

RHEINISCHE FACHHOCHSCHULE KÖLN

University of Applied Sciences

Fachbereich: Wirtschaft & Recht

Studiengang: Wirtschaftsinformatik (M.Sc.)



Masterthesis

Konzeption einer Reifegradmessung
der Software-Qualitätssicherung im öffentlichen Sektor

Thesis vorgelegt von:

Mike Wittsiepe

531182014

Sommersemester 2020

RHEINISCHE FACHHOCHSCHULE KÖLN

University of Applied Sciences

Fachbereich: Wirtschaft & Recht

Studiengang: Wirtschaftsinformatik (M.Sc.)



Masterthesis

Konzeption einer Reifegradmessung
der Software-Qualitätssicherung im öffentlichen Sektor

Thesis vorgelegt von: Mike Wittsiepe

531182014

1. Prüfer: Prof. Dr. Ralph Trittman

2. Prüfer: Prof. Dr. Thomas Barth

Sommersemester 2020

Vorwort

Diese Arbeit ist mit der Absicht verfasst worden die Softwarequalitätssicherung als solches zugänglicher zu gestalten und möglicherweise sogar Vorurteilen entgegen zu wirken.

Vorab noch ein paar Worte des Dankes an die Menschen in meinem Umfeld:

Ein großer Dank geht an das Bundesverwaltungsamts und vor allem an das Referat, in dem die Erprobung durchgeführt werden durfte. Die Ergebnisse stellten einen wertvollen Beitrag zu dieser Thesis dar. Ebenfalls gilt mein Dank an den Fachbereich IT-Verfahren Öffentliche Sicherheit, die Projektgruppe Smart Borders sowie Dr. Prpitsch die die Anfertigung der Thesis in dieser Form erst ermöglichten. Ein ganz besonderer Dank gilt Dr. Oliver Kortendick, der die Arbeit auf Seiten des BVA betreute. Oliver, vielen Dank!

Auch gilt ein Dank an Freunde und Arbeitskollegen, die in der gesamten Studienzeit viel Verständnis und Rücksicht zeigten und auch mit dem einen oder anderen Rat zur Seite standen.

Und da auch am 21. März 2020 meine Hochzeit stattfand, gilt abschließend selbstverständlich mein größter Dank meiner Ehefrau Nadine, die mir auch in den schwersten Zeiten den Rücken stärkte und mir, wenn notwendig, umgangssprachlich auch mal in den Allerwertesten trat.

Kurzfassung

Diese Masterthesis stellt das Konzept eines Verfahrens zur Messung des Reifegrades von Softwarequalitätssicherungstätigkeiten zur Anwendung in Entwicklungsverfahren des öffentlichen Sektors vor. Ziel ist es unterschiedlichen Organisationseinheiten ein leichtgewichtiges und fokussiertes Verfahren zur Verfügung zu stellen. Schlussendlich sollen die Ergebnisse aus diesem Verfahren dabei helfen, neue oder potentielle Verbesserungsmöglichkeiten im Bereich der eigenen Softwarequalitätsmaßnahmen aufzudecken. Auch wird im Rahmen dieser Masterthesis eine Erprobung dieses Verfahrens durchgeführt.

Das hier entwickelte Verfahren erhebt nicht den Anspruch Konkurrenz zu bestehenden Testprozessverbesserungsmethoden zu sein. Es soll vielmehr einen vereinfachten Einstieg und einen ersten Schritt in die Welt der Testprozessverbesserung ermöglichen.

Abkürzungsverzeichnis

Abkürzung	Darstellung des abgekürzten Begriffes
a. d.	an der
Abs.	Absatz
Abschn.	Abschnitt
ALM	Application-Lifecycle-Management System
BMA	Brandmeldeanlage
BMI	Bundesministerium des Innern, für Bau und Heimat
BVA	Bundesverwaltungsamt
bzw.	beziehungsweise
ca.	circa
CD	Compact Disc
CMM	Capability Maturity Model
CMMI	Capability Maturity Model Integration
CTFL	Certified Tester Foundation Level
d.h.	das heißt
DAAD	Deutscher Akademischer Austauschdienst
Dev	Developer
Dr.	Doktor
e.g.	exempli gratia (zu Deutsch: zum Beispiel)
e.V.	eingetragener Verein
EMA	Einbruchmeldeanlage
etc.	et cetera
FDA	Food and Drug Administration
ForTeQ	Forum für Test und Qualitätssicherung
ggf.	gegebenenfalls
GmbH	Gesellschaft mit beschränkter Haftung
GPM	Deutsche Gesellschaft für Projektmanagement e. V
GTB	German Testing Board
HP	Hewlett-Packard
ID	Identifikator
IEC	International Electrotechnical Commission
IEEE	Institute of Electrical and Electronics Engineers
IOT	Internet of Things
IS	Information system
ISO	Internationale Organisation für Normung
ISTQB	International Software Testing Qualifications Board
IT	Informationstechnik
ITZ	Informationstechnikzentrum
Kap.	Kapitel
KEx	kombinatorische Erkenntnis x
KONSENS	Koordinierte Neue Software-Entwicklung der Steuerverwaltung
M. Sc.	Master of Science
MF	Micro Fokus

Min.	Minuten
MS	Microsoft
Nr.	Nummer
o. J.	ohne Jahr
PDF	Portable Document Format
Prof.	Professor
QS	Qualitätssicherung
RFH	Rheinische Fachhochschule
RG	Reifegrad
RGx	Reifegrad x
ROBASO	rollenbasierten Oberfläche
ROM	Read-Only Memory
s.	siehe
SB	Smart Borders
SEI	Software Engineering Institute
SG	Specific Goal
SLES	SUSE Linux Enterprise Server
sog.	sogenannte
SP	Specific Practices
SPICE	Software Process Improvement and Capability Determination
TH	technische Hochschule
TMA	test maturity assessment
TMap	Test Management Approach
TMMi	Test Maturity Model integration
TPI	Test Process Improvement
u.a.	unter anderem
vgl.	vergleiche / vergleichsweise
vs.	versus
z. B.	zum Beispiel

Inhaltsverzeichnis

Vorwort.....	I
Kurzfassung	II
Abkürzungsverzeichnis	III
1 Einleitung	1
1.1 Problemstellung und Forschungsfrage	2
1.2 Forschungsantwort und Ziel der Thesis	4
1.3 Umfeld der Masterthesis.....	7
1.3.1 Öffentlicher Sektor.....	8
1.3.2 Bundesverwaltungsamt.....	9
1.3.3 Organisationseinheit.....	10
1.4 Methodisches Vorgehen und Forschungsdesign.....	10
2 Stand der Forschung und Wissensbasis	12
2.1 Softwarequalitätssicherung	13
2.2 Testprozess Reifegradmessungs- bzw. Bewertungsverfahren	15
2.2.1 TPI / TPI-Next.....	17
2.2.2 CMMI / TMMi	19
3 Entwicklung eines Verfahrens zur Reifegradmessung. 25	
3.1 Ziel eines Verfahrens zur Reifegradmessung.....	25
3.2 Bestimmung der Verfahrensart	27
3.3 Population und Stichprobe einer Befragung	32
3.4 Reifegradmodell.....	34
3.5 Gestaltung eines Fragebogens zur Reifegradmessung	38
3.5.1 Definierung der Form und Verteilung.....	38
3.5.2 Gestaltung einer Einleitung	38
3.5.3 Entwicklung der Merkmalsbereiche und Umfang	41
3.5.4 Entwicklung der Fragestellungen.....	44
3.5.5 Pre-Test des Fragebogens	69

3.6	Beschreibung der Reifegradstufen.....	70
3.6.1	Testprozess (Item 1 - 3).....	71
3.6.2	Testmanagement (Item 4 – 8).....	72
3.6.3	Reviews (Item 9 - 11).....	72
3.6.4	Fehlermanagement (Item 12 – 15).....	73
3.6.5	Testwerkzeuge und Automatisierung (Item 16 – 19).....	73
3.6.6	Soziale Kompetenz und Teamzusammensetzung (Item 20 – 22).....	74
3.7	Abschluss der Artefakt-Entwicklung.....	74
4	Erprobung des entwickelten Werkzeugs.....	75
4.1	Auswahl des Untersuchungsgegenstandes.....	75
4.2	Abschluss und Dokumentation der Ergebnisse.....	76
4.3	Ergebnisse der Reifegrad Bewertung.....	76
4.4	Analyse der Datenbasis.....	77
4.4.1	Merkmal: Testprozess (Item 1 – 3).....	78
4.4.2	Merkmal: Testmanagement (Item 4 – 8).....	79
4.4.3	Merkmal: Review (Item 9 – 11).....	81
4.4.4	Merkmal: Fehlermanagement (Item 12 – 15).....	82
4.4.5	Merkmal: Testwerkzeuge und Automatisierung (Item 16 – 19).....	84
4.4.6	Merkmal: Soziale Kompetenz und Teamzusammensetzung (Item 20 – 22).....	85
4.5	Kombination der Erkenntnisse einzelner Items.....	86
5	Validierung des Reifegradverfahrens.....	89
5.1	Interviewgestaltung.....	90
5.1.1	Interviewabschnitt Eins.....	90
5.1.2	Interviewabschnitt Zwei.....	90
5.1.3	Interviewabschnitt Drei.....	91
5.2	Ergebnisse des Interviews.....	91
5.3	Analyse der Interviewergebnisse und Validierung der Reifegradmessung...	94
5.3.1	Validierung Reifegradmessung.....	95
5.3.2	Validierung der kombinatorischen Ergebnisse.....	96
6	Abschlussbewertung.....	98

6.1 Kritische Würdigung und Entwicklungsräume.....	98
6.1.1 Gewählte Themen der Softwarequalitätssicherung.....	98
6.1.2 Detailtiefe der Items	99
6.1.3 Eingeschränkte Evaluation.....	99
6.1.4 Struktureller Aufbau der Befragung bzw. Fragebogens	99
6.2 Fazit	100
Abbildungsverzeichnis	101
Tabellenverzeichnis.....	102
Anhangsverzeichnis.....	103
Literaturverzeichnis.....	153
Eigenständigkeitserklärung	

1 Einleitung

Diese Masterthesis wird im Rahmen des vierten Semesters des Studiengangs Wirtschaftsinformatik (Master of Science) an der Rheinischen Fachhochschule Köln erarbeitet. Im Kern wird sich in dieser Masterthesis mit der Entwicklung eines Verfahrens zu Reifegradmessung von Softwarequalitätssicherung auseinandergesetzt.

Jedes Softwareentwicklungsprojekt agiert in seinem Softwareentwicklungslebenszyklus innerhalb seiner eigenen Rahmenbedingungen und Organisationsstrukturen. Dennoch besteht dieser Zyklus mindestens aus den folgenden Tätigkeiten:

- Anforderungsaufnahme
- Design und Entwicklung
- Qualitätssicherung

Diese Aufzählung erhebt hierbei keinen Anspruch auf Vollständigkeit. In der Regel beziehen sich Softwareentwicklungsprojekte oder -verfahren auf ein Vorgehensmodell, welches diese Tätigkeiten in Art und Durchführung beschreibt (z. B. Wasserfall, V-Modell, Agile)¹. Jede dieser Tätigkeiten stellt Unternehmen oder ähnliche Organisationen vor Herausforderungen. Konkrete Betrachtung innerhalb dieser Thesis findet daraus der Bereich um die Qualitätssicherung von Softwareentwicklungsprojekten bzw. Softwareentwicklungsverfahren im öffentlichen Sektor.

Im weiteren Verlauf der Einleitung wird auf die Problemstellung und die konkrete Lösungsidee eingegangen. Daraufhin werden kurz die Rahmenbedingungen der Verfahrensentwicklung beschrieben sowie konkreter auf das methodische Vorgehen eingegangen. Zur Einführung des Hauptteils dieser Masterthesis erfolgt eine kurze Erläuterung des Themas Softwarequalitätssicherung und bestehender Testprozessverbesserungsmethoden. Kapitel 3 bis 5 stellen den Kern dieser Thesis dar. Konkret bearbeitet Kapitel 3 die Entwicklung und Herleitung des Reifegradverfahrens. Nach abgeschlossener Entwicklung wird in Kapitel 4 auf die Erprobung des Verfahrens in einer Organisationseinheit des öffentlichen Sektors eingegangen. Kapitel 5 greift diese Ergebnisse auf und validiert diese. Zum Abschluss erfolgt eine Abschlussbewertung der Thesis.

¹ Erläutert in (Spillner und Linz 2019: 53–62).

1.1 Problemstellung und Forschungsfrage

Für die Erarbeitung der Problemstellung bis hin zur Forschungsantwort wird sich im Folgenden an dem Werk „The craft of research“ orientiert².

Auf Basis persönlicher Erfahrung bzw. Einschätzung und Beobachtungen des Verfassers scheint gerade der Bereich Softwarequalitätssicherung im Bundesverwaltungsamt ein hohes Maß an Entwicklungspotential aufzuweisen. Der bestehende Einsatz von Qualitätssicherungstätigkeiten wirkt unkoordiniert und heterogen. Verschiedene Austausche auf Arbeitsebene innerhalb der Behörde des Verfassers bestätigen diesen Eindruck zusätzlich und zeigen auf, dass viele Mitarbeiter unsicher sind über ihr weiteres Vorgehen. Die wenigen bestehenden Ressourcen möchten sie zurückhaltend und gezielt einsetzen. Nach Rücksprache innerhalb des BVA mit einigen Verantwortlichen bestehen aus deren Sicht keine oder wenig Informationen darüber, welche Entwicklungsverfahren in welchem Umfang und Intensität Qualitätssicherung betreiben. Weitere Nachforschung ergab, dass bereits der Bundesrechnungshof in seinem Bericht über die Querschnittsprüfung zur Sicherung der Softwarequalität in IT-Vorhaben der Bundesverwaltung aus dem Jahr 2008 vergleichbares feststellte. Konkret wurden folgende Probleme benannt³:

- Es bestanden keine übergreifenden Qualitätssicherungsprozesse oder -modelle. Jedes Verfahren musste eigene Vorgehensweisen im Bereich der Qualitätssicherung entwickeln.
- Aufgaben im Bereich der Qualitätssicherung die an externe Dienstleister vergeben worden sind, wurden unzureichend geregelt, was wiederum zu weiteren Kosten führte.
- Fehlende Vorgaben für Dokumente und die Kontrolle von Dokumenten führten zu zeitlichen Verzögerungen und Mehraufwände.
- Anforderungen an die Produktqualität und Ansprüche an den Entwicklungs- sowie Qualitätssicherungsprozess von externen Dienstleistern wurden häufig nur sehr vage formuliert.
- Qualitätsmeilensteine in Projektplanungen entsprachen nicht dem realen Projektverlauf und wurden nicht angepasst.
- Fehlende Metriken im Bereich der Fehlernachverfolgung (z.B. Fehlerbehebungszeiten, Fehlerquellen, etc.).

² Siehe Kapitel 4 und die dort beschriebene Struktur des Praxisproblem, Forschungsfrage, Forschungsproblem und Forschungsantwort. (Booth u. a. 2016: 57–43).

³ Siehe dazu (Bundesrechnungshof 2008: 3–5).

- Nur selten wurden bestehende Regressionstests zur Testautomatisierung überführt.

Es lässt sich daraus schließen, dass sich die Verantwortlichen in einer Situation befinden, in der sie nicht einschätzen können, welche Qualitätssicherungsmaßnahmen aktuell in Ausführung sind, welche Wirkung diese erzielen und wo ggf. neue Ansätze notwendig sind. Weitere Recherchen im Bereich der Behörden über das Bundesverwaltungsamt hinaus weisen ebenfalls auf vergleichbare Herausforderungen hin. Eine Bemerkung des Bundesrechnungshofes aus dem Jahr 2009 zeigte verschiedene Probleme der Bundeswehr bei der Abnahme und infolgedessen Verzögerungen bei der Planung eines IT-Systems⁴. Ein weiterer Bericht⁵ des Bundesrechnungshofes aus dem Jahr 2018 bemängelte bei der Bundesagentur für Arbeit eine unzureichende Anforderungsumsetzung und insgesamt deutliche Verzögerungen, die auf Grund von technischen Schwierigkeiten an Schnittstellen entstanden sind. Weiter bemängelt wurden die fehlerhafte Umsetzung von Geschäftsprozessen und dass Erkenntnisse über Fehler nicht ausreichend in die Projektplanung mit einbezogen worden waren. Folgende Aussage lässt den Rückschluss zu, dass hier das genutzte Vorgehen in der Entwicklung und die Qualitätssicherung zu diesem Verlauf wesentlich beigetragen hatten:

„Insgesamt führt die Bundesagentur das späte Erkennen der Probleme, die zum Scheitern bei ROBASO führten, auf die gewählte Projektmethodik zurück. Sie habe versucht zunächst alle Kernprozesse zu entwickeln, um erst am Schluss zu testen, ob diese funktionierten. Nach der Erfahrung bei ROBASO würde sie heute die Entwicklung der Software in einzelnen Teilmodulen vornehmen und testen (sog. agile Softwareentwicklung).“⁶

Verstärkt wird dieser Eindruck zusätzlich durch eine Antwort der Bundesagentur für Arbeit auf eine Anfrage von heise.de⁷. Im Jahr 2019 folgte ein Bericht⁸ des Bundesrechnungshofes zum aktuellen Stand und zu den Fortschritten des Zusammenwirkens von Bund und Ländern im Vorhaben KONSENS. In diesem wurde auf das Fehlen eines anforderungsgerechten Evaluationskonzeptes hingewiesen. Dieses sollte laut dem Bericht qualitative und quantitative Kennzahlen für das Anforderungsmanagement, die Produktivsetzung und die Qualität der Software beinhalten. Abschließend wurde im Bericht erläutert, dass solch eine Kennzahl auch für den Gesetzgeber relevant sein könnte. All

⁴ Vgl. (Bundesrechnungshof 2010).

⁵ Vgl. (Bundesrechnungshof 2018).

⁶ (Bundesrechnungshof 2018: 19)

⁷ Siehe dazu Artikel auf heise.de (Kannenberg 2017).

⁸ Vgl. (Bundesrechnungshof 2019).

diese Berichte zeigen mindestens einen der typischen Gründe⁹ auf, aus denen Softwareentwicklungsprojekte potentiell beendet werden. Die Schlussfolgerungen von Kweku E-wusi-Mensah aus dem Jahr 1997 lauteten damals:

„Executives should seize every opportunity to become fully informed and knowledgeable about all the major issues and aspects of the development process.“¹⁰.

Diese grundsätzliche Aussage, eigene Erfahrungen und die zahlreichen Berichte führen den Verfasser dazu folgendes Praxisproblem abzuleiten:

Es liegen wenig bis keine Informationen darüber, vor welche Qualitätssicherungsreife-grade in den unterschiedlichen behördlichen Softwareentwicklungsverfahren bestehen.

Die sich daraus ableitende Forschungsfrage lautet wie folgt:

Wie kann im öffentlichen Sektor der Reifegrad einer Softwarequalitätssicherung be-stimmt werden?

1.2 Forschungsantwort und Ziel der Thesis

Mit Blick auf die Forschungsfragen bietet es sich an, die Methoden (u.a. TPI / TPI-Next oder TMMi) zur Testreifegradmessung und Bewertung als mögliche Antwort in Betracht zu ziehen. Verschiedene Arbeiten beschäftigten sich bereits mit diesen Methoden, de-ren Ziel die Verbesserung und Reifegradmessung von Testprozessen ist. Folgende Herausforderungen bzw. Hindernisse für die Einführung und Nutzung solcher Metho-den wurden bereits identifiziert¹¹:

- Fehlende Ressourcen.
- Fehlendes Know-How.
- Verbesserungen werden als Mehraufwand wahrgenommen.
- Scheinbar keine erkennbaren Vorteile.
- Unklare Ziele oder Fokus.
- Fehlende Verantwortlichkeiten und Verbindlichkeiten.
- Andere.

⁹ Vgl. (Ewusi-Mensah 1997).

¹⁰ (Ewusi-Mensah 1997: 80)

¹¹ Siehe dazu (Garousi u. a. 2017: 22).

Ebenfalls wird anscheinend¹² für eine erfolgreiche Anwendung umfassendes Wissen über das Softwaretesten als solches vorausgesetzt. Es lässt sich hiervon ableiten, dass gerade für Einsteiger und unerfahrene Organisationen in Kombination mit den aus dem Behördenbereich beispielhaft dargestellten Situationen diese Methoden nicht erfolgsversprechend scheinen. Gestärkt wird dies durch folgende Aussage von Spillner über diese Vorgehensweisen:

„Wie gut sich ein bestimmtes Modell eignet, kann nur individuell im jeweiligen Unternehmen entschieden werden. Eine generelle Empfehlung kann nicht ausgesprochen werden.“¹³

Es ist allerdings zu beachten, dass damit nicht ein grundsätzlicher Erfolg dieser Methoden in Organisationen und Situationen in Frage gestellt wird. Aktuelle Umfragen lassen allerdings eine fehlende Relevanz der Verfahren vermuten. So könnte in Betracht gezogen werden, von diesen Methoden gänzlich abzusehen. Eine Umfrage¹⁴ für den Zeitraum September 2015 bis Januar 2016 der Hochschulen Bremerhaven, TH Köln und Bremen, behandelt unter anderem auch Fragenstellungen zu diesen Themen. Zehn Prozent der Teilnehmer ordneten sich hier der Branche „öffentliche Hand“ zu. Das Ergebnis ist, dass gerade im unternehmerischen sowie Forschungskreisen Audits wie TPI / TPI Next, TMMI und andere, keine oder nur wenig Relevanz besitzen. 69,9% gaben an, dass sie keine Audits in dieser Form durchführen. Dies scheint auch keine neuerliche Entwicklung zu sein, da in der Umfrage¹⁵ von 2011 „Softwaretest in der Praxis“ 56% der Teilnehmer angaben, keine Audits ihrer Testprozesse durchzuführen. Einige Freitextmeldungen¹⁶ aus der Umfrage von 2016 lassen vermuten, dass durchaus interne Formen von Audits stattfinden, die sich an verschiedenen bekannten Normen der Softwarequalitätssicherung orientieren. An den Umfragen von 2011 und 2016 waren Teilnehmer aus den Ländern Deutschland, Österreich und der Schweiz beteiligt. Es gilt allerdings zu berücksichtigen, dass die meisten Teilnehmer in beiden Umfragen aus Deutschland stammten¹⁷. Das ISTQB stellte 2018 ihre Umfrage¹⁸ „ISTQB® Worldwide Software Testing Practices Survey 2017-18“ vor. Hier nahmen aus 92 Ländern über 2000 Personen

¹² Vgl. (Afzal u. a. 2016: 31).

¹³ (Spillner u. a. 2014: 392)

¹⁴ Siehe dazu die Zusammenfassung für Forschung (Spillner u. a. 2016), für Unternehmen (Simon u. a. 2016) und im Detail (Spillner u. a. 2017b) sowie (Spillner u. a. 2017a).

¹⁵ Siehe dazu (Vosseberg u. a. 2013).

¹⁶ Siehe dazu (Spillner u. a. 2017a: 106–110).

¹⁷ 2011 Deutschland mit 77% vgl. (Haberl u. a. 2012: 4) und 2016 Deutschland mit 90% vgl. (Spillner u. a. 2017b: 3).

¹⁸ Siehe dazu (Linetzki u. a. 2018).

teil. 27% der Teilnehmer stammten aus Europa. Davon ordneten sich allerdings nur 3,1 dem in der Umfrage benannten „Public Sector“ zu. In dieser Umfrage wird unter anderem auf Testprozesse im Allgemeinen eingegangen. 46,5% der Befragten gaben an, dass sich ihr Wissen über ihre Testprozesse zu vorher verbessert hatte. Und obwohl dies ein hoher Wert zu sein scheint, gaben in derselben Umfrage 22% der Befragten an, dass sie das sogenannte Thema¹⁹ „Test Process Improvement“ als ein Hauptthema für die Zukunft sehen. Aus diesen Informationen lässt sich schließen, dass das Thema Testprozesse als solches sowie die Testprozessverbesserung ein aktuelles Thema für die Teilnehmer der Umfrage darstellte. Die Forschungsfrage und das konkrete Praxisproblem in Zusammenspiel mit den erläuterten Sachverhalten zur Testreifegradmessung führen daher zu folgendem Forschungsproblem:

Es existiert kein begründetes Verfahren, um eine Reifegradmessung der Softwarequalitätssicherung im öffentlichen Sektor durchzuführen.

