 14

W

Walkthrough 99

Wartbarkeit 73

Werkzeugart 201

Werkzeugunterstützung 198

Wettbewerbsanalyse 92

Z

Zeitverhalten 100

Zugänglichkeit 87

Zurechenbarkeit 113

Zusammenarbeit 40, 47, 189–190

Zuverlässigkeit 72