Die sinkende Relevanz, der hohe Aufwand, die hohe Komplexität der Bestandsverfahren, aber auch das konstante Interesse an einer Verbesserung weisen darauf hin, dass gerade mit Blick auf den öffentlichen Sektor ein einfaches und leichtgewichtiges Verfahren hier Abhilfe schaffen kann. Die vorgeschlagene Forschungsantwort sowie das Ziel dieser Masterthesis sind somit wie folgt:

Konzeption eines begründeten Verfahrens zur Reifegradmessung der Softwarequalitätssicherung im öffentlichen Sektor.

Zielgruppe für solch ein Verfahren sind grundsätzlich Mitarbeiter einer Organisation im öffentlichen Sektor (z.B. einer Behörde). Solch ein Mitarbeiter könnte beispielsweise ein Qualitätsbeauftragter, ein Sachgebietsleiter, ein Komponentenverantwortlicher, ein Fachverfahrensentwickler oder auch Referatsleiter sein. Die Ergebnisse eines solchen Verfahrens würden dabei helfen mit einer einfachen Methode einen initialen Eindruck über den Reifegrad der eigenen Qualitätssicherungsvorgehen zu ermöglichen. Es sollte allerdings beachtet werden, dass ein solches Verfahren und dessen Ergebnisse nicht direkt die zuvor beschriebenen Situationen vermeiden oder konkrete Problemlösungen dafür anbieten. Es wäre viel mehr als Hilfestellung zu betrachten. Aus dessen Erkenntnissen können die behördlichen Organisationen einen initialen Eindruck ihres aktuellen Qualitätssicherungsvorgehens erhalten ohne umfangreiche und je nach Organisationsgröße ggf. sehr aufwändige Testreifegradmessungen durchführen zu müssen. Die Erkenntnisse können daraufhin potentiell dazu dienen vergleichbaren Situationen entgegen zu wirken. Welche Maßnahmen am Schluss aus den Ergebnissen abgeleitet werden und wie diese in der Organisation umgesetzt werden, kann durch dieses Verfahren nicht

¹⁹ Vgl. (Linetzki u. a. 2018: 38).

beeinflusst werden. In Abbildung 1-1²⁰ findet sich eine grafische Zusammenfassung zu den Kapiteln 1.1 sowie 1.2.

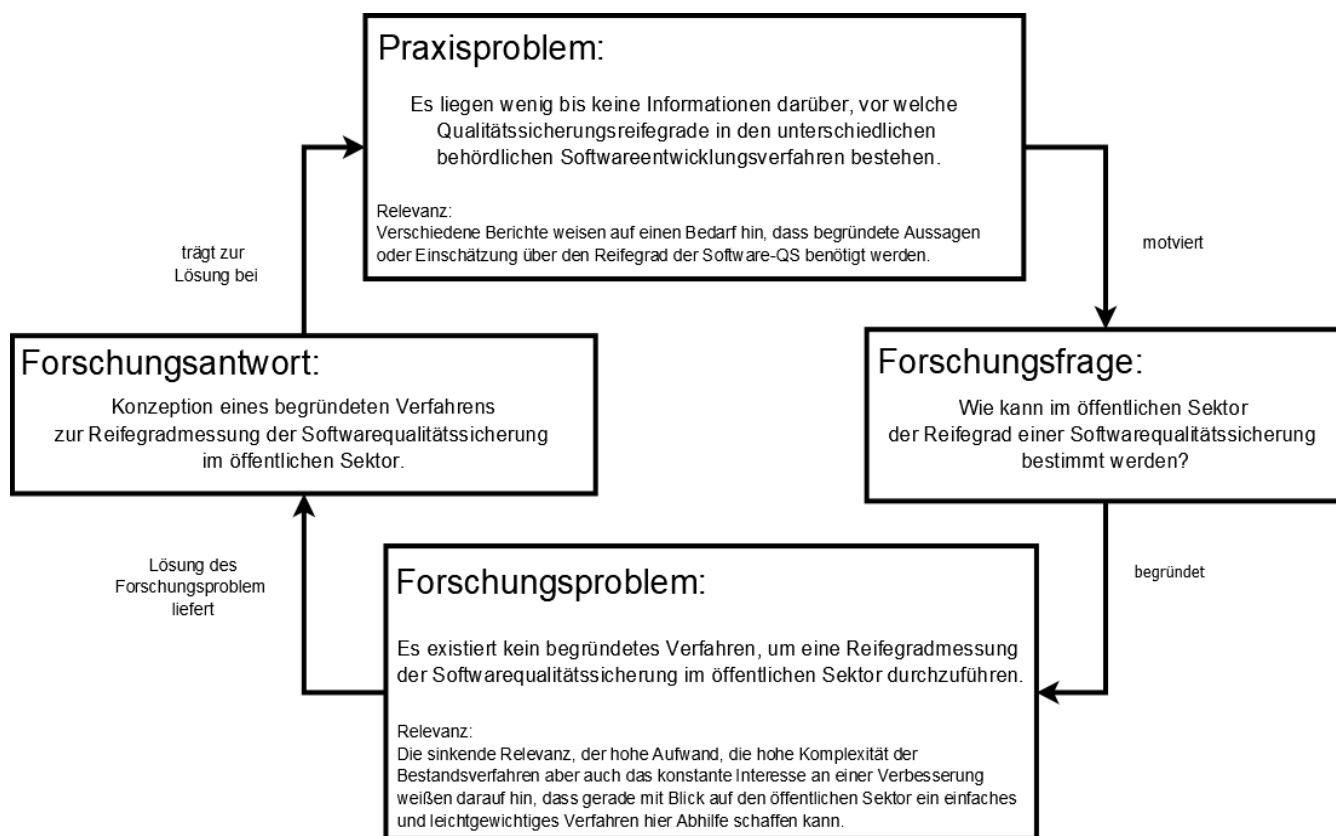


Abbildung 1-1: Herleitung Forschungsantwort

1.3 Umfeld der Masterthesis

Der Autor dieser Masterthesis befindet sich zurzeit in einer Anstellung im Bundesverwaltungsamt. Dort ist er seit 2017 als Testmanager der Projektgruppe Smart Borders²¹ für die Qualitätssicherungsprozesse der Softwareentwicklung verantwortlich. Weitere Details können dem Lebenslauf in Anhang 33 entnommen werden. Die Aufgaben des Testmanagement werden durch Spillner wie folgt zusammengefasst:

²⁰ Aufbau der Abbildung übernommen von (Booth u. a. 2016: 59).

²¹ Die Projektgruppe Smart Borders (kurz: SB) ist Teil des „Fachbereich IT-Verfahren Öffentliche Sicherheit innerhalb der IT-Abteilung“ des BVA. Siehe weitere Details im Organigramm des BVA (Bundesverwaltungsamt 2020).

„Das Testmanagement umfasst klassische Methoden des Projektmanagements und des Risikomanagements sowie das Wissen um den zweckmäßigen Einsatz wohldefinierter Testentwurfsverfahren.“²².

Das bedeutet, dass ein Testmanagement mit den gleichen Werkzeugen und Prozessen wie ein Projektmanagement vorgeht. Im Weiteren werden nun die Begriffe „öffentlicher Sektor“ aus dem Titel der Masterthesis und Organisationseinheit erläutert sowie kurz auf die Hintergründe des Bundesverwaltungsamts eingegangen.

1.3.1 Öffentlicher Sektor

Der öffentliche Sektor definiert sich nach dem Handbuch des ökonomischen Systems der Bundesrepublik Deutschland wie folgt:

„Der öffentliche Sektor einer Volkswirtschaft umfasst alle wirtschaftlichen Aktivitäten staatlicher Institutionen.

...

Die in Deutschland nach der Definition der (->) Volkswirtschaftlichen Gesamtrechnung eindeutig zum öffentlichen Sektor gehörenden Institutionen sind die Regierungen der Gebietskörperschaften, Fachbürokratien wie die Deutsche Bundesbank, das Bundeskartellamt und die (->) Bundesanstalt für Arbeit sowie die Träger der Gesetzlichen (->) Sozialversicherungen.“²³

Das schließt alle Behörden des Bundes, des Landes sowie Kommunen und Gemeinden ein. Behörde ist hierbei im Sinne des §1 Abs. 4 Verwaltungsverfahrensgesetz zu verstehen:

„Behörde im Sinne dieses Gesetzes ist jede Stelle, die Aufgaben der öffentlichen Verwaltung wahrnimmt.“

Zur Vereinfachung schließt der Verfasser für diese Thesis ebenfalls Leistungen privatwirtschaftlicher Unternehmen mit ein, die durch eine Behörde beauftragt werden, im weiteren Kontext auch externe Dienstleister genannt. Diese Leistungen können dabei zum Beispiel auch vor Ort innerhalb der Behörden erbracht werden²⁴. Innerhalb dieser Thesis werden alternativ für den Begriff „öffentlicher Sektor“ folgende Begriffe ebenfalls genutzt: Organisation, behördlich sowie Behördenbereich.

²² (Spillner u. a. 2014: 1)

²³ (Schubert 2005: 311–312)

²⁴ Eine eindeutige Unterscheidung muss nach (Schubert 2005: 311–312) dabei im Einzelfall erfolgen.

1.3.2 Bundesverwaltungsamt

Am 28. Dezember 1959 wurde mit dem „Gesetz über die Errichtung des Bundesverwaltungsamtes“ die Einrichtung des Bundesverwaltungsamtes (kurz BVA) im Bundestag beschlossen. Die allgemeinen Aufgaben des BVA wurden damals in §1 Absatz 2 und 3 beschrieben. Damit wurde das BVA dem Bundesministerium des Inneren, für Bau und Heimat (BMI) unterstellt und wird dadurch grundsätzlich dem öffentlichen Sektor zugeschrieben. Konkrete Aufgaben des BVA konnten zu Anfang wie folgt zusammengefasst werden:

- Beratung von Auswanderungs- bzw. Einwanderungswilligen inkl. Vorbereitung einer Aus- /bzw. Einwanderung, Beobachtung Aus- bzw. Einwanderbewegungen. (§2).
- Aufgaben als Bundesausgleichsstelle (§3).
- Führen des Ausländerzentralregisters (§6).
- Regelung von Fragen im Rahmen des Staatsangehörigkeitsgesetzes (§5).
- Weitere Leistungs-, Abrechnungs- und Versorgungsaufgaben (§4 und §7).

Seit März 2010 ist Christoph Verenkotte Präsident des BVA²⁵. Derzeit befindet sich der Hauptsitz des BVA in Köln. In den verschiedenen Außenstellen deutschlandweit werden rund 5.500 Mitarbeiter beschäftigt. Mit diesen werden mehr als 150 Fachaufgaben bearbeitet²⁶. Hierzu zählen unter anderem²⁷:

- Aufgaben im Bereich der Modernisierung und Digitalisierung der öffentlichen Verwaltung
- Aufgaben im Bereich der Öffentlichen Sicherheit.
- Beteiligung am Visaverfahren.
- Ausländerschulwesen.
- Vergabe von Bildungskrediten.
- Ausbildungsbehörden für den mittleren Dienst.
- Organisation für Ministerien und andere Behörden.
- Querschnittsaufgaben im Geschäftsbereich des BMI (z.B. Reisemanagement, Beihilfe, Bezügebearbeitung und Personalgewinnung).
- Seit 2017: Teilweise Übernahme von Aufgaben des Bundesamtes für zentrale Dienste und offene Vermögensfragen.

²⁵ Angabe aus dem vom BVA veröffentlichten Lebenslauf von Hr. Verenkotte, siehe dazu (Bundesverwaltungsamt o. J.).

²⁶ Angaben aus (Bundesverwaltungsamt 2018: 1).

²⁷ Angaben aus (Bundesverwaltungsamt 2018: 1–2).

- Kunstverwaltung und Münzwesen des Bundes.

Die technische Infrastruktur für Test- und Betriebsumgebung werden dem BVA durch das ITZ Bund, dem IT-Dienstleister des Bundes, zur Verfügung gestellt und betreut. Aus eigener Erfahrung bedienen sich die meisten Softwareverfahren sowie Projektgruppen dieser bestehenden IT-Strukturen.

1.3.3 Organisationseinheit

Im Verlauf dieser Masterthesis findet der Begriff Organisationseinheit mehrfach Verwendung. Konkret kann dieser Begriff je nach Anwendung des Fragebogens und der vorliegenden Organisationsstruktur definiert werden. Oftmals unterliegen Organisationen im öffentlichen Sektor festen Strukturen und Hierarchien an denen sich Ressourcen, Prozesse, Arbeitsplätze, Rollenverteilung etc. orientieren müssen. Selbige Einschränkungen gelten auch grundsätzlich für die Ausführung einer Befragung. Wenn man sich beispielsweise das Organigramm²⁸ des BVA ansieht, stellt die höchste Ebene die des Präsidenten des BVAs dar. Auf der anderen Seite ist die kleinste Ebene vor einem Mitarbeiter innerhalb dieser Grafik das Referat oder die Projektgruppe. Ein Referat bzw. Projektgruppe kann aus eigener Beobachtung wiederum aus mehreren Sachgebieten bzw. Teilprojekten bestehen. Eine Organisationseinheit im Sinne dieser Thesis wäre im Idealfall als eine Sachgebietseinheit oder eine Referatseinheit zu verstehen. Es wird dabei angenommen, dass dasselbe Verständnis über eine Organisationseinheit sich leicht auf vergleichbare Strukturen anwenden lässt.

1.4 Methodisches Vorgehen und Forschungsdesign

Im Fokus dieser Thesis steht die Konzipierung eines Instrumentes zur Messung des Reifegrades der Softwarequalitätssicherung. Um dieses Ziel zu erreichen wird auf Basis wissenschaftlicher Methoden gearbeitet. Die sieben Prinzipien des Design Science²⁹ stellen in weiten Teilen einen geeigneten Einsatz dar, da in ihrem Fokus die Entwicklung eines Artefakts steht. Mittels dieser Methode werden Artefakte erarbeitet, die zugleich eine unternehmerische Lösung sowie einen wissenschaftlichen Beitrag darstellen. Im Weiteren werden diese Artefakte evaluiert, verifiziert und zum Abschluss der Allgemeinheit zur Verfügung gestellt werden:

²⁸ Siehe (Bundesverwaltungsamt 2020).

²⁹ Siehe dazu (Hevner u. a. 2004) und (Simon 1996).

„Design science, as the other side of the IS research cycle, creates and evaluates IT artifacts intended to solve identified organizational problems.“³⁰

Ein Artefakt wird nach Hevner wie folgt definiert:

„It is broader in the sense that we include not only instantiations in our definition of the IT artifact but also the constructs, models, and methods applied in the development and use of information systems. However, it is narrower in the sense that we do not include people or elements of organizations in our definition nor do we explicitly include the process by which such artifacts evolve over time. We conceive of IT artifacts not as independent of people or the organizational and social contexts in which they are used but as interdependent and coequal with them in meeting business needs.“³¹

Diese Definition könnte nach Ansicht des Verfassers auch auf ein Reifegradmessverfahren zutreffen. Daher kann sich auch die Entwicklung eines solchen Verfahrens an diesen Prinzipien orientieren. Die sieben Prinzipien sind in Tabelle 1-1³² beschrieben.

Tabelle 1-1: Design-Science Research Guidelines

Guideline	Description
Guideline 1: Design as an Artifact	Design-science research must produce a viable artifact in the form of a construct, a model, a method, or an instantiation.
Guideline 2: Problem Relevance	The objective of design-science research is to develop technology-based solutions to important and relevant business problems.
Guideline 3: Design Evaluation	The utility, quality, and efficacy of a design artifact must be rigorously demonstrated via well-executed evaluation methods.
Guideline 4: Research Contributions	Effective design-science research must provide clear and verifiable contributions in the areas of the design artifact, design foundations, and/or design methodologies.
Guideline 5: Research Rigor	Design-science research relies upon the application of rigorous methods in both the construction and evaluation of the design artifact.
Guideline 6: Design as a Search Process	The search for an effective artifact requires utilizing available means to reach desired ends while satisfying laws in the problem environment.
Guideline 7: Communication of Research	Design-science research must be presented effectively both to technology-oriented

³⁰ (Hevner u. a. 2004: 77)

³¹ (Hevner u. a. 2004: 82–83)

³² Tabelle übernommen von (Hevner u. a. 2004: 83).

Aus Sicht des Verfassers sind die großen Stärken dieses erkenntnistheoretischen Ansatzes die Problemorientierung, die Artefakt-Gestaltung sowie die Feedbackloops mit der Community. Rückblickend wurde Prinzip 2 bereits in Kapitel 1 dieser Thesis berücksichtigt. Das Artefakt im Kontext dieser Thesis stellt ein Verfahren zur Reifegradmessung inkl. Modell und der hier entwickelten Werkzeuge dar. Zur Entwicklung dieses Artefakt findet das Mittel der Deduktion³³ auf Basis von anerkannten Dokumenten, Standards und Werken im Bereich der Softwarequalitätssicherung Anwendung. Ebenfalls wird sich für die Entwicklung im weiteren Verlauf eine Anlehnung am Werk von Kallus³⁴ erlaubt. Für diese Thesis muss ferner betrachtet werden, dass eine Subjektivität³⁵ des Verfassers nicht gänzlich ausgeschlossen werden kann. Reichertz beschreibt diesen Umstand wie folgt:

„Mit ‚Subjektivität‘ meine ich in diesem Kontext (und damit unterscheidet es sich von einem Subjektbegriff, der immer die Handlungsfähigkeit und die Handlungsmächtigkeit in den Blick nimmt) nicht die Einzigartigkeit der Identität des Forschers bzw. der Forscherin, dass sie also biografisch und kulturell an einem einzigartigen Ort stehen, sondern meint die jedem Menschen aufgrund seiner Lebensgeschichte je eigene psychische, emotionale und kognitive Verfasstheit, die jedes Handeln deutlich beeinflusst und damit auch als Ausdruck der Subjektivität erkennbar werden lässt. Subjektivität ist so etwas wie die eigene Handschrift, die sich lebensgeschichtlich als körperliche Praktik herausgebildet hat und die, sobald man schreibt (unabhängig vom Inhalt des Geschriebenen), sichtbar wird.“³⁶

2 Stand der Forschung und Wissensbasis

In diesem Kapitel werden die Grundlagen zu den Themen Softwarequalitätssicherung und zugehörige Reifegradmessungs- bzw. Bewertungsverfahren zusammengefasst. Da es sich hier im Einzelnen um sehr komplexe Themengebiete handelt, kann nur eine grobe Zusammenfassung erfolgen. Die Tatsache, dass Softwarequalitätssicherung sich als einen Teil der Kette von Softwareentwicklungsverfahren versteht und das Wissen entsprechender Verfahren bzw. Modelle wie zum Beispiel SCRUM, Wasserfallmodell oder V-Modell XT werden dem Leser als bereits bekannt vorausgesetzt und im Rahmen dieser Masterthesis nicht weiter erläutert.

³³ Siehe dazu (Reichertz 2016: 129–130).

³⁴ (Kallus 2016)

³⁵ Siehe dazu (Reichertz 2016: 78–82).

³⁶ (Reichertz 2016: 79)

2.1 Softwarequalitätssicherung

Jedes Softwareprojekt oder jede Entwicklung eines Softwareproduktes muss sich früher oder später mit den drei Zielkonflikten (Kosten, Leistung und Zeit), dem sogenannte „Magischen Dreieck“, auseinandersetzen. Alternativ wird dabei im sogenannten „Teufelsquadrat“ Leistung in Quantität aka Funktionsumfang und Qualität aufgeteilt³⁷. Daraus ergeben sich vier Zielgrößen, nach denen sich Softwareprojekte orientieren können, siehe dazu Abbildung 2-1³⁸.

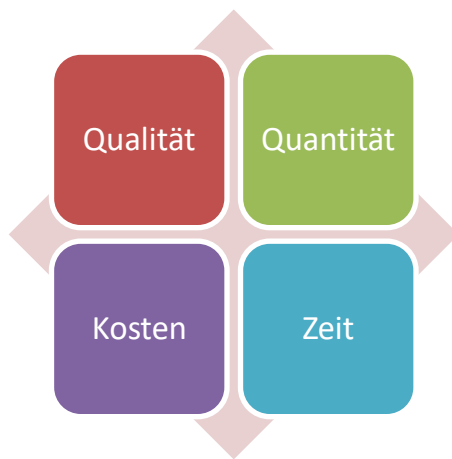


Abbildung 2-1: Teufelsquadrat

Es ist das grundsätzliche Ziel, diese einzelnen Bereiche soweit wie möglich im Gleichgewicht zu halten. In einem Artikel in der OBJEKTspektrum zum Thema Qualitätssicherung und dem „International Software Testing Qualifications Board“ (Kurz: ISTQB) stellen die Autoren ebenfalls den Zusammenhang zwischen Qualitätssicherung und Wirtschaftlichkeit fest:

„Doch nicht nur die Planung, der Entwurf und die Durchführung solcher Projekte stellen an die involvierten Personen besondere Anforderungen, auch die fortwährende Qualitätssicherung (QS) muss entsprechend skalierende Methoden und Verfahren anwenden,

³⁷Vgl. (Sneed 1987: 42 und 171).

³⁸ Erläutert im Werk von (Sneed 1987).

um eine adäquate Gesamtqualität bei gleichzeitiger Wirtschaftlichkeit sicherstellen zu können.³⁹.

Im Fokus von Softwarequalitätssicherungsmaßnahmen stehen also Inhalt, Methodik und Wirtschaftlichkeit. Der Klassiker „The Art of Testing“ definiert die zehn Prinzipien des Testens⁴⁰. Diese dienen in der Branche der Softwarequalitätssicherung als Grundpfeiler. Die aktuelle ISO Norm zum Thema Softwarequalitätssicherung stellt die ISO 29119 dar:

„The purpose of the ISO/IEC/IEEE 29119 series of software testing standards is to define an internationally-agreed set of standards for software testing that can be used by any organization when performing any form of software testing.“⁴¹.

Weitere Normen sind unter anderem IEEE 829-2008, ISO 9126, sowie die ISO 25000. Eine Säule des Themas Softwarequalitätssicherung stellt das „International Software Testing Qualifications Board“ (kurz: ISTQB) dar. Das ISTQB ist ein weltweiter Zusammenschluss verschiedener nationaler Testing Board Mitglieder⁴². Das „German Testing Board“ (kurz: GTB) zählt ebenfalls dazu⁴³. Im Folgenden ein Auszug über die eigens gesetzte „Mission“ des ISTQB:

- „1. We promote the value of software testing as a profession to individuals and organizations.*
- 2. We help software testers to be more efficient and effective in their work, through the certification of competencies.*
- 3. We enable testers to progress their career through a Professionals' Code of Ethics and a multi-level certification pathway that provides them with the skills and knowledge they need to fulfil their growing responsibilities and to achieve increased professionalism.*
- 4. We continually advance the Testing Body of Knowledge by drawing on the best available industry practices and the most innovative research, and we make this knowledge freely available to all.*
- ...*
- 7. We are committed to expanding software testing certifications around the world, by admitting member boards into the ISTQB®. These boards adhere to the constitution, bylaws, and processes defined by the ISTQB®, and participate in regular audits.*

³⁹ (Metzger u. a. 2013: 58)

⁴⁰ Siehe auch (Myers u. a. 2012: 12–18).

⁴¹ (ISO 29119-1 2013: vi)

⁴² Siehe dazu die Webseite des ISTQB (ISTQB 2019).

⁴³ Siehe dazu die Webseite des GTB (German Testing Board e.V. 2019).

8. *We nurture an open international community, committed to sharing knowledge, ideas, and innovations in software testing.*⁴⁴

Aus einer Wirksamkeitsumfrage mit weltweit mehr als 2000 Teilnehmern des ISTQB von 2016-2017 geht folgendes über die Anerkennung der Zertifizierung des ISTQB hervor:

*„More than half of the managers who answered the survey are considering 80% or more of their staff obtaining ISTQB® CTFL certification. There is a small decline (58% vs 67%) in managers still strongly preferring their team to reach high levels of certification coverage.*⁴⁵

Es liegt der Schluss nahe, dass eine Vielzahl von Unternehmen auf ISTQB setzen, dessen Zertifizierungen und Dokumente für den Bereich Softwarequalitätssicherung stehen. Zertifizierungen des ISTQB werden von vielen Unternehmen und Universitäten in Deutschland, u.a. an der RFH-Köln, ermöglicht⁴⁶. ISTQB-Werke und die verschiedenen ISO-Normen (u.a. ISO 9126, ISO 251010, ISO 29119) werden als Standardbezugspunkt für eine Vielzahl von Werken zur Softwarequalitätssicherung genutzt. ISTQB bezieht sich meist ebenfalls auf die genannten ISO-Normen zur Softwarequalitätssicherung. Die Syllabus des ISTQB bieten zum Thema Softwarequalitätssicherung, deren Begrifflichkeiten sowie Vorgehensweisen und Best-Practices eine in der Branche anerkannte grundsätzliche Übersicht und Informationsquelle an⁴⁷. Im Rahmen dieser Masterthesis wird sich ebenfalls auf die Werke des ISTQB bezogen.

2.2 Testprozess Reifegradmessungs- bzw. Bewertungsverfahren

In diesem Kapitel wird auf einen Auszug der bestehenden Testprozessverbesserungsverfahren eingegangen. In Abbildung 2-2⁴⁸ ist zu erkennen, dass seit 1988 eine Vielzahl von Verbesserungsverfahren bzw. Reifegradverfahren entwickelt wurden. Es ist allerdings im Rahmen dieser Masterthesis auf Grund der Menge nicht zielführend, alle Verfahren mit einer Beschreibung zu würdigen. Es beschäftigten sich bereits in der Vergangenheit verschiedene Literaturreviews⁴⁹ und Vergleiche hinreichend mit vielen dieser

⁴⁴ (ISTQB 2020)

⁴⁵ (Amaral u. a. 2017: 17)

⁴⁶ Weiteres dazu im Jahresbericht 2018 (German Testing Board e.V. 2018).

⁴⁷ Siehe dazu (ISTQB u. a. 2012) und (ISTQB u. a. 2018).

⁴⁸ Abbildung übernommen von (Garousi u. a. 2018: 89).

⁴⁹ Siehe dazu (Afzal u. a. 2016), (Garousi u. a. 2017) oder (Ron Swinkels 2000).

Verfahren. Daher wird sich auf die beiden Verfahren konzentriert, die bereits in der zuvor erwähnten Befragung als relevant galten:

„Zur Verbesserung des Testprozesses werden keine systematischen Audits durchgeführt (77 %). Sehr zurückhaltend werden genutzt (Mehrfachauswahl): TPI®/TPI-Next® (12 %), CMMI® (12 %), Automotive SPICE® (6 %) und TMMi® (3 %) sowie andere Referenzmodelle (18 %).“⁵⁰

Konkret sind das TPI/TPI-Next, sowie CMMI. Da TMMi als Ergänzung mit mehr Fokus auf Softwaretests zur CMMI gilt⁵¹, werden CMMI und TMMi gemeinsam betrachtet.

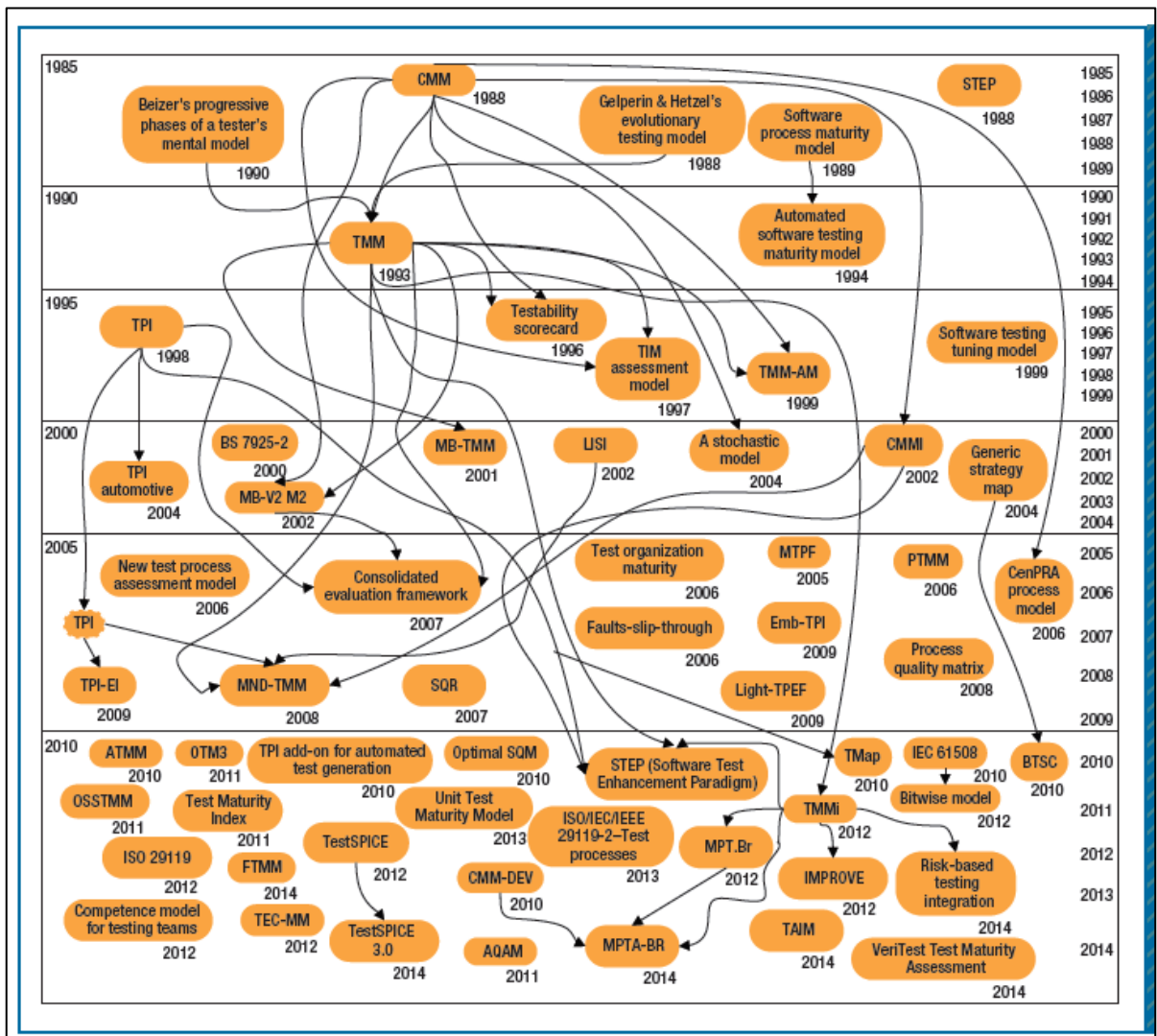


Abbildung 2-2: The evolution of TMA and TPI models and their relationships

⁵⁰ (Spillner u. a. 2017b: 16)

⁵¹ Vgl. (van Veenendaal 2018: 6).

2.2.1 TPI / TPI-Next

TPI bzw. TPI-Next steht für „Business Driven Test Process Improvement“⁵² und wurde von der Firma Sogeti entwickelt⁵³. TPI-Next baut auf dem ursprünglichen TPI auf, dieses wurde in den 90er Jahren entwickelt und 1998 veröffentlicht⁵⁴. TPI-Next unterscheidet für den Testprozess 16 sogenannte Kernbereiche⁵⁵. Diese Kernbereiche sind in drei Kategorien zugeordnet⁵⁶: Stakeholder-Beziehung, Testmanagement und Testkompetenz. Jeder Kernbereich kann in einen der vier Reifegrade bewertet werden: „

- **Initial**
Der Testprozess ist unstrukturiert und besteht aus Ad-hoc-Aktivitäten
- **Kontrolliert**
Alle Aktivitäten des Testprozesses werden geplant und durchgeführt. Der Testprozess ermöglicht ausreichende Erkenntnisse über die Qualität des getesteten Objektes
- **Effizient**
Die Aktivitäten des Testprozesses werden so aufeinander abgestimmt, dass es zu einem günstigeren Kosten-Nutzen-Verhältnis kommt.
- **Optimierend**
Der erreichte Zustand des Testprozesses wird ständig gepflegt, analysiert und an sich ändernde Erfordernisse angepasst.⁵⁷

Darüber hinaus sind für jeden Kernbereich sogenannte Kontrollpunkte bzw. Cluster⁵⁸ und Verbesserungsvorschläge definiert:

„Können alle jeweiligen Kontrollpunkte positiv beantwortet werden, dann gilt der Reifegrad des Kernbereiches als erreicht. Neben den Kontrollpunkten und Enablern sind jeweils konkrete Verbesserungsvorschläge aufgeführt, die eine zusätzliche Unterstützung zum Erreichen des Reifegrades anbieten.“⁵⁹

Ergänzend zur Nutzung der Kontrollpunkte folgende Beschreibung:

⁵² Siehe dazu (Spillner u. a. 2014: 368).

⁵³ Vgl. (Sogeti Deutschland GmbH 2020) und (sogeti.com 2020).

⁵⁴ Siehe dazu Kapitel Preface (Cap Gemini Sogeti (Firm) 2009).

⁵⁵ Siehe dazu (Spillner u. a. 2014: 368–373) und (Otter u. a. 2014b)

⁵⁶ Vgl. (Spillner u. a. 2014: 372–373).

⁵⁷ (Spillner u. a. 2014: 368–369)

⁵⁸ Siehe dazu (Otter u. a. 2014a)

⁵⁹ (Spillner u. a. 2014: 372)

„The requirements for each level are defined in the form of Checkpoints: questions that need to be answered positively in order to classify for that level. Based on the checkpoints a test process can be assessed, and for each key area the proper level can be established. As each next level of a key area is considered an improvement, this means that the checkpoints are cumulative: in order to classify for level B the test process needs to answer positively to the checkpoints both of level B and of level A.“⁶⁰

Dazu weiter im Detail eine kurze Beschreibung und die TPI-Bewertungsmatrix in Abbildung 2-3⁶¹:

„The main purpose of the matrix is to show the strong and weak sides of the current test process and to support prioritising actions for improvement. A filled in matrix offers all participants a clear view of the current situation of the test process. Furthermore, the matrix helps in defining and selecting proposals for improvement.“⁶²

#	Key area	H	N	L	Initial	Controlled			Efficient			Optimizing				
01	Stakeholder commitment	x				A	B	B	C	F	H	H	K	M	M	
02	Degree of involvement	x				A	B	C	E	H	H	J	L	L	L	
03	Test strategy	x				A	A	B	E	F	F	H	K	L	L	
04	Test organization	x				A	D	D	E	I	I	J	J	K	L	L
05	Communication	x				B	C	C	D	F	F	J	M	M	M	
06	Reporting	x				A	C	C	C	F	G	G	K	K	K	
07	Test process management	x				A	A	B	B	G	H	J	J	K	M	M
08	Estimating and planning	x				B	B	C	C	G	H	I	I	K	L	L
09	Metrics	x				C	C	C	D	G	H	H	I	K	L	L
10	Defect management	x				A	A	B	D	F	F	H	J	K	L	L
11	Testware management	x				B	B	D	E	I	I	J	J	L	L	L
12	Methodology practice	x				C	D	D	E	F	H	J	J	M	M	M
13	Tester professionalism	x				D	D	E	E	G	G	I	I	K	K	M
14	Test case design	x				A	A	A	E	F	I	I	J	K	K	M
15	Test tools	x				E	E	E	E	F	G	G	I	L	M	M
16	Test environment	x				C	D	D	E	G	H	J	J	L	M	M

Abbildung 2-3: TPI-Bewertungsmatrix

Ein paar Beispiele aus den 157 Assessment Kontrollpunkten des TPI lauten wie folgt:

„The principal stakeholder is defined (not necessarily documented) and known to the testers.“⁶³

„Reporting is done in writing.“⁶⁴

„For further handling of defects the responsibilities are defined.“⁶⁵

⁶⁰ (Koomen u. a. 1998: 9)

⁶¹ Übernommen aus dem TPI-Tool (Otter u. a. 2014b).

⁶² (Koomen u. a. 1998: 10)

⁶³ Sequence 01.c.1 (Otter u. a. 2014b)

⁶⁴ Sequence 06.c.3 (Otter u. a. 2014b)

⁶⁵ Sequence 10.c.3 (Otter u. a. 2014b)

Jeder Kontrollpunkt muss mit „Ja“, „Nein“ oder „Nicht Verfügbar“ beantwortet werden. Aus der Kombination der Antworten leitet sich dann der Reifegrad der Kernbereiche ab. Weitere Details finden sich in:

- TPI® NEXT - Business Driven Test Process Improvement⁶⁶
- TPI® NEXT - Geschäftsbasierte Verbesserung des Testprozesses⁶⁷.

2.2.2 CMMI / TMMi

CMMI steht für „Capability Maturity Model Integration“⁶⁸ welches vom Software Engineering Institute (SEI) der Carnegie Mellon University entwickelt wurde und befindet sich aktuell in der Version 1.3.

„CMMI-Modelle (Capability Maturity Model Integration) sind eine Sammlung guter Praktiken, die Organisationen helfen, ihre Prozesse zu verbessern.“⁶⁹

CMMI ist ein Nachfolger des CMM Verfahrens. Der Entwicklungsverlauf von CMMI wird in Abbildung 2-4⁷⁰ dargestellt.

⁶⁶ Siehe dazu (Cap Gemini Sogeti (Firm) 2009).

⁶⁷ Siehe dazu (Ewijk u. a. 2011).

⁶⁸ Vgl. (Carnegie Mellon University 2011).

⁶⁹ (Carnegie Mellon University 2011: 4)

⁷⁰ Abbildung übernommen von (Carnegie Mellon University 2011: 18).

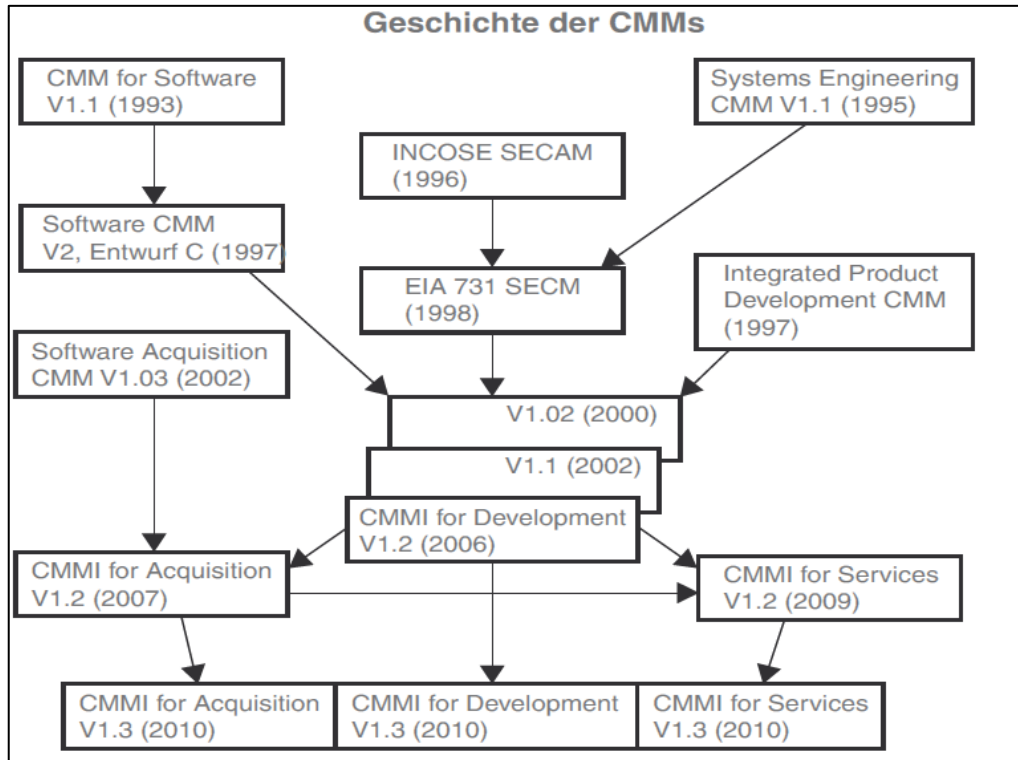


Abbildung 2-4: CMM/CMMI Entwicklung

CMMI Development oder CMMI Entwicklung (im weiteren CMMI-Dev genannt) beinhaltet u. a. die Softwareentwicklung⁷¹ und bietet zwei Ansätze zur Verbesserung an⁷²: Die Fähigkeitsgrade und die Reifegrade:

„Es handelt sich dabei um die »Darstellung in Fähigkeitsgraden« und die »Darstellung in Reifegraden«. Mithilfe der Darstellung in Fähigkeitsgraden können Sie bestimmte Fähigkeitsgrade erreichen, mit der Darstellung in Reifegraden bestimmte Reifegrade.

Um einen bestimmten Grad zu erreichen, muss eine Organisation alle Ziele der für eine Verbesserung vorgesehenen Prozessgebiete erfüllen. Dies gilt unabhängig davon, ob es sich um einen Fähigkeits- oder einen Reifegrad handelt.“⁷³

⁷¹ Siehe dazu (Carnegie Mellon University 2011: 19–20).

⁷² Weitere Details in (Carnegie Mellon University 2011: 33–34).

⁷³ (Carnegie Mellon University 2011: 33–34)

CMMI-Dev unterscheidet für seine Reifegrade in fünf Abstufungen⁷⁴: 1 - Initial, 2 - Geführt, 3 - Definiert, 4 - Quantitativ geführt und 5 - Prozessoptimierung. Die Fähigkeitsgrade werden in vier Abstufungen⁷⁵ (0-3) unterschieden: 0 - Unvollständig, 1 - Durchgeführt, 2 - Geführt und 3 - Definiert. Allerdings schränkt Spillner dahingehend ein, dass CMMI-Dev den Bereich Softwarequalitätssicherung nur in einem begrenzten Rahmen abdeckt:

„Der Testmanager findet keine sehr detaillierten Beschreibungen oder Hilfestellungen für seine Aufgaben, die über seine Kenntnisse des fundamentalen Testprozesses (s. [Spillner 12, Abschnitt 2.2] und Kap. 2) hinausgehen.“⁷⁶

TMMi steht für „Test Maturity Model integration“⁷⁷ und entstand aus dieser Abgrenzung heraus:

"The CMMI is often regarded as the industry standard for software process improvement. Despite the fact that testing often accounts for at least 30-40% of the total project costs, only limited attention is given to testing in the various software process improvement models such as the CMMI. As an answer, the testing community has created its own improvement models.“⁷⁸

Aktuell befindet TMMi sich in der Version 1.2, diese ist als komplementär zur CMMI Version 1.3 angesetzt⁷⁹. Aus diesem Grund sind ebenfalls die Struktur der fünf Reifegradstufen⁸⁰ des TMMi vergleichbar zu denen des CMMI⁸¹, siehe dazu Abbildung 2-5⁸².

⁷⁴ Siehe dazu (Carnegie Mellon University 2011: 38–41).

⁷⁵ Siehe dazu (Carnegie Mellon University 2011: 36–37).

⁷⁶ (Spillner u. a. 2014: 356)

⁷⁷ Siehe dazu (van Veenendaal 2018).

⁷⁸ (van Veenendaal 2018: 6)

⁷⁹ Siehe dazu (van Veenendaal 2018: 6).

⁸⁰ Siehe im Detail zu den Reifegradstufen (van Veenendaal 2018: 9–12).

⁸¹ Vgl. (van Veenendaal 2018: 13) und (Spillner u. a. 2014: 366–367).

⁸² Abbildung übernommen von (van Veenendaal 2018: 9).

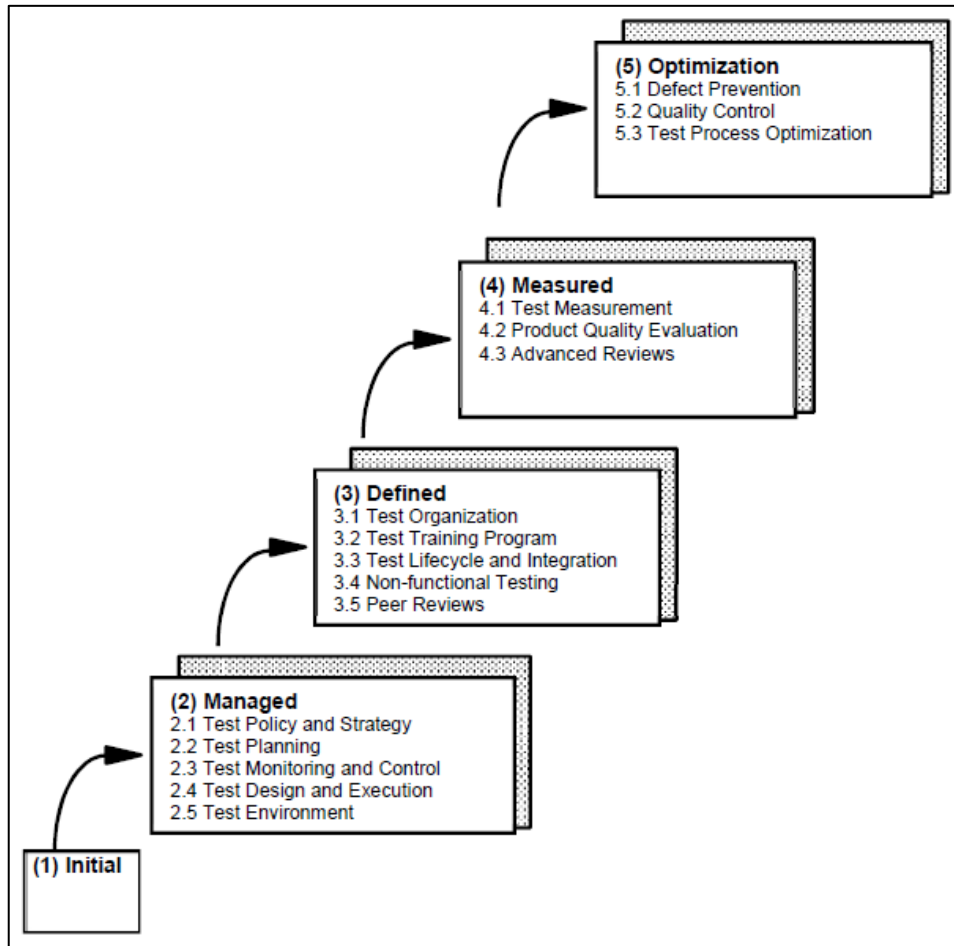


Abbildung 2-5: CMMI/TMMi Reifegradstufen

„As stated with the exception of level 1, each maturity level consists of several process areas that indicate where an organization should focus to improve its test process. Process areas identify the issues that must be addressed to achieve a maturity level. Each process area identifies a cluster of test related activities.“⁸³

TMMi unterscheidet also verschiedene Prozessbereiche welche wiederum konkrete Cluster beinhaltet. Die Prozessbereiche sind ebenfalls in Abbildung 2-5 erkennbar und ordnen sich immer einem Reifegrad unter. Im Kern werden von TMMi konkrete Ziele für jeden Prozessbereich vorgegeben. Hierbei wird zwischen spezifischen und generischen Zielen unterschieden⁸⁴. Dem gegenüber stehen spezifische und generische Praktiken und stellen den Weg dar, ein Ziel zu erreichen⁸⁵. Um einen Reifegrad zu erreichen, muss

⁸³ (van Veenendaal 2018: 13)

⁸⁴ Siehe dazu (van Veenendaal 2018: 14).

⁸⁵ Siehe dazu (van Veenendaal 2018: 15).

ein definiertes Set an Zielen erfüllt sein. Es folgt als Beispiel ein Auszug aus dem Prozessbereich „2.1 Test Policy and Strategy“ der zweiten Reifegradstufe. Dieser Prozessbereich teilt sich in drei spezifische Ziele auf⁸⁶:

- SG 1 - Establish a Test Policy
- SG 2 - Establish a Test Strategy
- SG 3 - Establish Test Performance Indicators

SG steht für „Specific Goal“ und SP für „Specific Practices“. Ein spezifisches Ziel besteht aus spezifischen Praktiken:

„SG 1 Establish a Test Policy

SP 1.1 Define test goals

SP 1.2 Define test policy

SP 1.3 Distribute the test policy to stakeholders“⁸⁷

SP 1.1 ist wie folgt definiert:

„Define and maintain test goals based upon business needs and objectives.“⁸⁸

Innerhalb von SP 1.1 sind konkrete „Sub-practices“ definiert:

„1. Study business needs and objectives

Examples of business needs and objectives to be studied include the following:

- *Mission statement*
- *Business and user needs regarding the products*
- *Business drivers*
- *Main goals of a quality program*
- *Business (quality) policy*
- *Type of business, e.g., risk level of products being developed*

2. Provide feedback for clarifying business needs and objectives as necessary

3. Define test goals traceable to business needs and objectives

Examples of test goals include the following:

- *Validate products for 'fit-for-use'*
- *Prevent defects from occurring in operation*

⁸⁶ Vgl. (van Veenendaal 2018: 24).

⁸⁷ (van Veenendaal 2018: 24)

⁸⁸ (van Veenendaal 2018: 25)

- *Verify compliance to external standards*
- *Provide visibility regarding product quality*
- *Shorten test execution lead-time*

4. Review the test goals with stakeholders

5. Revisit and revise the test goals as appropriate, e.g., on a yearly basis⁸⁹

⁸⁹ (van Veenendaal 2018: 25)

3 Entwicklung eines Verfahrens zur Reifegradmessung

Nachdem auf einen Auszug der bestehenden Verfahren eingegangen worden ist, bearbeitet dieses Kapitel die Entwicklung eines Verfahrens zur Reifegradmessung der bestehenden Qualitätssicherung innerhalb einer Organisationseinheit des öffentlichen Sektors. Dafür steht vor allem die konkrete Lösung der zuvor beschriebenen Problematik im Fokus, aber auch die Aufwände, die solch ein Verfahren zur Ausführung beansprucht. Die in Kapitel 1.4 genannten Prinzipien der Design Science 1, 4, 5 und 6 finden in diesem Kapitel Anwendung. Zu Beginn wird konkret die Zielsetzung und Art des anzustrebenden Verfahrens definiert. Im Anschluss findet für den Bereich des öffentlichen Sektors eine Eingrenzung möglicher Teilnehmerkreise statt, auf die das Verfahren Anwendung finden sollte. Die Grundlagen des Verfahrens werden das daraufhin konzipierte Reifegradmodell und die Verfahrensinhalte darstellen.

3.1 Ziel eines Verfahrens zur Reifegradmessung

Die Umsätze mit Software sind seit Jahren im ständigen Wachstum⁹⁰. Ebenso steigt die Anzahl von Computernutzern⁹¹. Ein Pressebericht der Bitkom aus dem Jahr 2018 kommt auf Basis des Bitkom-Branchenbarometers zu folgenden Erkenntnissen:

„Der wichtigste Trend für die Digitalunternehmen bleibt wie in den Vorjahren die IT-Sicherheit, die zwei Drittel der Befragten (67 Prozent) nennen. „Immer mehr Geräte sind digital vernetzt, das macht sie für Cyberkriminelle zu interessanten Angriffszielen. IT-Sicherheit gewinnt daher an Bedeutung, vor allem in Bereichen wie Energie, Verkehr oder auch Medizin wird sie herausragend wichtig“, so Berg. Auf dem zweiten Platz rangiert Cloud Computing (61 Prozent), dahinter folgen mit deutlichem Abstand das Internet der Dinge (IoT, 48 Prozent), Industrie 4.0 (47 Prozent) sowie Big Data (43 Prozent). Das Thema Digitale Plattformen, das im vergangenen Jahr erstmals im Ranking aufgeführt wurde, bleibt auf dem sechsten Platz (33 Prozent).“⁹²

Das lässt den Schluss zu, dass die Erwartungen an Software, Funktionsumfang und ihre Zuverlässigkeit sich ebenfalls im Wachstum befinden. Es ist eine logische Schlussfolgerung, dass infolgedessen auch die Ansprüche an Softwarequalitätssicherung steigen. Mit Blick auf das in Kapitel 1 beschriebene Praxisproblem könnte dies gerade für den öffentlichen Sektor eine große Herausforderung darstellen. Konkret weist der Bereich der Qualitätssicherung mutmaßlich starke Schwächen auf. Im Fokus des hier entwickel-

⁹⁰ Vgl. (Bitkom e.V. und EITO 2020).

⁹¹ Vgl. (Statistisches Bundesamt 2018).

⁹² (Bitkom e.V. 2018)

ten Verfahrens steht eine Reifegradmessung solcher Qualitätssicherung. Ziel einer Reifegradmessung ist im Allgemeinen eine Verbesserung⁹³ der bestehenden Situation zu erreichen. Ziele für eine Verbesserung können laut von Veenendaal⁹⁴ folgende sein:

„As stated earlier, the context within which test process improvement takes place includes any business or organizational process improvement and any IT or software process improvement. The following list includes some typical reasons for business improvements that influence testing:

- Improve product quality-Reduce time to market but maintain quality levels*
- Save money, improve efficiency*
- Improve predictability*
- Meet customer requirements*
- Be at a capability level; e.g., for outsourcing companies*
- Ensure compliance to standards; e.g., FDA in the medical domain (see section 2.5.4)⁹⁵*

Wie zuvor erläutert existieren einige Testprozessverbesserungsmethoden. Zu Beginn muss also eine geeignete Methode zur Reifegradmessung ausgewählt werden. Aus dem Literaturreview von Garousi et al.⁹⁶ geht hervor, dass die wenigen relevanten Methoden zur Durchführung und Analyse der Ergebnisse zeit- und ressourcenaufwändig sind. Dies wäre aber gerade in kleinen Organisationseinheiten angesichts des gegenübergestellten direkten Nutzens oftmals nicht vertretbar. Ebenfalls werden die Nachvollziehbarkeit, die Verhältnismäßigkeit und der Änderungswiderstand der Beteiligten sowie weitere Punkte als Herausforderung beschrieben⁹⁷. Daher zielt das hier konzipierte Verfahren auf einen vereinfachten Rahmen ab, welcher auch kurzfristig sowie transparent zur Ausführung gebracht werden kann. Eine Durchführung soll für kleine und große Organisationseinheiten möglich sein. Ziel ist es, den Bedarf an Ressourcen so gering wie möglich zu halten und auf der anderen Seite verwertbare Aussagen über den Reifegrad der Qualitätssicherung zu erhalten. Im weiteren Verlauf könnten diese Ergebnisse als erste Anhaltspunkte für Maßnahmen der Verbesserung genutzt werden. Das Verfahren soll Managern des öffentlichen Sektors im Bereich der Softwareentwicklung helfen eine initiale Reifegradmessung durchzuführen. ISTQB definiert den Aufgabenbereich eines solchen Managers wie folgt:

⁹³ Siehe auch Ziele der Verbesserung (Spillner u. a. 2014: 350).

⁹⁴ Erik Van Vennendaal ist an der Verfahrensentwicklung von TMap sowie TMMi beteiligt.
Siehe dazu (van Veenendaal 2020).

⁹⁵ (Bath und Veenendaal 2014: 15)

⁹⁶ Siehe (Garousi u. a. 2017).

⁹⁷ Vgl. (Garousi u. a. 2017: 23–24).

„Eine zentrale Verantwortung eines Managers besteht darin, die Schaffung und Nutzung von benötigten Ressourcen (Personen, Software, Hardware, Infrastruktur usw.) sicherzustellen, um für Mehrwert schaffende Prozesse zu sorgen. Für Manager in der Software- und IT-Industrie sind diese Prozesse oft Teil eines Projekts oder Programms, das eine Software oder ein System für interne oder externe Zwecke liefern soll.“⁹⁸

Diese Manager könnten ein Qualitätssicherungsbeauftragter, ein Sachgebietsleiter, ein Referatsleiter oder vergleichbares sein. Das hier entwickelte Verfahren könnte und sollte dabei auch als Vorstufe einer umfassenderen Reifegradmessung oder QS-Analyse verstanden werden. Es wäre denkbar, dass die Ergebnisse zu der Erkenntnis führen, umfangreichere Reifegradanalysen durchzuführen. Es stellt also keinen Ersatz oder eine Konkurrenz zu den Bestandsverfahren (wie z.B. TPI-Next oder TMMi) dar. Durch die Darlegungen dieses Kapitels wurde Prinzip 4 der Design-Science abgedeckt.

3.2 Bestimmung der Verfahrensart

In Einklang mit Design-Science Prinzip 1, 5 und 6 wird in diesem und im nächsten Kapitel die Grundlage für das Verfahren gelegt. Das angestrebte Verfahren verfolgt den Zweck einer standardisierten initialen Reifegradmessung der Softwarequalitätssicherung, welches im Rahmen einer Organisationseinheit des öffentlichen Sektors durchgeführt werden kann. Ein Reifegrad wird wie folgt definiert:

„Der Reifegrad wird durch sogenannte >>Assessments<<⁴ ermittelt. Beim Assessment werden die tatsächlichen Abläufe mit den Anforderungen des Assessmentsmodell⁵ verglichen, Verbesserungspotenziale herausgearbeitet und eine entsprechende Einordnung in eine Reifegradstufe vorgenommen.“⁹⁹

Alternativ folgt eine Definition des CMMI:

„Reifegrade beziehen sich darauf, wie gut eine Organisation Prozessverbesserungen auf mehreren Prozessgebieten erreicht. Diese Grade dienen dazu, die Prozesse zu verbessern, die zu einer gegebenen Menge von Prozessgebieten (d.h. einem Reifegrad) gehören.“¹⁰⁰

⁹⁸ (ISTQB u. a. 2012: 20)

⁹⁹ (Spillner u. a. 2014: 354) Anm.: Fußnoten aus dem Original: ⁴ „Bei CMMI als Appraisal bezeichnet; ⁵ In der ISO/IEC 15504 (s. Abschnitt 14.2.1) ist im Teil 2 das Reifegradmodell und in Teil 5 das Assessmentmodell definiert. CMMI (s. Abschnitt 14.2.1) ist ein Assessmentmodell, das statt Teil 5 verwendet werden kann.“

¹⁰⁰ (Carnegie Mellon University 2011: 35)

Daraus lässt sich schließen, dass für eine Reifegradmessung im Sinne dieser Masterthesis ein Assessment notwendig ist. Auf Basis dieser Erkenntnisse erfolgt mit dem Ziel der Verbesserung eine Einordnung in einem Stufenmodell. Für das Assessment werden Informationen über die aktuellen Zustände bezüglich der Softwarequalitätssicherung innerhalb der Organisationseinheit benötigt. Dazu muss eine Datenbasis erhoben werden. Auf Basis dieser erfolgt die Bewertung (das Assessment). Zur Datenerhebung bieten sich folgende Verfahren an¹⁰¹: Beobachtung, Interview, Fragebogen oder Dokumentenanalyse. Damit auch Organisationseinheiten mit vergleichsweise geringer Ressourcenverfügbarkeit dieses Verfahren zur Ausführung bringen können, sollen Komplexität und erwartete Aufwände klein gehalten werden. Daher wird maximal ein Verfahren für die Datenerhebung in Betracht gezogen. Zur Auswahl eines Datenerhebungsverfahrens wird sich im weiteren Verlauf mit der Ergebnisform, der möglichen Datenbasis und der Klassifikation auseinandergesetzt. Die Ergebnisse der Reifegradbewertung dienen dazu, besser zu verstehen welche QS-Aktivitäten in welcher Intensität ausgeführt werden, um mit diesem Wissen letztendlich höhere Reifegrade zu erzielen. Die Kernidee hinter einem solchen Verfahren ist es, die tatsächlich ausgeführten Vorgehensweisen aufzudecken. Es liegt hier der Gedanke nahe dazu eine Prozessbeschreibungsdokumentation zu verwenden. Allerdings kann nicht davon ausgegangen werden, dass diese überhaupt existiert. Auch kann eine Abweichung zwischen Prozessbeschreibungsdokumenten und tatsächlich gelebten Prozessen nicht ausgeschlossen werden. Eine solche Abweichung kann aus sehr unterschiedlichen Gründen entstanden sein. Eine Dissertation über die Wirtschaftlichkeit von Wissenstransfers in der Softwareentwicklung musste ebenfalls diese Wirkung der sogenannten Scheinsicherheit für ihr Vorgehen berücksichtigen:

„Mit der Fokussierung auf materielle Wissensträger kann nach Bendt eine Scheinsicherheit entstehen die zu einer Vernachlässigung personaler Maßnahmen⁴³ beitragen kann. Im Zuge von Sicherheitsbestrebungen angeordnete Dokumentationsaufgaben können zudem kontraproduktiv wirken, wenn sie Demotivation und im Extremfall sogar eine Abwanderung von Mitarbeitern auslösen. Ein weiteres Problem stellt ihrer Ansicht nach die Veralterungsgefahr dar, da sich mit der Explizierung und Formalisierung der Aufwand und die Zeitdauer für die Aktualisierung von Wissen erhöhen können.“¹⁰²

¹⁰¹ Siehe dazu Kapitel 10 (Döring und Bortz 2016: 321–578).

¹⁰² (Trittmann 2004: 138). Anm.: Fußnote aus dem Original: ⁴³ „Dazu zählen beispielsweise Mentorenkonzepte wie „sempai-kohai“ aber auch Maßnahmen zur Reduzierung der Mitarbeiterfluktuation. Vgl. ausführlich Romhardt/ Wissensperspektive/ 252-279.“

Daher schließt sich die Dokumentenanalyse an dieser Stelle aus. Für ein ergänzendes Bild empfiehlt es sich dennoch die Dokumentenanalyse zusätzlich zu dem hier konzipierten Verfahren auszuführen. Ein solches Vorgehen ist aber nicht Teil dieser Masterthesis und sollte im Einzelfall bei Bedarf betrachtet werden.

CMMI/TMMi empfiehlt zur Datenerhebung, Analyse und in Folge auch um die Prozessverbesserung sicherzustellen, ein Team einzurichten, welches die relevanten Stakeholder repräsentiert¹⁰³. Und folgende Aussage zum TPI/TPI-Next deutet auf eine Form der Befragung und Dokumentenanalyse hin:

„Durch Befragung von Beteiligten und Beurteilung von Dokumenten werden die einzelnen Kernbereiche auf Grundlage der Kontrollpunkte bewertet. Wird festgestellt, welche Fragen positiv beantwortet werden können und welche nicht oder nur teilweise.“¹⁰⁴

Es wird daraus die Annahme aufgestellt, dass mindestens in jeder Organisationseinheit die Mitarbeiter selbst in der Lage sind, einzuschätzen, welche Vorgehensweisen und Prozesse gelebt werden. Als Datenbasis für dieses Verfahren wird sich an TPI/TPI-Next angelehnt und ebenfalls auf die Aussagen der Mitarbeiter gestützt, welche damit auch aktiv in das Verfahren eingebunden werden. Aus diesem Grund werden den Teilnehmern Hintergrund und Ziele des Verfahrens transparent dargelegt. Unterschiedliche Quellen empfehlen hier eine Einleitung bzw. Einladung¹⁰⁵ begleitend zum Verfahren einzusetzen. Nach wissenschaftlicher Auffassung kann sich für eine qualitative oder quantitative Vorgehensweise entschieden werden¹⁰⁶. Eine quantitative Form bietet sich hier an, da sie es ermöglicht, dass Verfahren skalierbar in Organisationseinheiten und Hierarchien unterschiedlicher Größen einzusetzen. Zusätzlich können die Ergebnisse durch

¹⁰³ Siehe dazu (Carnegie Mellon University 2011: 70–71).

¹⁰⁴ (Spillner u. a. 2014: 374–375)

¹⁰⁵ Siehe auch (Bedenk u. a. 2004: 5), (Pierenkemper 2016: 6), (Porst 2014: 36–37).

¹⁰⁶ Siehe dazu Kapitel 1.4.1 in (Hussy u. a. 2013: 20–26).

die Darstellung empirischer Sachverhalte vergleichbar gemacht werden¹⁰⁶. Daraus ableitend wird für dieses Verfahren ein quantitativ¹⁰⁷ reaktiver¹⁰⁸ Ansatz der Datenerhebung verfolgt. Es ist allerdings zu beachten, dass sich durch diese Vorgehensweise eine Verzerrung in den Daten wiederfinden kann:

„Die im Zuge der Datenerhebung generierten Daten unterliegen somit unterschiedlichen Verzerrungen (z. B. Interviewdaten zur Internetnutzung können durch sozial erwünschtes Antworten verzerrt sein)“¹⁰⁹

Im Rahmen dieses Verfahren soll den Teilnehmern daher die Anonymität zugesichert werden¹¹⁰. Eine Einleitung und die Anonymität der Teilnehmer streben in erster Linie eine Mitigation der Datenverzerrung an, aber zusätzlich soll einer möglichen Sorge vor „geheimen Arbeitskontrollen“ entgegengewirkt werden. Ebenfalls wird sich ein Nebeneffekt erhofft, der das Bewusstsein der Teilnehmer als solches zum Thema Software-QS über den Verlauf des Verfahrens hinaus schärft. Die Erhebungsmethode eines Interviews oder einer Beobachtung würde je nach Größe der Organisationseinheit einen deutlichen Mehraufwand in der Ausführung, Dokumentation und Auswertung bedeuten. Es besteht die Möglichkeit, dass die Teilnehmer an der zuvor versprochenen Anonymität zweifeln und somit ggf. ehrliche Antworten im Interview vermeiden¹¹¹. Auch kann nicht angenommen werden, dass jede Organisationseinheit sich räumlich an einem Platz oder innerhalb eines Gebäudes befindet. Gerade in Bundesbehörden kann es durchaus vorkommen, wie es nach eigener Beobachtung auch im BVA der Fall ist, dass Mitarbeiter einer Organisationseinheit sich über verschiedene Standorte in Deutschland verteilen. Dies würde im Falle von Beobachtungen oder Interviews zu einem organisatorischen Mehraufwand des Verfahrens führen¹¹¹. Weiterhin muss davon ausgegangen werden, dass technische Mittel für ein mögliches Telefon- oder Onlineinterview nur ungleichmäßig in den unterschiedlichen Behörden zur Verfügung stehen. Die ganze Situation würde sich nochmals verschärfen je mehr Abteilungen und oder Organisationseinheiten an

¹⁰⁷ Definition quantitativ: *„Bei quantitativen Datenerhebungsmethoden ist dagegen der Strukturierungsgrad sehr hoch und es wird mit vollstandardisierten Instrumenten gearbeitet. Im vollstrukturierten quantitativen Interview sind z. B. nicht nur Anzahl, Reihenfolge und Wortlaut aller Fragen exakt vorgegeben, sondern auch die Antwortvorgaben ausformuliert, so dass die Befragten nur die auf sie am besten zutreffenden Antworten auszuwählen haben.“* (Döring und Bortz 2016: 322–323).

¹⁰⁸ Definition reaktiv: *„Bei reaktiven Datenerhebungsmethoden wissen die untersuchten Personen, dass sie an einer Studie teilnehmen.“* (Döring und Bortz 2016: 323).

¹⁰⁹ (Döring und Bortz 2016: 323)

¹¹⁰ Siehe auch (Hussy u. a. 2013: 59–60).

¹¹¹ Siehe dazu (Döring und Bortz 2016: 357).

dem Verfahren beteiligt werden. Im Schluss bedeutet dies, dass als Datenerhebungsmethode die Form eines Fragebogens genutzt wird. Folgende Eigenschaften dieser Methode unterstützen diesen Schluss¹¹²:

- Ein Fragebogen kann den Teilnehmer in unterschiedlichen Situationen ohne zwingende Anwesenheit des Forschenden zur Verfügung gestellt werden.
- Es können Aspekte des Erlebens, sowie des vergangenen oder privaten Verhaltens erfasst werden, die nicht direkt beobachtbar und auch nicht in Dokumenten beschrieben sind.
- Zeitaufwand und Schulungsaufwand sind geringer
- Die diskrete und anonyme Situation führt zu einer höheren Bereitschaft der Teilnahme.
- Der Fragebogen ist für quantitative Befragungen geeignet und bevorzugt¹¹³.

Diese Auswahl bringt auf der anderen Seite auch negative Eigenschaften mit die für den weiteren Verlauf zu berücksichtigten sind:

- Eine quantitative Befragung schränkt die Antwortmöglichkeiten und damit auch einen möglichen Informationsgewinn ein.
- Die zuvor schon angesprochene mögliche Verzerrung von Antworten.
- Aufgrund der fehlenden Kontrolle über die Bedingungen des Verfahrens können externe Einflüsse nicht ausgeschlossen werden (z.B. Tagesstimmung in der Organisation, Beeinflussung durch andere Personen, usw.)¹¹⁴. Auch vergleichbar mit einer Datenerhebung im Internet¹¹⁵.

Konkret handelt es sich somit um eine Befragung¹¹⁶, spezifisch um eine Mitarbeiterbefragung mittels Fragebogen. Um über die vorgegebenen Antworten für die quantitative Bewertung hinaus Einblicke in die Vorgehensweisen und Sichtweisen der Teilnehmer zu erhalten, werden in den Fragebogen halb-offene Fragen integriert. Auf diese Weise lassen sich Aspekte der qualitativen Form einbauen.

¹¹² Erläutert in (Döring und Bortz 2016: 398–399).

¹¹³ Siehe auch (Hussy u. a. 2013: 223).

¹¹⁴ Siehe dazu „Fehler durch die Testperson“ in (Döring und Bortz 2016: 436) und „Objektivität“ in (Hussy u. a. 2013: 23–24).

¹¹⁵ Siehe Abschnitt 2.6 (Hussy u. a. 2013: 104–112).

¹¹⁶ Weiteres dazu in (Hussy u. a. 2013: 71–77).

„Die Offenheit des Vorgehens soll eine optimale Annäherung an den Untersuchungsgegenstand gewährleisten und die Chance erhöhen, dass sich auch neue und unerwartete Inhalte in den erhobenen Daten zeigen.“¹¹⁷

Das Verfahren im Kern bleibt daher quantitativ und kann auch nicht die inhaltlichen Ergebnisse und Tiefe eines qualitativen Verfahrens erreichen. Vielmehr soll dieser offene Ansatz die Reifegradbewertung nicht beeinflussen. Dies kann aber in Kombination mit diesem die anschließende Analyse für das weitere Vorgehen positiv beeinflussen und so einen gewissen Interpretationsspielraum ermöglichen. Durch die halbstrukturierte Form lässt sich das konzeptionierende Verfahren den Mixed Methods¹¹⁸ zuordnen. Dies wird auch durch folgende Aussage gestützt:

„Denn im Zuge einer quantitativen Studie und im Verständnis des Kritischen Rationalismus lassen sich qualitative Datenerhebungen ohne weiteres integrieren – typischerweise im Entdeckungszusammenhang zur Generierung von neuen Theorien, die anschließend im Begründungszusammenhang quantitativ-statistisch geprüft werden“¹¹⁹

Das bedeutet konkret für den Fragebogen, dass die Antworten weitgehend standardisiert sein werden, aber dort, wo zielführend, auch Freitextfelder als Antwortoption dem Teilnehmer zur Verfügung gestellt werden. Diese Felder dürfen vom Teilnehmer im Kontext der Frage frei ausgefüllt werden. Im Anschluss der Datenerhebung folgt das Assessment bzw. die Bewertung der Datenbasis bezüglich einer möglichen Einordnung in ein Reifegradmodell. Das Reifegradmodell wird im in Kapitel 3.4 dieser Masterthesis entworfen.

3.3 Population und Stichprobe einer Befragung

Das hier entwickelte Verfahren kann grundsätzlich von jeder Organisation des öffentlichen Sektors durchgeführt werden. Jede dieser Organisationen könnte theoretisch das Verfahren in einem für sich spezifischen Rahmen mit einer spezifischen Zielpopulation und Stichprobe ausführen. Eine Zielpopulation¹²⁰ bzw. Grundgesamtheit kann sich also nur im Kontext der Organisationseinheiten befinden. In Folge dessen bewegt sich eine

¹¹⁷ (Döring und Bortz 2016: 322)

¹¹⁸ „Mixed Methods bezeichnet eine Forschungsmethode, die eine Kombination von Elementen qualitativer und quantitativer Forschungstraditionen beinhaltet, typischerweise (aber nicht notwendig) innerhalb einer Untersuchung“ Definition nach (Hussy u. a. 2013: 290).

¹¹⁹ (Döring und Bortz 2016: 72)

¹²⁰ Definition Zielpopulation: „Unter Population bzw. Zielpopulation („population“; „target population“) versteht man die Gesamtmenge aller Untersuchungseinheiten (z. B. Personen, Verhaltensweisen, Objekte), über die im Rahmen einer quantitativen empirischen Studie wissenschaftliche Aussagen getroffen werden sollen.“ nach (Döring und Bortz 2016: 292).

mögliche Stichprobe nur innerhalb dieser Zielpopulation. Daher empfiehlt es sich grundsätzlich eine Stichprobe zu definieren:

„Die meisten qualitativen und quantitativen sozialwissenschaftlichen Studien untersuchen weder sämtliche interessierenden Personen, Objekten oder Verhaltensweisen noch einen singulären Einzelfall, sondern arbeiten mit Stichproben („samples“), d. h. mit einer Auswahl von Fällen.“¹²¹

Beschrieben für dieses Verfahren wird die Grundgesamtheit daher wie folgt:

Grundsätzlich wird davon ausgegangen, dass ein hypothetischer Teilnehmer, unabhängig von der Datenerhebungsart, die Grundvoraussetzungen zur Teilnahme an dieser besitzt (z.B. Fähigkeit zur Verbalisierung, Fähigkeiten für Lesen und Schreiben usw.). Auch wird bei dem Teilnehmer eine grundsätzliche Auskunftswilligkeit angenommen. Ebenfalls wird von Personen mit einem rechtlich gültigen Arbeitsverhältnis oder Auftragsverhältnis in Bezug zur Organisationseinheit ausgegangen. Teilnehmer mit körperlichen oder psychischen Einschränkungen müssten ggf. bei der Teilnahme unterstützt werden. Mit Blick auf das Thema Qualitätssicherung wäre es sinnvoll, dass sich Teilnehmer aus der gesamten Softwareentwicklungskette in der Population wiederfinden und diese auch über einen relevanten Zeitraum in ihrer regelmäßigen Tätigkeit an dieser Kette beteiligt sind. Mit dieser Auffassung werden als Merkmale für eine mögliche Population für dieses Verfahren folgende Eigenschaften definiert:

- Direkt oder indirekt beteiligt an einem Softwareentwicklungsverfahren oder Softwarewartungsverfahren innerhalb der gegebenen Organisationseinheit.
- Mindestens sechs Monate in gegebener Organisationseinheit tätig.
- Beherrschen der Sprache, in dem der Fragebogen verfasst wurde.

Es wird empfohlen, dass die Teilnahme für Mitarbeiter freiwillig ist¹²². Dadurch wäre eine Vollerhebung¹²³ theoretisch nicht möglich, es sei denn, die Organisation verpflichtet ihre Mitarbeiter zur Teilnahme am Verfahren. Die Mitarbeiter aus dieser Population (Grundgesamtheit), die sich zur Teilnahme bereit erklären, stellen damit die sogenannte Stich-

¹²¹ (Döring und Bortz 2016: 292)

¹²² Gemäß der vierten Richtlinie der ethischen Prinzipie zur Durchführung von Untersuchungen. (Hussy u. a. 2013: 45). Anm.: An dieser Stelle kann nur eine Empfehlung ausgesprochen werden, da es nicht auszuschließt ist das Mitarbeiter mittels Arbeitsauftrag zur Teilnahme verpflichtet werden.

¹²³ Definition Vollerhebung: „Unter Vollerhebung (Totalerhebung, „total population survey“) versteht man die Untersuchung aller Objekte einer definierten Population.“ (Döring und Bortz 2016: 293).

probe dar. Die Größe der Stichprobe kann allerdings nicht vorhergesagt werden. Idealerweise nehmen so viele Personen der Organisationseinheit wie möglich an der Befragung teil. Es ist aber davon auszugehen, dass auch Organisationseinheiten existieren, die nur eine sehr geringe Anzahl von potentiellen Teilnehmern aufweisen.

„Welche Stichprobenmerkmale anzustreben sind, hängt von der Zielsetzung einer Studie ab: Für qualitative wie quantitative Erkundungsstudien (explorative Studien) sind kleine, nicht-zufällige Stichproben ausreichend.“¹²⁴

Dies lässt den Rückschluss zu, dass für das Verfahren eine solche Situation grundsätzlich kein Hindernis darstellen würde. Allerdings wäre die Wahrung der Anonymität bei einer sehr niedrigen Stichprobe wahrscheinlich nicht mehr zu gewährleisten. Es wird sich daher an einer Empfehlung des Berufsverbandes Deutscher Markt- und Sozialforscher e.V. für Online-Mitarbeiterbefragungen¹²⁵ angelehnt, die für die Stichprobe einen Richtwert von mindestens acht Personen vorgibt.

3.4 Reifegradmodell

Nach Abschluss der Datenerhebung via Befragung und Fragebogen erfolgt das sogenannte Assessment, also eine Bewertung der ermittelten Datenbasis. Für eine solche Bewertung wird im Weiteren das Reifegradmodell konzipiert. Das Reifegradmodell stellt einen Kernaspekt des nach Design-Science entwickelten und angestrebten Artefakts dar. Ein Reifegrad ist wie folgt definiert:

„Der Reifegrad bezieht sich auf die im Entwicklungsmodell definierten Aktivitäten im Vergleich zu sogenannten >> Best Practices<< und deren Umsetzung. Mit Best Practices werden Vorgehensweisen bezeichnet, die sich über viele Jahre in der Praxis als erfolgreich erwiesen haben. Je genauer die durchzuführenden Aktivitäten im zu bewertenden Entwicklungsmodell beschrieben sind und je vollständiger und exakter sie umgesetzt werden, desto höher ist sein ausgewiesener Reifegrad. Er wird oft in Form von Reifegradstufen angegeben“¹²⁶.

Konkret wird sich an bestehenden Modellen von CMMI/TMMi und TPI/TPI Next, auf die bereits in Kapitel 2.2 eingegangen wurde, angelehnt. Auch wenn eine Anlehnung an diese Modelle erfolgt, können die Ergebnisse dieses Modells mit denen der Bestandsverfahren nicht verglichen werden. Das leitet sich in erster Linie von den unterschiedlichen Vorgehensweisen und Kriterien der Verfahren ab. Das bedeutet ebenfalls, dass eine Reifegradstufe dieses Modells inhaltlich und in seiner Bewertung nicht mit einer

¹²⁴ (Döring und Bortz 2016: 297)

¹²⁵ Siehe auch (Bedenk u. a. 2004: 9).

¹²⁶ (Spillner u. a. 2014: 354)

Stufe eines anderen Reifegradmodells vergleichbar ist, auch wenn diese Reifegrade ggf. auf Grund einer Anlehnung gleiche oder ähnliche Abstufungen, Beschreibungen oder Bezeichnungen besitzen.

Die zuvor erläuterten Reifegradmodelle von TPI/TPI Next und CMMI/TMMi nutzen zur Einordnung der Reifegrade eine Ordinal-Skala¹²⁷. Dies wird für das hier konzipierte Reifegradmodell ebenfalls übernommen. Es wird hierbei sich an die vierstufige Variante¹²⁸ der Rangordnung¹²⁹ des TPI-Next orientiert. Daraus ergeben sich folgende Abstufungen des Reifegradmodells: „1. Initial, 2. Mittel, 3. Gut, 4. Sehr gut“. Die niedrigste Bewertung stellt „1 – Initial“ dar und die höchste Bewertung stellt „4 – Sehr gut“ dar. Diese Reifegradstufen werden im weiteren Verlauf mit „RG“ oder „RGx“ abgekürzt, x steht dabei für die zugeordnete Reifegradstufe. Als Beispiel: RG1 bedeutet eine Einstufung auf der ersten Reifegradstufe nach diesem Modell. Wie auch in den bestehenden Modellen werden Merkmale herausgearbeitet, die innerhalb dieser Stufen bewertet werden können. Da TPI-Next und das hier entwickelte Modell sich inhaltlich unterscheiden, kann die konkrete Beschreibung von TPI-Next für die Reifegrade nicht übernommen werden. Die konkreten Inhalte zur Messung des Reifegrades sind zu diesem Zeitpunkt noch nicht ermittelt und werden zu einem späteren Zeitpunkt in Kapitel 3.6. definiert. Aus der Reifegradbewertung aller Merkmale wird ein Mittelwert ermittelt. Dieser Wert stellt die Bewertung für den Gesamtreifegrad dar. Ein Merkmal setzt sich aus verschiedenen Themenpunkten zusammen, welche im weiteren Verlauf erarbeitet werden. Aus den Reifegradbewertungen der Themenpunkte wird ebenfalls ein Mittelwert gebildet. Dieser Mittelwert stellt die Bewertung für den Merkmalsreifegrad dar. Jeder Themenpunkt soll dabei als Fragestellung formuliert und die zugehörigen Antworten einem Reifegrad zugeordnet werden. Alle Zwischen- und Endergebnisse werden nach dem üblichen Verfahren für die Ganzzahl auf- bzw. abgerundet. Das hier beschriebene Modell befindet sich in Abbildung 3-1.

¹²⁷ Vgl. (Porst 2014: 73–75).

¹²⁸ Vgl. (Spillner u. a. 2014: 368–369).

¹²⁹ Siehe dazu auch (Kallus 2016: 72).

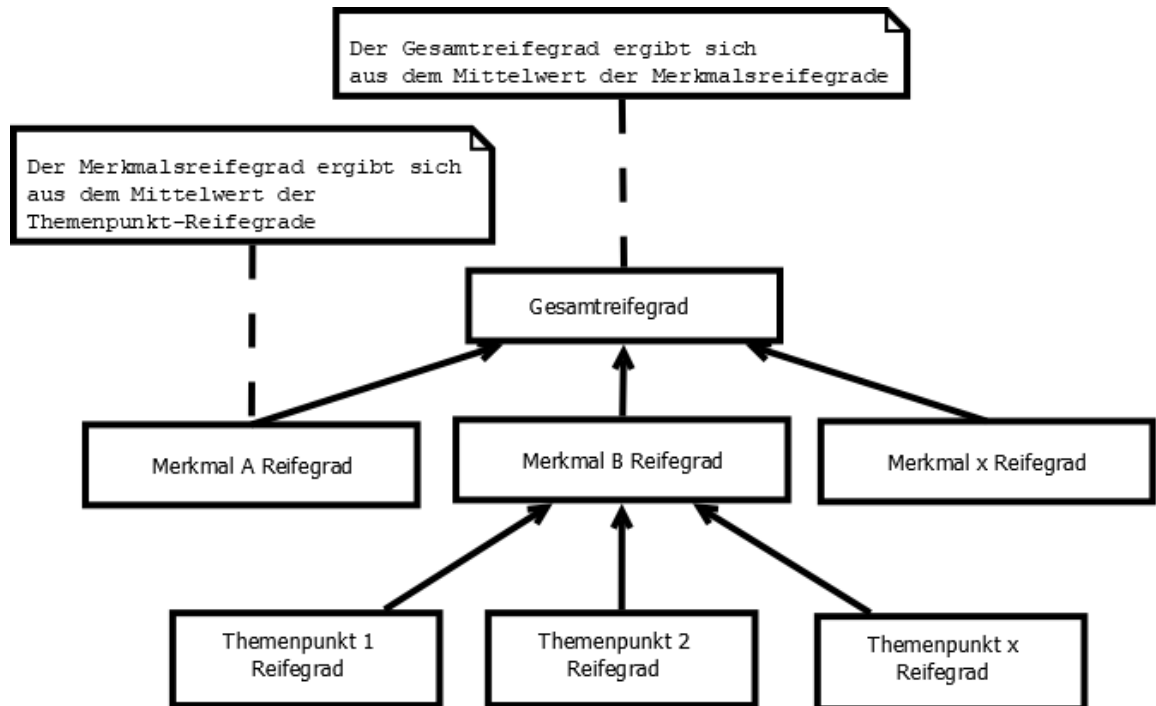


Abbildung 3-1: Reifegradmodell

Eine Fragestellung wird in der Regel mit „Ja“ oder „Nein“ zu beantworten sein. Es kann auch eine abgewandelte Form zur Nutzung kommen, in der eine Mehrfachauswahl oder Einfachauswahl eingesetzt wird. Für jede Fragestellung wird dazu auf Basis der möglichen Antworten eine Übersetzung in RG definiert. Diese Auswahlwahlmöglichkeiten der Fragestellungen werden damit entsprechend einem RG zugeordnet werden. Allgemeine Vorgaben für die Bewertung sind wie folgt:

- Für Fragestellungen mit einer „Ja / Nein“-Ausprägung wird für „Ja“ ein RG4 vergeben und für „Nein“ ein RG1 vergeben. Mögliche Freitextfelder werden in der Regel in diesem Kontext zur Bewertung nicht berücksichtigt.
- Stehen mehrere Ausprägungen innerhalb einer Fragestellung zur Verfügung, muss ein spezifisches Mapping auf eine RGx erfolgen.
- Sollten Antworten nicht verwertbar sein, werden diese mit RG1 gewertet.

Jeder eingereichte Fragebogen stellt einen Datensatz da. Für jeden Datensatz wird pro Item der Reifegrad anhand des Mappings ermittelt. Daraufhin wird der Reifegradmittelwert über alle Datensätze hinweg pro Item bestimmt. Dies stellt wie in Abbildung 3-2 zu erkennen einen Themenpunktreifegrad dar.

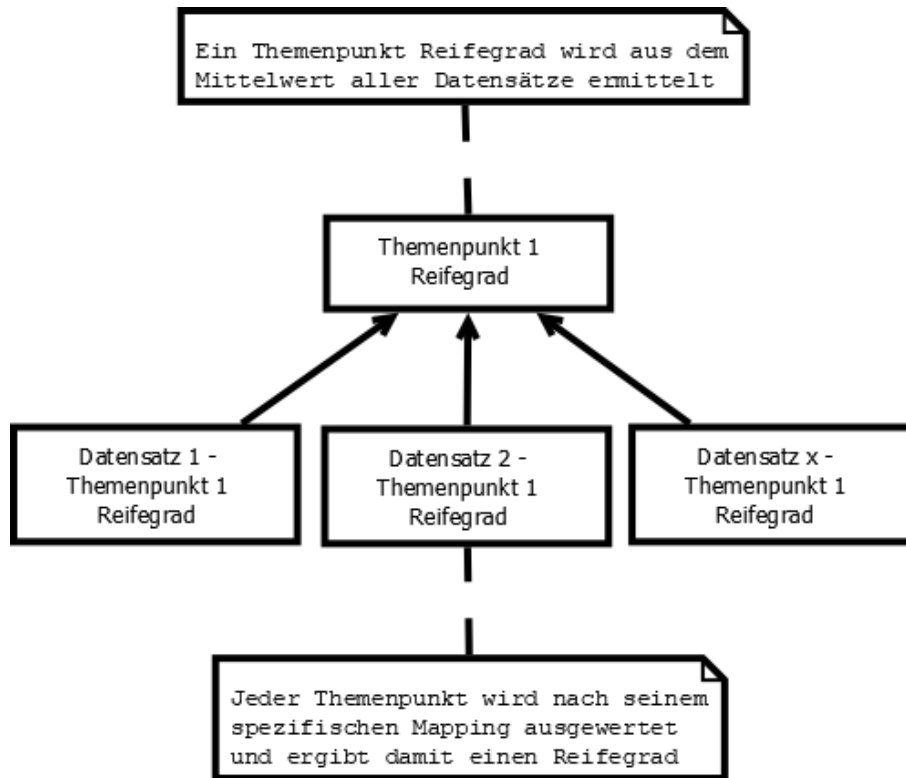


Abbildung 3-2: Themenpunkt Reifegrad

Ähnlich wie CMMI/TMMi könnten die Fragestellungen ebenfalls als Ziele ausgelegt werden, um eine bestimmte Reifegradstufe zu erreichen. Eine Zielerreichung wirkt sich durch die Nutzung des Mittelwerts immer direkt auf den Reifegrad eines Merkmales aus. Im Fokus dieses Modells stehen die reine Reifegradbewertung und daher keine Maßnahmenentwicklung oder Maßnahmenvorschläge. Die Ergebnisse können aber zu gezielteren Analysen genutzt werden oder in die Gestaltung von Maßnahmen mit einfließen. Welche konkreten Maßnahmen auf Basis dieser Einstufungen abgeleitet werden können, steht immer im Kontext zur Organisationseinheit auf die das Modell angewendet wurde. Durch die quantitative Art des Verfahrens und die zu erwartende Form der Datenerhebung ist zusätzlich eine Auswertung pro Item möglich. Diese Auswertung wirkt sich nicht direkt auf den Reifegrad aus und ist daher optional. Die Erkenntnisse können wie die Reifegradbewertung selbst ebenfalls zur weiteren Analyse oder Ableitung von Maßnahmen genutzt werden. Wird das Verfahren mit dem gleichen Fragebogen und der hier beschriebenen Bewertungsweise in mehreren Organisationseinheiten angewendet, so sind die Ergebnisse vergleichbar.

3.5 Gestaltung eines Fragebogens zur Reifegradmessung

Nachdem nun die Art und Weise der Datenerhebung und das Reifegradmodell definiert wurden, wird in diesem Kapitel die Konzeption des Fragebogens erarbeitet. Als inhaltliche Basis für die Fragestellungen dienen die Syllabus des ISTQB sowie ergänzende Werke im Bereich der Softwarequalitätssicherung. Dies soll den inhaltlich fachlichen Teil sicherstellen.

3.5.1 Definierung der Form und Verteilung

Für das hier entwickelte Verfahren wird die elektronische Form¹³⁰ vorgesehen. Nach Erfahrung des Verfassers ist es zu erwarten, dass verschiedene IT-Sicherheitsvorschriften der Behörden den Zugang zu externen Umfragesystemen verbieten könnten. Daher wird kein zentrales Umfragesystem wie z.B. SurveyMonkey genutzt. Es würde sich theoretisch anbieten ein vergleichbares System innerhalb des Behördennetzwerkes aufzubauen. Dies wäre aber nur mit erheblichem Aufwand möglich und lies sich wahrscheinlich nicht in jeder Behörde im gleichen Maß wiederholen. Auch beugt das ggf. datenschutzrechtliche Bedenken vor, da auf diese Weise keine Daten das Organisationsnetzwerk verlassen. Aus diesem Grund wird eine Verteilung mittels E-Mail angestrebt. Innerhalb einer Organisation kann man davon ausgehen, dass die dienstliche E-Mailadresse der Mitarbeiter bekannt sein wird. Damit kann jeder bei Bedarf einzeln angeschrieben werden. Auf diese Weise ist für die anderen Teilnehmer nicht einsehbar, wer einen Bogen erhalten hat. Jeder Bogen erhält dazu innerhalb der Kopfzeile eine ID. Der ausgefüllte Bogen soll via E-Mail zurückgesendet werden. Dies dient dazu, dass maximal nur so viele Fragebogen eingesammelt werden wie, ursprünglich verteilt worden sind. Auf diese Weise können keine Bogen mehrfach eingereicht werden und damit nicht die Ergebnisse durch potentielle mehrfache Einreichung verfälscht werden. Durch die elektrische Form via E-Mail kann allerdings eine Manipulation nicht ausgeschlossen werden.

3.5.2 Gestaltung einer Einleitung

Wie zuvor erläutert wird ebenfalls eine Einleitung verfasst. Diese sollen Teilnehmer der Befragung die Hintergründe, die Zielsetzung und das konkrete Vorgehen der Befragung erläutern. Auch soll dies mögliche Vorbehalte bei den Teilnehmern mindern und in Folge Verzerrung bei der Datenerhebung entgegenwirken. Gleichzeitig dient die Einleitung ebenfalls als Titelseite/Deckblattes des Verfahrens¹³¹.

¹³⁰ Vgl. (Döring und Bortz 2016: 400).

¹³¹ Siehe auch (Porst 2014: 33–46).

„Die Titelseite des schriftlichen Fragebogens sollte – unter dem Gesichtspunkt der Information – eine Reihe von Angaben enthalten, die der potentiellen Befragungsperson anzeigen, womit sie es eigentlich zu tun hat.“¹³²

Zur Gestaltung wurde sich an den Werkern von Porst¹³³ und Kallus¹³⁴ zur Erstellung von Fragebögen orientiert. Die Einleitung beginnt mit einer Begrüßung, Vorstellung und ein paar einführenden Worten:

„Sehr geehrte Damen und Herren,

ich würde mich Ihnen gerne vorstellen: Mein Name ist Mike Wittsiepe. Vielleicht kennen Sie mich bereits aus dem „Forum für Test und Qualitätssicherung (ForTeQ)“ im BVA. Erstmal Vielen Dank, dass Sie sich die Zeit nehmen an dieser Befragung zum Thema „Reifegradmessung in der Softwarequalitätssicherung“ teilzunehmen. Im Rahmen meiner Master-These entwickle ich dieses Werkzeug. Für Ihre Teilnahme bin ich Ihnen daher sehr dankbar, da die Erprobung dieses Werkzeugs einen wichtigen Beitrag zu meiner Abschlussarbeit leistet.

Im Folgenden werden Ihnen kurz die Hintergründe, der Ablauf des Verfahrens und die Zielsetzung dargelegt. Bitte lesen Sie diese aufmerksam und in Ruhe durch.“

Dieser Teil ist spezifisch für die Durchführung im Rahmen dieser Masterthesis verfasst worden und sollte für die Durchführung in einem anderen Kontext angepasst werden. Es folgen daraufhin verallgemeinert die Hintergründe und Zielsetzung der Qualitätssicherung und Reifegradmessung:

„Hintergründe

Die Erwartungen an die Qualität und Leistungsfähigkeit von Software steigen seit der Einführung des Internets und der damit einhergehenden Vernetzung immer weiter an. Gleichzeitig erhöht sich die Komplexität von Software immer mehr. Während man früher eine Anwendung mit verhältnismäßig wenig Aufwand vollständig verstehen und bedienen konnte, werden dazu heutzutage mitunter ganze Teams beschäftigt. Dadurch wachsen die Herausforderungen an die Qualitätssicherung immens. Vor einigen Jahren konnte man vielleicht noch sagen „Das wird schon irgendwie hinhauen“ und wenn mal im Betrieb was in Schieflage geraten war, konnte dies wahrscheinlich ohne große Auswirkung behoben werden. Heutzutage können Fehleranalysen und -korrekturen Wochen in Anspruch nehmen Die mögliche Folge daraus ist: Digitale Prozesse und Dienstleistungen deutschlandweit ggf. sogar europaweit, werden im schlimmsten (Fehler-) Fall blockiert und laufen nicht. Menschen oder Unternehmen warten vielleicht dringend auf Geld oder

¹³² (Porst 2014: 36)

¹³³ Siehe dazu (Porst 2014).

¹³⁴ Siehe dazu (Kallus 2016: 86–89).

Zahlungen. Und wenn Software im Bereich der öffentlichen Sicherheit verwendet wird, könnten sogar Menschenleben auf dem Spiel stehen. Es ist allgemein anerkannt, dass in jeder Software Fehler stecken. Idealweise werden vor einer Inbetriebnahme so viele Fehler wie möglich entdeckt. Die Mittel dazu sind allerdings begrenzt. Es kommt also immer darauf an mit den gegebenen Mitteln mindestens die kritischen und gefährlichen Fehler frühzeitig aufzudecken, obwohl man vollkommen fehlerfreie Software vermutlich niemals herstellen kann.

Zielsetzung

Ziel ist es also innerhalb eines Softwareentwicklungsverfahrens oder -wartungsverfahrens die gegebenen Mittel so kostengünstig und effizient wie möglich zu nutzen. Vor dem Hintergrund der steigenden Erwartungen ist dies häufig eine Gratwanderung und etwas wo jede Organisation sich konstant weiterentwickeln kann. Der Zweck dieser Befragung ist es diesen Entwicklungsprozess im Rahmen der Qualitätssicherung zu unterstützen. Dazu wird ein sogenannter Reifegrad erhoben. Damit ist es möglich Entwicklungspotentiale, Herausforderungen und mögliche Erleichterungen einfach aufzudecken. Und wer könnte besser bei der Aufdeckung helfen als Mitarbeiterinnen und Mitarbeiter die in ihrer Arbeit täglich diese Software entwickeln bzw. warten. Bitte beachten Sie, es existieren weder „richtige“ noch „falsche“ Antworten. Bitte lesen Sie sich die Fragen aufmerksam durch und treffen, wenn möglich, spontan eine Entscheidung“

Dieser Anteil der Einleitung kann grundsätzlich in der weiteren Anwendung des Verfahrens wiederverwendet werden. Zum Abschluss erfolgt noch eine kurze Erläuterung zum Ablauf:

„Ablauf des Verfahrens

Nachdem Sie das Dokument digital ausgefüllt haben, speichern Sie es bitte im Docx- oder PDF-Format ab und senden es bitte an die folgende E-Mailadresse: „mike.wittsiepe@bva.bund.de“.

Das Verfahren verläuft vollständig anonymisiert. Ihr ausgefüllter Bogen wird nach Empfang von der E-Mail getrennt und gespeichert. Die E-Mail selbst wird daraufhin gelöscht. Jeder Bogen erhält eine einmalige ID, diese dient allein dazu, dass ein Fragebogen nicht mehrfach eingereicht werden kann, da dies sonst die Ergebnisse verfälschen würde. Ihre mögliche Teilnahme an dieser Befragung wurde im Vorfeld mit Ihrem Vorgesetzten abgestimmt und wird von diesem unterstützt. Es ist ebenfalls abgestimmt, dass Ihre Vorgesetzten keinen Einblick in die ausgefüllten Fragebogen erhalten. Die Teilnahme ist selbstverständlich freiwillig und kann jederzeit abgebrochen werden.

Vielen Dank und viele Grüße

Mike Wittsiepe“

In diesem Abschnitt sollten je nach konkretem Verlauf und Planung des Verfahrens Anpassungen erfolgen. Da der Fragebogen in diesem Fall via E-Mail an die Teilnehmer verteilt wird, zählen die Inhalte dieser E-Mail ebenfalls als Teil dieser Einleitung. Diese lautete im konkreten Fall wie folgt:

„Sehr geehrte Damen und Herren,

Sie haben sich bereit erklärt an einer Befragung zum Thema Reifegrad der Softwarequalitätssicherung teilzunehmen. Alle weiteren Informationen und Anweisungen finden Sie in der angehängenen Word-Datei "Befragung.docx". Die PDF Datei "Befragung_Beiblatt.pdf" dient zur Erläuterung bestimmter Begrifflichkeiten. Falls eine Meldung über Markos in Word erscheint, bestätigen Sie diese bitte mit "Inhalt aktivieren".

Nachdem Sie das Dokument digital ausgefüllt haben, speichern Sie es bitte im Docx- oder PDF-Format ab und senden es bitte an die folgende

E-Mailadresse: mike.wittsiepe@bva.bund.de

Sie können dafür auch die Antworten Funktion verwenden.

Die Frist zur Einreichung ist nächste Woche Montag, der 18.05.2020.

Sollten Sie Probleme bei der Ausfüllung des Fragebogens haben, so können Sie gerne jederzeit auf mich zu kommen.

Vielen Dank

Mit freundlichen Grüßen

Mike Wittsiepe“

Die E-Mail enthält ebenfalls eine kurze Begrüßung aber auch wichtige Angaben wie einen Ansprechpartner, eine Frist und ggf. technische Hinweise.

3.5.3 Entwicklung der Merkmalsbereiche und Umfang

Grundsätzlich bietet sich für die methodische Entwicklung des Fragebogens das Werk von K. Wolfgang Kallus „Erstellung von Fragebogen“ an:

„Ziel eines psychometrischen Fragebogens ist es, unterschiedliche Merkmalsausprägungen von Eigenschaften, Einstellungen und Meinungen, Bewertungen, Verhaltenstendenzen, Zuständen oder längerfristigen Reaktions- oder Stimmungslagen bei Personen oder Personengruppen repräsentativ zu erfassen und zahlenmäßig wiederzugeben.“¹³⁵

In diesem Werk werden in einem sehr hohen Detailgrad Items entwickelt. Dieser Detailgrad kann auf Grund von zeitlichen Einschränkungen innerhalb dieser Masterthesis nicht auf demselben Niveau erfolgen. Unter anderem wird nicht auf Themen wie Iteminterkorrelationen, -validitäten und -reliabilität sowie der Trennschärfe von Items eingegangen werden. Das Werk wird vielmehr zur Orientierung und zur Anlehnung des hier entwickelten Fragebogens verwendet. Es werden dort 13 Schritte beschrieben die im Folgenden verkürzt erläutert werden¹³⁶:

1. Abgrenzung und Klärung des Merkmalbereiches und des Teilnehmerkreises
2. Spezifizierung des Merkmalbereiches
3. Gruppierung von Teilbereichen
4. Festlegung von Varianten und Antwortdimensionen
5. Formulierung von Items
6. Prüfen der Items (Sprachlich, Verständlichkeit, Einfachheit)
7. Festlegung von Vorgaben
8. Fixierung der Item Reihenfolge
9. Durchführung einer Studie zur Prüfung der Subtest/Items
10. Kürzung/Ergänzung des Fragebogens
11. Festlegung von Varianten
12. Normierung und Auswertrichtlinien
13. Validierung und Interpretationsrichtlinien

Die Schritte 1 – 3 erfolgen innerhalb dieses Kapitels. Um also den inhaltlichen Umfang des Fragebogens zu bestimmen werden initial Kategorien bzw. Merkmale ermittelt. Aus diesen Kategorien werden Items (Konkrete Fragen und deren Antworten)¹³⁷ und sogenannte Subtests¹³⁸ abgeleitet. Als fachlich inhaltliche Grundlage bietet sich das Inhaltsverzeichnis des Syllabus des „ISTQB Certified Tester Advanced Level Syllabus Test

¹³⁵ (Kallus 2016: 20)

¹³⁶ Vgl. (Kallus 2016: 14–15).

¹³⁷ Definition für Item: „Ein Fragebogen besteht aus systematisch zusammengestellten Frage-Antwort-Einheiten. Diese Frage/Feststellung-Antwort-Kombinationen werden mit dem Begriff Fragebogenitem oder kurz Item bezeichnet.“ (Kallus 2016: 18–19).

¹³⁸ Definition für Subtest: „Zu diesem Zweck werden bei einem Fragebogen die Werte für zusammengehörige Items zu einem Messwert, dem Subtestwert, verrechnet. Diese Zusammenfassung von

Manager“ in der aktuellen Version von 2012¹³⁹ an. Die Kapitel der obersten Ebene des Inhaltsverzeichnisses bestehen aus¹³⁹:

- ~~0. Einführung in den Advanced Level Syllabus~~
- 1. Testprozess
- 2. Testmanagement
- 3. Reviews
- 4. Fehlermanagement
- ~~5. Verbesserung des Testprozesses~~
- 6. Testwerkzeuge und Automatisierung
- 7. Soziale Kompetenz und Teamzusammensetzung
- ~~8. Referenzen~~
- ~~9. Index~~

Die Kapitel 0, 8 und 9 werden nicht verwendet (daher oben bereits durchgestrichen), da sie nur zur Strukturbeschreibung des Syllabus gehören. Kapitel 5 betrifft die Verbesserung des Testprozesses. Dort werden Ansätze wie zum Beispiel TMMi, TPI Next, etc. beschrieben. Die Voraussetzung für die Befragung ist es allerdings nicht, einen dieser Ansätze zur Ausführung zu bringen. Aus diesem Grund wird dieser Punkt ebenfalls gestrichen. Das ergibt sechs Kategorien, die zur Ableitung von konkreten Items genutzt werden. Zum Umfang einer Befragung findet sich folgendes in der Literatur:

„Bei der Festlegung der Testlänge darf die Praktikabilität des Tests und insbesondere die Motivationslage der Probanden nicht aus den Augen verloren werden. Je länger der Test, desto mehr ist damit zu rechnen, dass die Items nicht mehr konstruktgemäß bearbeitet werden.

...

Auch die zur Bearbeitung der Aufgaben vorgesehene Testzeit muss sorgfältig überlegt werden. Hierbei sind vor allem die Testart, die Zielgruppe und der intendierte Geltungsbereich zu berücksichtigen.“¹⁴⁰

Für das Durchlesen der Einleitung und des Beiblattes werden ca. 10 Minuten veranschlagt. Die Ausfüllung des Fragebogens eines Teilnehmers soll bei maximal 30 Minuten liegen. Es wird davon ausgegangen, dass für die Beantwortung eines Items im Schnitt 60 Sekunden benötigt werden. Damit wäre der mögliche Rahmen für den Fragebogen 30 Items. Es sollte allerdings eine Pufferzeit vorgesehen werden, weswegen dieser Rah-

Items zu einem sog. Subtest bildet die Grundlage, um die Güte der Fragen mit den Verfahren der Klassischen Testtheorie (Lord & Novick, 1968) zu analysieren.“ (Kallus 2016: 20).

¹³⁹ Siehe dazu (ISTQB u. a. 2012: 4–5).

¹⁴⁰ (Moosbrugger und Kelava 2012: 35)

men nicht voll ausgereizt werden darf. Damit dem entsprochen wird, wird sich pro Merkmal auf drei Items, die zur Reifegradbewertung genutzt werden, beschränkt. Auf diese Weise besteht auch noch ein Raum für ergänzende Items, die nicht unmittelbar in die Reifegradbewertung einfließen, sondern die Interpretation und Erkenntnisgewinne unterstützen können. Damit würde ein Teilnehmer maximal 40 Minuten für die Ausfüllung benötigen, auch wenn davon auszugehen ist, dass diese Zeit nicht benötigt wird. Dies stellt nach Auffassung des Verfassers einen tragbaren Aufwand für die Teilnehmer dar.

3.5.4 Entwicklung der Fragestellungen

Im Folgenden wird nun auf die zuvor spezifizierten Merkmale bzw. Kategorien eingegangen. Die konkrete Item Entwicklung deckt die Schritte 4 – 8 nach Kallus¹⁴¹ ab. Dabei werden Fragestellungen bzw. Items aus den jeweiligen Themenbereichen des Syllabus abgeleitet. Als Basis dafür werden der „ISTQB Certified Tester Advanced Level Syllabus Test Manager“ in der aktuellen Version von 2012 genutzt. Für Details werden sich ggf. auf den „Certified Tester Foundation Level Syllabus“ in der Version 2018, auf die zugehörigen ISO-Standards oder ergänzende Literatur bezogen.

Da es Ziel ist, im Gegensatz zu den Verfahren TPI/TPI-Next oder CMMI/TMMi, ein vereinfachtes Verfahren zu entwickeln, kann nur ein Anteil von dem gemessen werden was diese Verfahren leisten würden. Die Vielzahl an unterschiedlichen Themen macht es unmöglich eine Auswahl zu jedem Aspekt der Softwarequalitätssicherung darzustellen. Die zuvor erfolgte Eingrenzung auf spezifische Merkmale ändert diesen Umstand nicht. Es kann daher nur eine Annäherung vereinzelter Themenbereiche für die konkrete Reifegradmessung erfolgen. Die Themenbereiche orientieren sich an den zuvor genannten Werken, es kann aber eine Subjektivität¹⁴² des Verfassers bezüglich dieser Themen nicht ausgeschlossen werden. Persönliche und professionelle Erfahrungswerte des Verfassers üben naturgemäß einen Einfluss auf die eingeschätzte Richtigkeit bzw. Einordnung dieser Themenpunkte innerhalb der Qualitätssicherung aus.

Das Ziel der Datenerhebung und somit für den Fragebogen ist der Einblick in die bestehenden Vorgehensweisen einer Organisationseinheit. Ein Item darf also nicht zu sehr in die Detailtiefe eintauchen, da sonst keine standardisierten Antworten entwickelt werden könnten. Dem gegenüber steht, dass Antworten auf oberflächliche Fragestellungen kein verwertbares Ergebnis im Sinne der Zielsetzung liefern. Auch sollten sie inhaltlich in ihrem Charakter nicht den sieben Grundsätzen¹⁴³ des Testens widersprechen. Bei der

¹⁴¹ Siehe auch (Kallus 2016: 14–15).

¹⁴² Beschrieben in Kapitel 2.3 von (Reichert 2016: 78–82).

¹⁴³ Vgl. (ISTQB u. a. 2018: 16–17).

Entwicklung und Formulierung der Items wird sich soweit wie möglich an die Vorgehensweise und Leitlinien von Kallus¹⁴⁴ orientiert, siehe dazu Tabelle 3-1¹⁴⁵.

Tabelle 3-1: Checkliste der Leitlinien zur Itemformulierung

1	<i>Grammatikalisch einfache Sätze/Fragen formulieren.</i>
2	<i>Einfache Aussagen formulieren.</i>
3	<i>Items sollten leicht verstehbar sein.</i>
4	<i>Items sollten gut lesbar sein.</i>
5	<i>Präzise Aussagen/Fragen formulieren.</i>
6	<i>Aussagen klar formulieren.</i>
7	<i>Bei der Wortwahl und Formulierung auf eine klare Bedeutung der Begriffe achten und sicherstellen, dass die Bedeutung der Aussage für alle Antwortenden gleich ist.</i>
8	<i>Unklare Begriffe sollten präzisiert werden.</i>
9	<i>Seltene/ausgefallene und regionale Begriffe vermeiden.</i>
10	<i>Auf die Sprache der Untersuchungspopulation eingehen. Fachbegriffe vermeiden.</i>
11	<i>Affirmativ formulieren und komplexe Negationen, insbesondere doppelte Negationen, vermeiden.</i>
12	<i>Items sollten leicht beantwortbar sein und keine komplexen logischen Schlussfolgerungen und keine komplexen Gedächtnisleistungen von den Antwortenden verlangen.</i>
13	<i>Das geeignete Spezifitätsniveau wählen.</i>
14	<i>Items sollten Bezug zum Erleben und Verhalten der antwortenden Person herstellen.</i>
15	<i>So verhaltensnah wie möglich formulieren.</i>
16	<i>Tabuitems gehören nicht in einen Fragebogen.</i>
17	<i>Die angesprochenen Merkmalsfacetten sollen in der Population möglichst wenig extrem sein.</i>
18	<i>Items sind so zu formulieren, dass sie ohne den Kontext der anderen Items beantwortbar sind.</i>
19	<i>Pro Item nur eine Aussage.</i>
20	<i>Semantisch ergänzende Bereiche systematisch über die Items verteilen.</i>

Items können nach ihrer Antwortmöglichkeit wie folgt unterschieden werden:

- *„Items mit offener Beantwortung,*

¹⁴⁴ Siehe dazu (Kallus 2016: 41–77).

¹⁴⁵ Die Tabelle 3-1 wurde direkt aus dem Werk übernommen (Kallus 2016: 70).

- *Items mit halboffener Beantwortung und*
- *Items mit Antwortvorgaben.*¹⁴⁶

Für die Itementwicklung wird in der Regel die feste bzw. geschlossene Antwortmöglichkeit bevorzugt eingesetzt. Diese sollen die Objektivität und Reliabilität sicherstellen¹⁴⁷. Es wird dabei zwischen Single-Choice-Items (Einfachauswahl) und Multiple-Choice-Items (Mehrfachauswahl) unterschieden¹⁴⁸. Für Single-Choice bieten sich hier u.a. die Ja/Nein-Antwortmöglichkeit an:

*„Aufgrund der einfachen und detaillierten Datenstrukturen bieten Fragebogen mit Ja/Nein-Antworten aber andererseits die breiteste Option für anspruchsvolle Auswertungsmodelle.“*¹⁴⁹

Da in dieser Datenerhebung konkretes Wissen der Teilnehmer abgefragt wird, kann es in Einzelfällen effektiver sein, dem Teilnehmer konkrete Stichworte oder eine Mehrfachauswahl zu Verfügung zu stellen¹⁵⁰. Da wo es vom Verfasser als sinnvoll erachtet wird, werden feste Antwortmöglichkeiten durch die Bereitstellung von Freitextfeldern ergänzt. Teilnehmer erhalten somit die Möglichkeit ggf. fehlende Antwortmöglichkeiten zu ergänzen. Auf diese Weise können ebenfalls aus den Freitextfeldern Anhaltspunkte für weitere Schritte im Sinne der Verbesserung abgeleitet werden. Dies führt zu Items mit halboffener Beantwortung:

*„Bei Items mit halboffener Beantwortung werden die untersuchten Personen gebeten, eine vorgegebene Formulierung mit eigenen Worten zu vervollständigen bzw. zu ergänzen. Auch bei diesen Items kann es zu Problemen bezüglich der Objektivität und der Reliabilität kommen.“*¹⁵¹

Zur Sicherstellung von Objektivität und Reliabilität dieser Items werden die Freitextfelder nicht in die Reifegradbewertung mit einbezogen. Das bedeutet, obwohl es halb-offene Antwortmöglichkeiten geben kann, werden nur die „geschlossenen“ Anteile zur Reifegradbewertung genutzt.

¹⁴⁶ (Döring und Bortz 2016: 87)

¹⁴⁷ (Hussy u. a. 2013: 87–88)

¹⁴⁸ Vgl. (Porst 2014: 53–56).

¹⁴⁹ (Kallus 2016: 49)

¹⁵⁰ Siehe auch (Kallus 2016: 41) und (Porst 2014: 54).

¹⁵¹ (Döring und Bortz 2016: 87)

Im weiteren Verlauf folgt je Merkmal ein Unterkapitel und wird zu Beginn grob vorgestellt. Daraufhin folgt die Vorstellung der einzelnen Items des übergeordneten Merkmals. Jedes Item beinhaltet in Form einer Tabelle folgende Punkte:

- Item Nummerierung
- Kategorie / Merkmal
- Fragestellung
- Antwortmöglichkeiten bzw. Ausprägungen
- Zuordnung über die Art der Fragestellung und ggf. Art der Antworten
- Das Mapping zur Einstufung der Antworten innerhalb des Items bezogen auf das übergreifende Reifegradmodell

Im Anschluss zur Tabelle eines Items erfolgt eine Erläuterung, aus welchen Quellen eine Ableitung vollzogen wurde und mit welcher Sichtweise das Item als solches und ggf. im Gesamtkontext zu betrachten ist. Zur späteren Auswertung wird parallel eine Exceldatei verfasst. Diese wird ein Tabellenblatt als zusammenfassende Ergebnisseite, sowie je Item ein Tabellenblatt zur Auswertung erhalten (Anhang 30 und auf der beiliegenden CD-ROM im Verzeichnis „2_Anhänge“). Diese Tabelle deckt in Teilen Schritt 12 „Normierung und Auswertungsrichtlinien“ nach Kallus ab.

3.5.4.1 Testprozess

Ein Testprozess besteht nach ISTQB Lehrplan aus den folgenden Aktivitäten¹⁵²:

- Testplanung und Teststeuerung
- Testanalyse und Testentwurf
- Testrealisierung und Testdurchführung
- Bewertung von Endkriterien und Bericht
- Abschluss der Testaktivitäten

Diese Aktivitäten gelten unabhängig vom genutzten Softwareentwicklungsvorgehen¹⁵³.

Tabelle 3-2: Item Nr. 1

Item Nr. 1	
Kategorie:	Testprozess

¹⁵² Vgl. (ISTQB u. a. 2012: 9).

¹⁵³ Siehe dazu Kapitel 2.1 zum „Software Development Lifecycle Models“ des CTFL Syllabus (ISTQB u. a. 2018: 28–29).

Frage:	Existiert für Ihre Organisationseinheit eine konkrete Beschreibung (z.B. Teststrategie, Testkonzept) über die eingesetzten Testprozesse? Wenn Ja, in welcher Form liegen diese vor?
Antworten:	Nein (1) oder Ja (4), mit Freitextangabe
Einordnung:	Halboffen, Einfachauswahl
RGx	Angabe in Klammern hinter der möglichen Antwort.

In ISO 29119-3 wird eine Reihe von Dokumentationen vorgestellt¹⁵⁴. Unter anderem gehören Teststrategie und Testkonzept dazu. In diesen sollten die Testprozesse beschrieben sein. Die Verfügbarkeit eines solchen Dokumentes lässt sich sicherlich leicht prüfen. Es kann allerdings nicht erwartet werden, dass Dokumentation in dieser Form grundsätzlich verfügbar ist. Sie könnten ebenfalls andere Titel tragen oder Teile von bestehenden Dokumenten und Vorgaben sein. Daher wird mit diesem Item konkret nach einer Beschreibung der Testprozesse gefragt, unabhängig von dessen Form oder Gestaltung. Mit diesem Item kann zusätzlich aufgedeckt werden, ob den Teilnehmern die bestehenden Beschreibungen grundsätzlich bekannt sind und falls Freitextangaben erfolgen, ob diese einheitlich sind und hier nicht ggf. unerwünschte Informationsquellen vorliegen. Im Fokus steht hier allein die Existenz einer Beschreibung. Inhaltlich kann diese Beschreibung hiermit nicht bewertet werden. Dies wäre nur in einer gesonderten Analyse möglich. Die Testprozessbeschreibung kann grundsätzlich dem Merkmal Testprozess oder Testmanagement zugeordnet werden¹⁵⁵. Da hier konkret die Beschreibung der Testprozesse und die Sicht der Arbeitsebene angesprochen werden, wird dieses dem Merkmal Testprozess zugeordnet. Bei Bedarf kann der Teilnehmer die ihm bekannten Formen von Beschreibungen angeben. Diese können, wenn nicht schon bekannt, für weitere Analysen und Maßnahmen einbezogen werden.

Tabelle 3-3: Item Nr. 2

Item Nr. 2	
Kategorie:	Testprozess
Frage:	Welche Testaktivitäten sind Ihnen bekannt und werden Ihres Wissens nach in Ihrer Organisationseinheit regelmäßig eingesetzt?
Antworten:	Datensicherungstests, Kompatibilitätstests, Katastrophenbewältigungstest (z.B. bei Ausfall), Funktionstests, Installationstest, Integrationstests, Lasttests / Stresstests, Zugänglichkeit- / Benutzbarkeitstests, Ablauftests, Sicherheitstests, Abnahmetests, Systemtests / Systemintegrationstests, Komponententests, White-Box-Tests (z.B. Pfad-, Entscheidungs-, Zustandsbasiert), Erfahrungsbasierte Tests, Explorative

¹⁵⁴ Eine Übersicht ist in Abbildung 5-2 von (Daigl und Glunz 2016: 87) zu finden..

¹⁵⁵ Siehe dazu (ISTQB u. a. 2012: 36) und (ISTQB u. a. 2018: 17–22).

	<p>Tests, Use-Case-Tests, Code-Reviews, Akzeptanztests, Regressionstests, Session-based testing, Pair testing.</p> <p>Alle Antwortmöglichkeiten werden in folgenden Ausprägungen auswählbar sein: Unbekannt, Bekannt, Eingesetzt.</p>
Einordnung:	Geschlossen, Mehrfachauswahl mit Ausprägungen
RGx	Jede gewählte Ausprägung wird mit 0 (Unbekannt), 0 (Bekannt) oder 1 (Eingesetzt) bewertet und summiert. Wurde keine oder mehr als eine Ausprägung pro Antwort ausgewählt, zählt diese nicht. Die Antworten werden daraufhin summiert. Die Werteskala ist wie folgt: 0-5 -> RG1, 6-10 -> RG2, 11-15 -> RG3, 16-21 ->RG4

Die Tätigkeiten des Testens stellt einen wichtigen Aspekt der Qualitätssicherung dar und wurde bereits 1979 im Klassiker „The Art of Testing“ wie folgt definiert:

„Testing is the process of executing a program with the intent of finding errors.“¹⁵⁶

Geht man davon aus, dass keine Übersicht über aktuelle Testaktivitäten innerhalb einer Organisationseinheit besteht, ist es wichtig, zu Anfang in Erfahrung zu bringen, welche Aktivitäten regelmäßig zur Ausführung kommen. Sollte bestehende Dokumentation existieren, können diese hiermit zusätzlich auf ihre Aktualität kontrolliert werden. Diese Testarten sind Teil des Testentwurfs, der Testrealisierung sowie der Testdurchführung. Grundsätzlich werden Testarten wie folgt abgegrenzt:

- „Funktionale Tests und nicht funktionale Tests
- Anforderungs- und strukturbasierte Tests“¹⁵⁷

Diese Unterscheidung ist zu grob für eine initiale Abfrage und sollte weiter verfeinert werden. Die Wahl der Testart in einer Organisation oder einem Projekt ist aber von einer Vielzahl von Faktoren abhängig¹⁵⁸ (z.B. Testobjekt, Ressourcen, Anforderungsart, Tools, etc.).

¹⁵⁶ (Myers u. a. 2012: 6). Anm.: Das Zitat stammte ursprünglich aus der ersten Edition aus dem Werk „The Art of Software Testing“ von Glenford J. Meyer aus dem Jahre 1979. Aus Gründen der Verfügbarkeit dieses Werkes wird sich auf die neuste Edition aus dem Jahre 2012 bezogen, dort ist dieser Satz ebenfalls zu finden.

¹⁵⁷ (Spillner und Linz 2019: 86)

¹⁵⁸ Siehe dazu (ISTQB u. a. 2018: 56) und (Spillner und Linz 2019: 215–218).

„Die einzusetzenden Testverfahren zur Erstellung der Testfälle können nicht standardmäßig festgelegt werden. Für die Auswahl der Verfahren ist eine wohldurchdachte Entscheidung erforderlich.“¹⁵⁹

Daher kann den Teilnehmern nur eine begrenzte Auswahl konkreter Testarten¹⁶⁰ bereitgestellt werden. Diese Auswahl besitzt in Folge auch keinen Anspruch auf Vollständigkeit. Der Verfasser nimmt weiterhin an, dass nicht alle Testarten den Teilnehmern bekannt sein werden. Für die Auswertung ist es wichtig zu berücksichtigen, dass ein unterschiedliches Verständnis über die genutzten Begrifflichkeiten zwischen Teilnehmer und Verfasser vorliegen könnten. Dies muss bei Interpretation der Ergebnisse berücksichtigt werden. Die Antworten werden in Unbekannt, Bekannt und Eingesetzt ausgeprägt. Es wird erwartet, dass durch diese Ausprägungen in Kombination mit ggf. verfügbarer Testprozessbeschreibung ein verwertbares Ergebnis für die weitere Analyse erzielen zu können. Auf Basis der Aussage

„Zur Erstellung der Testfälle nutzen Tester in der Regel eine Kombination aus verschiedenen Testverfahren (abhängig von Zeit und Budget) die besten Ergebnisse zu erzielen.“¹⁶¹

wird für die Übertragbarkeit auf das Reifegradmodell die Annahme aufgestellt: Je mehr unterschiedliche, sich ergänzende Testarten bekannt sind und zum Einsatz kommen, umso höher fällt in Folge der Reifegrad aus. Zusätzlich lässt sich durch die Ausprägungen ableiten, ob ein einheitliches Verständnis über die aktuellen Testaktivitäten innerhalb der Organisationseinheit besteht. Auch wäre ein Rückschluss auf den Wissensstand der Teilnehmer über die zur Verfügung gestellten Testaktivitäten möglich. Die unterschiedlichen Kombinationen können zu unterschiedlichen Interpretationen führen, welche wiederum gezieltere Maßnahmen zur Reifegradsteigerung fördern.

Tabelle 3-4: Item Nr. 3

Item Nr. 3	
Kategorie:	Testprozess
Frage:	Werden Ihrer Meinung nach nicht-funktionale Anforderungen im Rahmen des aktuellen Anforderungsprozesses ausreichend berücksichtigt?
Antworten:	Nein (1) oder Ja (4)

¹⁵⁹ (Spillner und Linz 2019: 218)

¹⁶⁰ Die Auswahl bezieht sich auf folgende Quellen: (Daigl und Glunz 2016: 142), (ISO 29119-4 2015: 33–37), (Spillner und Linz 2019: 62–93, 192–215), (ISTQB u. a. 2018: 30–47, 56–62). Anm.: Zwischen den Quellen ergeben sich auch Überschneidungen.

¹⁶¹ (Spillner und Linz 2019: 215)

Einordnung:	Geschlossen, Einfachauswahl
RGx	Angabe in Klammern hinter der möglichen Antwort.

Nicht-funktionale Anforderungen sind ein wichtiger Teil des Testprozesses, welche bereits frühzeitig in der Testanalyse und Testdesign berücksichtigt werden müssen.

„Non-functional testing of a system evaluates characteristics of systems and software such as usability, performance efficiency or security.

...

Contrary to common misperceptions, non-functional testing can and often should be performed at all test levels, and done as early as possible. The late discovery of non-functional defects can be extremely dangerous to the success of a project.“¹⁶²

Damit diese in die Testbasis einfließen können ist daher wichtig, dass diese ebenfalls im Anforderungsprozess entsprechend berücksichtigt werden.

3.5.4.2 Testmanagement

Das Testmanagement behandelt im Kern Themen wie die Steuerung sowie Gestaltungen der Testprozesse mit Blick auf das Umfeld und die unterschiedlichen Stakeholder:

„Da Testprozesse nur Mehrwert schaffen, wenn sie zum Erfolg des Projekts bzw. Programms beitragen (oder schwerere Fehlerwirkungen oder Ausfälle verhindern), muss der Testmanager die Prozesse entsprechend planen und steuern. Konkret heißt das, dass der Testmanager die Testprozesse, einschließlich der damit verbundenen Aktivitäten und Arbeitsergebnisse, entsprechend der Erfordernisse und Umstände der anderen Stakeholder, deren Aktivitäten (z.B. Softwareentwicklungslebenszyklus, in dem getestet wird) und deren Arbeitsergebnisse (z.B. Anforderungsspezifikationen) gestaltet.“¹⁶³

Tabelle 3-5: Item Nr. 4

Item Nr. 4	
Kategorie:	Testmanagement
Frage:	Welche Qualitätsmerkmale decken die genutzten Testaktivitäten innerhalb Ihrer Organisationseinheit ab?
Antworten:	Funktionale Eignung, Leistungsfähigkeit, Komptabilität, Benutzerfreundlichkeit, Verlässlichkeit, Sicherheit, Wartbarkeit, Portabilität. Alle Antwortmöglichkeiten werden in folgenden Ausprägungen auswählbar sein: Nicht-Abgedeckt, Abgedeckt.

¹⁶² (ISTQB u. a. 2018: 40)

¹⁶³ (ISTQB u. a. 2012: 20)

Einordnung:	Geschlossen, Mehrfachauswahl mit Ausprägungen
RGx	Jede Ausprägung mit 0 (Unbekannt), 0 (Bekannt) oder 1 (Abgedeckt) gewählt werden. Die Antworten werden daraufhin summiert. Die Werteskala ist wie folgt: 0-2-> RG1, 3-4 -> RG2, 5-6 -> RG3, 7-8 -> RG4

Eine Auflistung der Testaktivitäten durch die Teilnehmer mit Item 4 allein wird nicht ausreichen. Grund hierfür ist eine fehlende inhaltliche Tiefe. Für eine Einschätzung wäre eine vollständige Analyse der Testbasis notwendig, die aus Erfahrung jedoch viel Aufwand und Zeit benötigt. Um dennoch einen groben Einblick zu erhalten, werden die Teilnehmer aufgefordert die Abdeckung mit Sicht auf die Qualitätsmerkmale¹⁶⁴ einzuschätzen. Es gilt für die Auswertung zu berücksichtigen, dass den Teilnehmern die Begrifflichkeiten nicht vollständig bekannt sein könnten. Um den entgegen zu wirken, wird eine Beschreibung der QS-Merkmale als Beiblatt in den Fragebogen aufgenommen. Die Beschreibung der QS-Merkmale stammt vollständig aus einer englischen Internetseite¹⁶⁵ zur ISO 25000. Es kann nicht davon ausgegangen werden, dass alle Teilnehmer Englisch im ausreichenden Maß beherrschen. Also wurden die Texte zum besseren Verständnis mittels Google Übersetzer¹⁶⁶ ins Deutsche übersetzt. Der übersetzte Text wurde vom Verfasser dieser Thesis auf die inhaltliche sowie sprachliche Korrektheit und auf seine Verständlichkeit hin überprüft, darüber hinaus aber nicht angepasst. Das Beiblatt befindet sich in Anhang 1. Die Antworten werden für dieses Item in Nicht-Abgedeckt und Abgedeckt ausgeprägt. Die unterschiedlichen Kombinationen können zu unterschiedlichen Interpretationen führen welche wiederum gezieltere Maßnahmen zur Reifegradsteigerung fördern.

Tabelle 3-6: Item Nr. 5

Item Nr. 5	
Kategorie:	Testmanagement
Frage:	Welche Relevanz haben die Qualitätsmerkmale Ihrer Meinung nach innerhalb Ihrer Organisationseinheit?
Antworten:	Funktionale Eignung, Leistungsfähigkeit, Komptabilität, Benutzerfreundlichkeit, Verlässlichkeit, Sicherheit, Wartbarkeit, Portabilität. Jedes Merkmal soll dabei eine Einordnung in „Wichtig“, „Mittel“ und „nicht Wichtig“ erhalten.
Einordnung:	Geschlossen, Reihenfolgen herstellen.

¹⁶⁴ Qualitätsmerkmale nach (ISO 29119-4 2015: 32).

¹⁶⁵ (iso25000.com 2019)

¹⁶⁶ Siehe (Google Ireland Limited o. J.).

RGx	Keine Bewertung. Diese Antworten dienen allein dazu eine umfassendere Interpretation des Items 4 für mögliche QS-Maßnahmen vornehmen zu können.
-----	---

Die Teilnehmer werden gebeten, eine Aussage darüber zu treffen welche Qualitätsmerkmale¹⁶⁷ ihnen wichtig sind. Diese Aussage dient dazu eine Vorstellung von den ausgeführten Testaktivitäten und der Sichtweise auf diese zu erhalten und damit diesen Aspekt in zukünftige Reifegradsteigerungsmaßnahmen mit einfließen zu lassen. Ebenfalls wird die Interpretation des Items 4 unterstützt. Damit können ggf. widersprüchliche Aussagen aufgedeckt werden. Zum Beispiel, wenn das QS-Merkmal Komptabilität im Schnitt von den Teilnehmern als wichtig eingestuft wird, es aber in Item 4 durch einen vergleichbaren Anteil der Teilnehmer als „Nicht-Abgedeckt“ geschätzt worden ist. Diese Erkenntnis kann so in weitere Analysen einbezogen werden. Wie sich die Relevanz der QS-Merkmale für die Teilnehmer darstellt, lässt aber nicht auf den Reifegrad schließen. Für dieses Item wird ebenfalls das in Item 4 bereits erwähnte Beiblatt benötigt. Die Items 4 und 5 sind bewusst hintereinandergestellt, so dass ein erneutes Lesen nicht mehr notwendig sein sollte.

Tabelle 3-7: Item Nr. 6

Item Nr. 6	
Kategorie:	Testmanagement
Frage:	Von welchen Personengruppen werden Testaktivitäten Ihrer Meinung hauptsächlich innerhalb Ihrer Organisationseinheit ausgeführt?
Antworten:	Interner Mitarbeiter, Externe Mitarbeiter
Einordnung:	Geschlossen, Einfachauswahl
RGx	Keine Bewertung über den Reifegrad.

Dieses Item fragt nicht direkt den Reifegrad ab. Es dient vielmehr dazu eine Einschätzung darüber zu erhalten, wer hauptsächlich an den Testaktivitäten beteiligt ist. Wenn zum Beispiel der vollständige Testprozess an externe Mitarbeiter gegeben worden ist, aber nur interne Mitarbeiter befragt worden sind, liegt der Schluss nahe, die Befragung mit einer anderen Teilnehmergruppe erneut durchzuführen. Ob und inwiefern es zielführend ist, Testprozesse von internen oder externen Mitarbeitern durchführen zu lassen hängt, grundsätzlich von der Unternehmensstrategie als Ganzes ab. Unabhängig davon

¹⁶⁷ Qualitätsmerkmale nach (ISO 29119-4 2015: 32).

ist es dennoch wichtig, diesen Umstand innerhalb der Analyse für Verbesserungsmaßnahmen frühzeitig zu betrachten¹⁶⁸.

Tabelle 3-8: Item Nr. 7

Item Nr. 7	
Kategorie:	Testmanagement
Frage:	Welche Stakeholder wurden Ihrer Meinung nach für die bestehenden Qualitätssicherungsmaßnahmen Ihrer Organisationseinheit berücksichtigt?
Antworten:	Entwickler, Architekten, Endnutzer, Fachbereich, IT, Support bzw. Helpdesk, Sachgebietsleiter, Sicherheitsbeauftragter / Abteilung, Datenschutzbeauftragter / Abteilung, Releasemanagement, Sonstige: Freitextangabe.
Einordnung:	Halboffen, Mehrfachauswahl
RGx	Jede Auswahl zählt einen Punkt und wird zum Ergebnis addiert. Die Summe der Punkte ist wie folgt im Modell einzustufen: 0-3 -> RG1, 4-6 -> RG2, 7-9 -> RG3, 10-11 -> RG4

Eine weitere grundsätzliche Aufgabe des Testmanagements ist es, alle Stakeholder in die Themen der Qualitätssicherung mit einzubeziehen¹⁶⁹. Item 7 ermöglicht die Auswertung darüber, welche Stakeholder laut der Teilnehmer bereits eingebunden sind und kann damit ggf. noch nicht eingebundene Stakeholder aufdecken. Diese Standardliste¹⁶⁹ deckt eine Reihe von typischen Stakeholdern in Softwareentwicklungsverfahren ab. Es gilt erneut das Prinzip: Je mehr eingebundene Stakeholder, desto höher der Reifegrad. Da aber eine Vielzahl von unterschiedlichen Rollen- und Aufgabenkombinationen existiert, erhebt diese Liste keinen Anspruch auf Vollständigkeit. Bei Bedarf können die Teilnehmer eigene Rollen angeben.

Tabelle 3-9: Item Nr.8

Item Nr. 8	
Kategorie:	Testmanagement
Frage:	Welche Testmetriken finden Ihrer Ansicht nach innerhalb Ihrer Organisationseinheit Verwendung?

¹⁶⁸ Siehe dazu Mastertestkonzept (ISTQB u. a. 2012: 39–40) und u.a. Outsourcing (ISTQB u. a. 2012: 49–50).

¹⁶⁹ (ISTQB u. a. 2012: 20–21)

Antworten:	Anforderungsabdeckung, Abweichungsübersicht (Offen, Geschlossen, ...), Codeüberdeckung, Testfallstatus (Offen, Ausgeführt, Blockiert, in Arbeit, ...), Testumgebungsverfügbarkeit, Sonstige: Freitextangabe
Einordnung:	Halboffen, Mehrfachauswahl
RGx	Jede Auswahl zählt einen Punkt und wird zum Ergebnis addiert. Die Summe der Punkte ist wie folgt im Modell einzustufen: 0-1 -> RG1, 2 -> RG2, 3-4 -> RG3, 5-6 -> RG4

Mit Item 8 soll das Thema Metriken abgedeckt werden. ISTQB fasst die Aufgabe von Metriken wie folgt zusammen:

„Mit Metriken können Tester ihrem Management in gleichbleibender Art und Weise berichten und der Fortschritt im zeitlichen Verlauf lässt sich kohärent verfolgen. Häufig müssen sie diese in Besprechungen verschiedenen Ebenen von Stakeholdern, vom technischen Personal bis zur Geschäftsleitung, präsentieren.“¹⁷⁰

Die Auswahl deckt einen Teil der möglichen Metriken aus ISTQB¹⁷¹ ab, von dem der Verfasser der Auffassung ist, dass diese die Teilnehmergruppe betreffen könnten. Der vielseitige Einsatz von Metriken stärkt den Reifegrad der QS. Es besteht allerdings kein Anspruch auf Vollständigkeit dieser Liste. Darüber hinaus könnten auch weitere Metriken existieren, die aber der Teilnehmergruppe nicht bekannt sein müssen.

3.5.4.3 Review

Reviews gehören nach ISTQB zu den statischen Testverfahren:

„Grundidee der statischen Prüfung ist die Prävention: Fehler(zustände)¹ und Abweichungen sollen erkannt werden, noch bevor sie im weiteren Verlauf der Softwareentwicklung negativ zum Tragen kommen.“¹⁷²

Es werden drei Review-Arten unterschieden, jede mit ihrem eigenen Fokus¹⁷³:

- Management Reviews/Audits mit Blick auf den Projektverlauf

¹⁷⁰ (ISTQB u. a. 2012: 44)

¹⁷¹ Erläutert in (ISTQB u. a. 2012: 43–48) und detailliert in (Spillner u. a. 2014: 191–224).

¹⁷² (Spillner und Linz 2019: 103). Anm.: Fußnote ¹ aus dem Original „Die Schreibweise soll sowohl Fehlerzustände im Programmtext als auch Fehler in anderen Arbeitsergebnissen umfassen.“

¹⁷³ Vgl. (Spillner und Linz 2019: 119), (Spillner u. a. 2014: 312) und (ISTQB u. a. 2012: 53–54).

- Assessments mit Blick auf Prozesse
- Reviews mit Blick auf Dokumente bzw. Arbeitsergebnisse

Diese können formal oder informal durchgeführt werden¹⁷⁴. Die Zielgruppe für die Teilnehmer dieses Fragebogens ist die sogenannte „Arbeitsebene“ einer Organisationseinheit, daher wird sich im Weiteren auf die Reviews von Dokumenten und Arbeitsergebnissen konzentriert. Arbeitsergebnisse werden nach ISTQB wie folgt definiert:

„Beim Arbeitsergebnis kann es sich um eine Anforderungsspezifikation, eine User Story, ein technisches Dokument, einen Testfall, den Programmcode, oder um irgendein anderes Arbeitsergebnis handeln, das während der Softwareentwicklungs- oder Wartungsprozesse erstellt wird.“¹⁷⁵

Tabelle 3-10: Item Nr. 9

Item Nr. 9	
Kategorie:	Review
Frage:	Auf welchen Dokumenten oder Arbeitsergebnissen werden Reviews in Ihrer Organisationseinheit durchgeführt?
Antworten:	Anforderungen, Akzeptanzkriterien, Testfälle, Spezifikationsdokumente, Schnittstellenbeschreibungen, Testergebnisse, Fehlertickets, Programmcode, Konfigurationsdateien, Testdaten, Benutzerhandbücher, Sonstige Dokumente: Freitextangabe.
Einordnung:	Halboffen, Mehrfachauswahl
RGx	Jede Antwort erhält einen Punkt. Die Summe der Punkte ist wie folgt im Modell einzustufen: 0-3 -> RG1, 4-6 -> RG2, 7-9 -> RG3, 10-12 -> RG4

Mit Item 9 werden verschiedene Arbeitsergebnisse oder Dokumente abgefragt, die sich im Rahmen eines üblichen Softwareentwicklungsprojekts befinden könnten¹⁷⁶. Wie zuvor, wird die Einstufung des RG über die Anzahl der unterschiedlichen Dokumente die ein Review erfahren abgeleitet. Die Liste erhebt hierbei keinen Anspruch auf Vollständigkeit.

Tabelle 3-11: Item Nr. 10

Item Nr. 10	
Kategorie:	Review

¹⁷⁴ Vgl. (ISTQB u. a. 2018: 48).

¹⁷⁵ (ISTQB u. a. 2012: 60)

¹⁷⁶ Vgl. (ISTQB u. a. 2012: 53–54) und (ISTQB u. a. 2018: 46).

Frage:	Werden zur Unterstützung der Reviews jeweils einheitliche Checklisten eingesetzt?
Antworten:	Nein (1) oder (4), Wenn Ja, für folgenden Dokumente
Einordnung:	Halboffen, Einfachauswahl
RGx	Angabe in Klammern hinter der möglichen Antwort.

Der Verfasser ist der Auffassung, dass für die Durchführung von möglichst einheitlichen Reviews, Checklisten ein einfaches und strukturiertes Mittel sind. Diese unterstützen zusätzlich die Standardisierung von Dokumenten. Auch lassen sie sich erfahrungsgemäß leicht in Arbeitsabläufe einführen und können auch von unerfahrenen Mitarbeitern effektiv genutzt werden. Checklisten können schon bei der Erstellung dazu beitragen, die Qualität zu verbessern, da sie ein einheitliches Verständnis über die inhaltliche Nutzbarkeit sowie die dort genutzten Begrifflichkeiten von Dokumenten fördern. Wie wichtig so was sein kann, wurde auch im Rahmen von Fallstudien durch Avci (2008) beobachtet¹⁷⁷. Andere Reviewtechniken wie zum Beispiel die „Ad-Hoc-Methode“ sind möglich, sind aber nach Auffassung des Verfassers entweder zu unstrukturiert (siehe Ad-Hoc¹⁷⁸), zu aufwändig (siehe Szenario-Methode) oder beanspruchen zu viele Ressourcen. Außerdem empfiehlt der Bundesrechnungshof, in einem zuvor benannten Bericht, einheitliche Vorlagen und Prüflisten als einer der vorgeschlagenen Maßnahmen:

„Einheitliche Dokumentenvorlagen und Prüfkriterienlisten könnten die Qualitätssicherung bereits in der Konzeptphase unterstützen und die Erstellung erforderlicher Planungs- und Konzeptionsdokumente beschleunigen.“¹⁷⁹

Tabelle 3-12: Item Nr. 11

Item Nr. 11	
Kategorie:	Review
Frage:	Werden innerhalb Ihrer Organisationseinheit Tester an einem Reviewprozess der zu entwickelten Anforderungen beteiligt?
Antworten:	Nein (1) oder Ja (4)
Einordnung:	Geschlossen, Einfachauswahl

¹⁷⁷ Siehe dazu (Avci 2008: 258). Die Dissertation beschäftigte sich im Kern damit wie man aus Fehlern in der Softwareentwicklung lernen kann mit einem Fokus auf die Anforderungsanalyse und Qualitätssicherung.

¹⁷⁸ Siehe dazu (ISTQB u. a. 2018: 52–54).

¹⁷⁹ (Bundesrechnungshof 2008: 4)

RGx	Einstufung in Klammern bei den Antworten angegeben. Liegt kein Reviewprozess für Anforderungen vor, wird eine Bewertung von RG1 angenommen
-----	--

Erfahrungsgemäß lassen sich so bereits frühzeitig potentielle Abweichungen oder Anforderungslücken aufdecken¹⁸⁰. Grundsätzlich wird es auch empfohlen, Anforderungen frühzeitig auf ihre Testbarkeit hin zu überprüfen¹⁸¹. Eine solche Vorgehensweise würde zusätzlich einen der Punkte aus dem Praxisproblem ansprechen:

„Anforderungen an die Produktqualität wurden selten vollständig im Vertrag oder im Fachlichen Feinkonzept beschrieben. Oftmals waren sie auslegungsbedürftig. Ebenso wie die Qualitätsansprüche an den Entwicklungs- und den Qualitätssicherungsprozess wurden sie in der Regel erst bei aufkommenden Problemen im Zuge des Entwicklungs- oder Abnahmetests diskutiert.“¹⁸²

Daher wird aus der Erkenntnis, ob Tester am Reviewprozess beteiligt werden, auf den Reifegrad geschlossen.

3.5.4.4 Fehlermanagement

Fehler ist ein Oberbegriff, der wie folgt definiert wird:

*„1. Oberbegriff für Fehlhandlung, Fehlerzustand, Fehlerentwicklung.
2. Nichterfüllung einer festgelegten Anforderung.“¹⁸³*

Ein Fehler entsteht grundsätzlich durch eine Fehlerhandlung, die in Folge zu einem Fehlerzustand führt¹⁸⁴. Wurde ein Fehlerzustand beobachtet, stellt dies die Fehlerwirkung dar. Jegliche Art von Softwaretest dient in erster Linie dazu, Fehlerwirkungen aufzudecken¹⁸⁵. Allerdings führt nicht jeder Fehlerzustand in direkter Folge zu einer Fehlerwirkung. Aus diesem Fakt leitet sich auch das erste Grundprinzip des Testens ab:

*„1. Testing shows the presence of defects, not their absence
Testing can show that defects are present, but cannot prove that there are no defects.“*

¹⁸⁰ Siehe dazu (ISTQB u. a. 2018: 14) und (ISTQB u. a. 2012: 21).

¹⁸¹ Erläutert in (Spillner und Linz 2019: 34–35) im Kapitel Testanalyse.

¹⁸² (Bundesrechnungshof 2008: 4)

¹⁸³ (Spillner und Linz 2019: 313)

¹⁸⁴ Siehe Kapitel 2.1.1 (Spillner und Linz 2019: 10–12).

¹⁸⁵ Siehe dazu auch (Myers u. a. 2012: 5–9).

*Testing reduces the probability of undiscovered defects remaining in the software but, even if no defects are found, testing is not a proof of correctness.*¹⁸⁶

Dieses geht mit dem allgemeinen Verständnis einher, dass nie alle Fehlerzustände eines umfangreichen Systems aufgedeckt werden können.¹⁸⁷ Bei vernetzten Systemen ist ebenfalls möglich, dass sich Fehlerzustand und -wirkung nicht innerhalb eines Systems befinden. Für den Begriff Fehlerwirkung sind alternativ auch folgende Begriffe gebräuchlich: Fehler, Defect, Bug, Abweichung¹⁸⁸. Grundsätzlich kann jeder, der mit dem Softwareprodukt in Kontakt kommt, eine Fehlerwirkung beobachten, ggf. melden und damit einen Fehlerbericht erstellen. In erste Linie können das Personen aus den folgenden Bereichen sein: Entwicklung, Fachbereich, Tester, Endnutzer sowie Systeme, deren Schnittstellen ggf. genutzt werden. Fehlerberichte können eine Fehlerwirkung oder einen Fehlerzustand beschreiben¹⁸⁹. Informationen über potentiell aufgedeckte Fehlerwirkungen bzw. Fehlerzustände stellen ein Kernergebnis von Testaktivitäten dar. Konkret wird hier der Fokus auf das Fehlermanagement gelegt.

*„Der Fehlermanagementprozess eines Unternehmens und das verwendete Fehlermanagementwerkzeug sind von entscheidender Bedeutung, nicht nur für das Testteam, sondern für alle Teams, die an der Softwareentwicklung beteiligt sind. Die Daten und Informationen, die durch ein effektives Fehlermanagement gesammelt werden, geben dem Testmanager und den Stakeholdern Aufschluss über den Zustand eines Projektes während des ganzen Entwicklungslebenszyklus.“*¹⁹⁰

Fehlermanagement regelt also den Fehlerlebenszyklus, der die Prozesse und den Umgang mit aufgedecktem Fehler beschreibt¹⁹¹. Es stellt laut dem ISTQB einen Kernaspekt der Softwareentwicklung dar¹⁹².

Tabelle 3-13: Item Nr. 12

Item Nr. 12	
Kategorie:	Fehlermanagement

¹⁸⁶ (ISTQB u. a. 2018: 16). Anm.: „Defect“ entspricht der zuvor gegebenen Definition von Fehler.

¹⁸⁷ Siehe auch (Spillner und Linz 2019: 9).

¹⁸⁸ Siehe auch (Spillner u. a. 2014: 328–329) für einer Erläuterung zum Begriff „Abweichung“.

¹⁸⁹ Vgl. (Daigl und Glunz 2016: 110–111).

¹⁹⁰ (ISTQB u. a. 2012: 60)

¹⁹¹ Siehe dazu (ISTQB u. a. 2012: 60–62).

¹⁹² Siehe dazu (ISTQB u. a. 2012: 64–65).

Frage:	Existiert für Ihre Organisationseinheit eine konkrete Beschreibung über die eingesetzten Fehler- bzw. Abweichungsprozesse? Wenn Ja, in welcher Form liegen diese vor?
Antworten:	Nein (1) oder ja (4), mit Freitextangabe
Einordnung:	Halboffen, Einfachauswahl
RGx	Angabe in Klammern hinter der möglichen Antwort.

Ein Fehlerbericht dient ebenfalls zur Unterstützung der Testfortschrittsüberwachung¹⁹², da unter anderem zum Beispiel von der Anzahl und Schwere der Berichte auf den Fortschritt der Entwicklung und der Qualität rückgeschlossen werden kann. Damit dies möglich wird, werden konkrete Prozesse und Abläufe zur Erstellung und Verarbeitung von Fehlerberichten benötigt¹⁹³. Mit Item 1¹⁹⁴ wurden Testprozesse und deren Beschreibung bei den Teilnehmern abgefragt. Selbiges gilt auch für eine Abweichungsprozessbeschreibung, da sich diese in ihrem Zweck und Inhalt unterscheidet. Es kann nicht davon ausgegangen werden dass sich diese beiden Beschreibungen innerhalb desselben Dokuments befinden. Auch könnte eine Beschreibung der Prozesse bestehen, welche aber den Teilnehmern nicht bekannt ist.

Tabelle 3-14: Item Nr. 13

Item Nr. 13	
Kategorie:	Fehlermanagement
Frage:	Wird Ihrer Meinung nach innerhalb Ihrer Organisationseinheit für die Fehler- bzw. Abweichungsdokumentation (dem sogenannten Fehlerbericht) ein einheitliches Schema genutzt?
Antworten:	Nein (1) oder Ja (4)
Einordnung:	Geschlossen, Einfachwahl
RGx	Einstufung in Klammern bei den Antworten angegeben. Liegt kein Schema vor, wird eine Bewertung von RG1 angenommen.

Ein Fehlerbericht stellt die Dokumentation einer beobachteten Fehlerwirkung dar. Er kann eine Vielzahl von möglichen Informationen¹⁹⁵ zur Beobachtung selbst und deren Situation beinhalten. Auch wenn einige Standards (z.B. ISO 29119, ISO 9126, ISO 25000,

¹⁹³ Nach ISO 29119 entstehen Fehlerberichte innerhalb eines Fehlermeldeprozesses. Siehe dazu (Daigl und Glunz 2016: 83–84).

¹⁹⁴ Siehe Kapitel 3.5.4.1.

¹⁹⁵ Vgl. (Spillner u. a. 2014: 62–63) und (ISTQB u. a. 2018: 76–77).

IEEE 829, etc.) dazu existieren, bestehen für jedes Softwareentwicklungsverfahren, dessen Testobjekt und dessen spezifisches Umfeld unterschiedliche Anforderungen an solch einen Bericht.

„Allerdings sollten die wesentlichen Informationen, die im Laufe des Lebenszyklus gesammelt werden, einheitlich sein. Idealerweise sollten sie sogar projektübergreifend konsistent sein, um einen aussagekräftigen Vergleich von Fehlerdaten aller Projekte zu ermöglichen.“¹⁹⁶

Im Kontext dieser Fragstellung wird sich auf die Organisationseinheit bezogen und dass mindestens innerhalb dieser ein einheitliches Schema genutzt werden sollte. Die Fragestellung bezieht sich bewusst auf die Nutzung eines Schemas, weil es möglich ist, dass ein Schema zwar existiert, aber nicht genutzt wird oder unbekannt ist.

Tabelle 3-15: Item Nr. 14

Item Nr. 14	
Kategorie:	Fehlermanagement
Frage:	Wird innerhalb Ihrer Organisationseinheit ein Fehler-Workflow-Werkzeug eingesetzt?
Antworten:	Nein (1) oder Ja (4)
Einordnung:	Geschlossen, Einfachwahl
RGx	Einstufung in Klammern bei den Antworten angegeben. Liegt kein Schema vor, wird eine Bewertung von RG1 angenommen.

Unter der Prämisse, dass bei komplexer Software niemals alle Abweichungen einer Software aufgedeckt werden können, kann die Anzahl von aufgefundenen Abweichungen sehr hoch werden. Jeder Fehlerbericht stellt potentiell ein Risiko für die Software selbst, die Betreiber, aber auch die Endnutzer dar. Es liegt nahe, dass ohne ausreichende Werkzeugunterstützung die Übersicht früher oder später verloren ginge. Um dem entgegen zu wirken, nutzen laut einer Umfrage des ISTQB¹⁹⁷ 80,6% der Teilnehmer ein sogenanntes „Defect Tracking Tool“. Die folgenden Aussagen lassen auf eine eindeutige Empfehlung des ISTQB schließen, ein solches Workflow-Werkzeug inkl. Prozess einzuführen:

„In order to manage all defects to resolution, an organization should establish a defect management process which includes a workflow and rules for classification. This process

¹⁹⁶ (Spillner und Linz 2019: 63–64)

¹⁹⁷ Aus einer Umfrage des ISTQB aus Jahr 2018 (Linetzki u. a. 2018: 32).

*must be agreed with all those participating in defect management, including designers, developers, testers, and product owners.*¹⁹⁸

Sowie:

„Die meisten Testorganisationen verwenden für das Management der Fehlerberichte im gesamten Fehlerlebenszyklus ein entsprechendes Werkzeug (Fehler-Workflow-Werkzeug).“¹⁹⁹

Mit Item 14 wird daher dieser Aspekt mit Sicht auf den Reifegrad abgedeckt.

Tabelle 3-16: Item Nr. 15

Item Nr. 15	
Kategorie:	Fehlermanagement
Frage:	Wann wird Ihrer Erfahrung nach innerhalb Ihrer Organisationseinheit Fehler bzw. Abweichungen oder Spezifikationslücken am ehesten entdeckt?
Antworten:	Anforderungsanalyse bzw. -prozess, Softwareentwicklung, Softwaretests, während einer Inbetriebnahme, während des Normalbetriebes, Andere
Einordnung:	Geschlossen, Einfachwahl
RGx	Keine Bewertung über den Reifegrad

Item 15 dient zur Einschätzung, in welcher Phase des Softwarelebenszyklus²⁰⁰ nach Ansicht der Teilnehmer Abweichungen oder Spezifikationslücken bzw. -fehler²⁰¹ in der Regel entdeckt werden. Spezifikationslücken sind Erscheinungen, die einer Fehlerwirkung gleichen, aber dieser spezifischen Fehlerwirkung keine Anforderung gegenübersteht²⁰². Eine weitere Formulierung wäre: Es wurde aus fachlicher Sicht bei der Anforderungserhebung „etwas vergessen“ oder die damalige Wissensbasis schloss diese

¹⁹⁸ (ISTQB u. a. 2018: 76)

¹⁹⁹ (ISTQB u. a. 2012: 61)

²⁰⁰ Grob angelehnt an Wasserfallmodell, V-Modell und SCRUM (Spillner und Linz 2019: 54–60).

²⁰¹ Siehe ein Beispiel der Lufthansa (Hoffmann 2013: 55–57).

²⁰² Siehe auch (Spillner u. a. 2014: 328–329).

Konstellation nicht mit ein. Ziel ist es, einen Anhaltspunkt darüber zu gewinnen, an welcher Stelle des Softwarelebenszyklus aktuell Fehler häufig entdeckt werden, um ggf. gezieltere Analysen zur Verbesserung einzuleiten.

„Dass das Prüfen umso effizienter und effektiver ist, je früher es im Softwareentwicklungsprozess angewendet wird (Shift-Left), ist weitgehend anerkannt.“²⁰³

Dieses Item kann in diesem sehr groben Detailgrad nicht zur Bewertung des Reifegrades herangezogen werden und dient als Trendeinschätzung zur Effektivität der Bestandsmaßnahmen.

3.5.4.5 Testwerkzeuge und Automatisierung

Testwerkzeuge und Automatisierung stellen im Bereich der Qualitätssicherung ein häufig ungenutztes Potential für effizientere und umfangreichere Testaktivitäten dar. Testwerkzeuge werden in folgende Klassifikationen unterschieden:

- *„Improve the efficiency of test activities by automating repetitive tasks or tasks that require significant resources when done manually (e.g., test execution, regression testing)*
- *Improve the efficiency of test activities by supporting manual test activities throughout the test process (see section 1.4)*
- *Improve the quality of test activities by allowing for more consistent testing and a higher level of defect reproducibility*
- *Automate activities that cannot be executed manually (e.g., large scale performance testing)*
- *Increase reliability of testing (e.g., by automating large data comparisons or simulating behavior)²⁰⁴*

Die Einführung von Testautomatisierung führt auch immer zur Einführung eines Testwerkzeugs. Diese Werkzeuge stellen im Grunde Programme dar, welche mit dem Ziel entwickelt oder verwendet werden, das Testobjekt nach einem spezifischen Vorgehen automatisiert anzusprechen. Solche Tests können auf unterschiedlichen Wegen die Software ansprechen (z.B. durch Simulation von Maus- und Tastatureingaben, Ansprechen von technischen Zugängen/Schnittstellen). Testautomatisierung bringt eine Reihe von Vor- und Nachteilen mit²⁰⁵:

- Vorteile:

²⁰³ (Spillner u. a. 2017b: 16)

²⁰⁴ (ISTQB u. a. 2018: 79)

²⁰⁵ Auszug aus (ISTQB u. a. 2018: 81–82).

- Reduzierung von manuellen Tätigkeiten.
- Bessere Wiederholbarkeit von Tests.
- Objektivere Bewertung der Testausführung.
- Schnell viele Testkombinationen testbar.
- Einfache Anbindung von Informationssystemen zur Testüberwachung
- Nachteile:
 - Ggf. unrealistische Erwartungen an die Automatisierung.
 - Ggf. unterschätze Einführungskosten (in Aufwand, Zeit, oder Budget) und unterschätze Zeit bis zur Erreichung einer Rentabilität.
 - Wartungsaufwände.
 - Es könnte Werkzeugabhängigkeit entstehen.
 - Unklare Verantwortlichkeiten bezüglich des Automatisierungsvorhaben.

Tabelle 3-17: Item Nr. 16

Item Nr. 16	
Kategorie:	Testwerkzeuge und Automatisierung
Frage:	Wo und in welcher Form werden in Ihrer Organisationseinheit die Testfälle und Ergebnisse dokumentiert?
Antworten:	Stift und Papier (1), strukturiert - lokal auf verschiedenen Computern (z.B. als Dokument) (2), strukturiert - im Netzwerkspeicher (z.B. als Dokument) (3), in einem digitalen Ticketsystem oder vergleichbares (4), Unstrukturierte Form oder keine Dokumentation (1), Sonstige Form (2): Freitextangabe.
Einordnung:	Halboffen, Einfachauswahl
RGx	In Klammern in der Antwort angeben. Die Antwort „Sonstige Form“ wird gleichgesetzt mit der Antwort „Strukturiert, lokal auf dem PC“.

In ISO 29119-3 werden eine Reihe von Dokumentationsformen vorgestellt²⁰⁶. Unter anderem gehören Teststrategie und Testkonzept zu diesem. In Item 1 wurde auf die gesamte Beschreibung der Testprozesse abgezielt. Mit diesem Item wird konkret auf den Testabschluss und die dort enthaltene Testfalldokumentation und deren Werkzeugnutzung eingegangen. Diese versteht sich daher als Teil des Merkmals Testwerkzeuge und Automatisierung²⁰⁷. Es wird überprüft, ob und in welcher Form eine Dokumentation darüber vorliegt und ob das Verständnis bzw. Wissen darüber bei den Teilnehmern einheitlich ist. Durch die halboffene Fragestellung lassen sich zusätzlich ggf. bisher unbekannte oder uneinheitliche Dokumentationsformen oder Werkzeuge aufdecken.

²⁰⁶ Eine Übersicht ist in Abbildung 5-2 im Werk (Daigl und Glunz 2016: 87) zu finden.

²⁰⁷ Erläutert in (ISTQB u. a. 2012: 16–17) sowie (ISTQB u. a. 2018: 79–80).

Tabelle 3-18: Item Nr. 17

Item Nr. 17	
Kategorie:	Testwerkzeuge und Automatisierung
Frage:	Wie hoch ist Ihrer Meinung nach aktuell der Anteil an automatisierten Testfällen innerhalb Ihrer Organisationseinheit? (Bitte schätzen Sie)
Antworten:	0% (1), 1% - 10% (1), 11% - 20% (2), 21% - 30% (3), 31% - 40% (3), 41% - 50% (4), Mehr als 50% (4)
Einordnung:	Geschlossen, Einfachauswahl.
RGx	Angabe in Klammern hinter der möglichen Antwort.

Laut der ISTQB Umfrage von 2017-2018 sehen 63,7% der Befragten Testautomatisierung als Trendthema der nahen Zukunft an. Aus diesem Grund wird die Fragestellung:

„What is the percentage of automated test cases you use with respect to your overall test cases?“²⁰⁸

für Item 17 daraus übernommen. Die Syllabus der ISTQB geben keine konkrete Vorgabe für einen zielführenden Anteil an Testautomatisierung an. Daher wird sich für die Bewertung auf den Reifegrad an den Ergebnissen der Umfrage orientiert²⁰⁸.

Tabelle 3-19: Item Nr. 18

Item Nr. 18	
Kategorie:	Testwerkzeuge und Automatisierung
Frage:	Welche Testaktivitäten sind innerhalb Ihrer Organisationseinheit Ihres Wissens nach bereits automatisiert?
Antworten:	Datensicherungstests, Kompatibilitätstests, Katastrophenbewältigungstest (z.B. bei Ausfall), Funktionstests, Installationstest, Integrationstests, Lasttests / Stresstests, Zugänglichkeit- / Benutzbarkeitstests, Ablauftests, Sicherheitstests, Abnahmetests, Systemtests / Systemintegrationstests, Komponententests, White-Box-Tests (z.B. Pfad-, Entscheidungs-, Zustandsbasiert), Erfahrungsbasierte Tests, Explorative Tests, Use-Case-Tests, Code-Reviews, Akzeptanztests, Regressionstests, Session-based testing, Pair testing. Alle Antwortmöglichkeiten werden in folgenden Ausprägungen auswählbar sein: Automatisiert, Nicht-Automatisiert. Zusätzlich wird den Teilnehmer folgende Auswahlmöglichkeit zur Verfügung gestellt: Keine

²⁰⁸ (Linetzki u. a. 2018: 33)

	der genannten Testaktivitäten wird für Tests verwendet und daher auch nicht für die Automatisierung berücksichtigt (1).
Einordnung:	Geschlossen, Mehrfachauswahl mit zwei Ausprägungen
RGx	Jede gewählte Ausprägung wird mit 0 (Nicht-Automatisiert) oder 1 (Automatisiert) bewertet und summiert. Wurde keine oder mehr als eine Ausprägung pro Antwort ausgewählt zählt, diese nicht. Die Antworten werden daraufhin summiert. Die Werteskala ist wie folgt: 0-3 -> RG1, 4-5 -> RG2, 5-10 -> RG3, 11-20 -> RG4. Bei einer Sonderantwortmöglichkeit ist die Einstufung in das Reifegradmodell in Klammern angegeben.

Dieses Item dient dazu, ein konkretes Bild über die automatisierten Testaktivitäten zu erhalten. Dazu wird die Liste der Testaktivitäten aus Item 2²⁰⁹ übernommen. Dies ermöglicht bei Bedarf eine kombinierte Interpretation der Items. Die Möglichkeit einer Einführung und Wartung der Testautomatisierung variiert je nach Testobjekt und Ressourcenverfügbarkeit und kann daher nicht als selbstverständlich angesehen werden. Nach Umfrage des ISTQB ist die Testautomatisierung die Nummer eins der Top-Herausforderungen in agilen Projekten²¹⁰. Umfrageergebnisse der GTB²¹¹ weisen ebenfalls darauf hin. Daraus lässt sich ableiten, dass die wenigsten Unternehmen eine 100% Automatisierung ihrer Testaktivitäten erreichen bzw. anstreben. Das bedeutet, es kann für das Item 18 nicht die gleiche Einstufung bezüglich des Reifegrades herangezogen werden wie für Item 2. Zur Kompensation wird die Höchstwertung von RG4 bereits bei der Hälfte der maximalen Auswahlmöglichkeiten erreicht sein. Dies wäre nach Auffassung des Verfassers gerade in Bezug zum öffentlichen Sektor eine verhältnismäßige Bewertung.

Tabelle 3-20: Item Nr. 19

Item Nr. 19	
Kategorie:	Testwerkzeuge und Automatisierung
Frage:	Bestehen Pläne dazu, über das aktuelle Maß hinaus, Testautomatisierungsaktivitäten einzuführen?
Antworten:	Nein oder Ja
Einordnung:	Geschlossen, Einfachauswahl
RGx	Keine Bewertung über den Reifegrad

Dieses Item dient dazu zu erkennen, ob innerhalb der Organisationseinheit ein Interesse

²⁰⁹ Siehe Kapitel 3.5.4.1.

²¹⁰ Siehe dazu (Linetzki u. a. 2018: 22)

²¹¹ Siehe dazu (Spillner u. a. 2017b: 14–16)

oder Pläne für eine mögliche Einführung oder Verbesserung der Testautomatisierung bestehen. Diese Erkenntnis kann in Überlegungen für weitere Analysen oder Maßnahmen einbezogen werden. Sollte zum Beispiel durchweg in einer Organisationseinheit kein Interesse zur Einführung von Testautomatisierung existieren, so kann erwartet werden, dass eine Einführung mit deutlichen Hindernissen verbunden sein könnte. Oder aber es wird sich zu Beginn auf anderen Maßnahmen konzentriert, die ggf. mehr Zuspruch besitzen. Da es dieses Item nicht die tatsächliche Situation, sondern einen möglichen Entwicklungsraum und Stimmung der Organisationseinheit abfragt, kann es nicht zur Bewertung des Reifegrades genutzt werden.

3.5.4.6 Soziale Kompetenz und Teambzusammensetzung

Das Merkmal kann wie folgt beschrieben werden:

„Testmanager sind zuständig für das Rekrutieren von Mitarbeitern sowie für das Erhalten und Betreuen von Testteams, in denen die richtigen Fähigkeiten vertreten sind. Die benötigten Fähigkeiten können sich im Laufe der Zeit ändern. Es ist daher wichtig, nicht nur zu Beginn die richtigen Mitarbeiter zu rekrutieren, sondern auch für geeignete Fortbildungs- und Entwicklungsmöglichkeiten zu sorgen, um das Testteam zu erhalten und die besten Leistungen zu fördern.“²¹²

Der Fokus dieses Merkmals liegt auf dem bestehenden Wissen und Fähigkeiten der Mitarbeiter einer Organisation. Idealerweise würde man ein vollständiges Skill-Assessment für die Mitarbeiter einer Unternehmung durchführen, dieses wäre aber deutlich umfangreicher und würde nicht der Zielrichtung des hier entwickelten Verfahrens entsprechen. Ebenfalls gehört der Aspekt des Recruiting und die Einarbeitung von neuen Mitarbeitern für Testaufgaben dazu²¹³. Infolge dessen sollte man auch das Ausscheiden von Mitarbeitern berücksichtigen. Bekannte Modelle berücksichtigen ebenfalls Wissen, Fähigkeiten und Fortbildungen. Ein Beispiel wäre TMMi²¹⁴.

Tabelle 3-21: Item Nr. 20

Item Nr. 20	
Kategorie:	Soziale Kompetenz und Teambzusammensetzung
Frage:	Existiert für Ihre Organisationseinheit ein Einarbeitungskonzept bzw. -plan für neue Mitarbeiter mit Testaufgaben?

²¹² (ISTQB u. a. 2012: 81)

²¹³ Siehe dazu (ISTQB u. a. 2012: 83–84).

²¹⁴ Siehe dazu GP 2.5 „Test Training Programm“ und SP 4.2 „Plan for test staffing“ des TMMi (van Veenendaal 2018).

Antworten:	Nein (1) oder Ja (4)
Einordnung:	Geschlossen, Einfachauswahl
RGx	Angabe in Klammern hinter der möglichen Antwort.

Tabelle 3-22: Item Nr. 21

Item Nr. 21	
Kategorie:	Soziale Kompetenz und Teamzusammensetzung
Frage:	Existiert ein Ausgliederungskonzept (z.B. Übergabeprotokolle, Wissensdokumentation) für Mitarbeiter mit Testaufgaben die Ihre Organisationseinheit geplant verlassen?
Antworten:	Nein (1) oder Ja (4)
Einordnung:	Geschlossen, Einfachauswahl
RGx	Angabe in Klammern hinter der möglichen Antwort.

Item 20 und Item 21 decken die Themen Einarbeitungskonzept und Ausgliederungskonzept für das On- bzw. Offboarding von Mitarbeitern, die Tätigkeiten ausführen oder Rollen im Rahmen der Qualitätssicherung besitzen, ab. Für die Ausführung von Tätigkeiten sind bestimmte Kenntnisse über die Organisation, die Entwicklungsprozesse, testbezogene Kenntnisse (z.B. Testkonzept, Teststrategie, Testprozesse, Testplanung)²¹⁵ notwendig. Diese effizient einem neuen Mitarbeiter zu vermitteln, trägt indirekt zum Reifegrad bei:

„Neue Testteammitglieder müssen schnell vom Testteam integriert werden und eine angemessene Betreuung und Unterstützung erhalten. Jede Person im Testteam sollte ihre definierte Rolle haben, die sich aus einer individuellen Beurteilung ergeben kann. Ziel ist es, dass jedes Mitglied als einzelne Person erfolgreich ist und gleichzeitig zum Erfolg des gesamten Testteams beiträgt.“²¹⁶

Für den Fall, dass ein Mitarbeiter die Organisationseinheit verlässt, sollte ebenfalls ein Konzept existieren. Hier könnten beispielsweise Regelungen bezüglich Aufgaben- bzw. Rollenübergabe definiert werden, aber auch testbezogene Themen wie zum Beispiel die Übergabe von Lizenzen für Testtools, Deaktivierung von Testzugängen, Übergabe von: Testdaten, Testkonfigurationen, etc. Beide Punkte tragen zu einem effizienteren Einstieg bzw. Ausstieg in die Organisationseinheit bei und somit auch zum Reifegrad der QS.

²¹⁵ Siehe auch (ISTQB u. a. 2012: 81–84).

²¹⁶ (ISTQB u. a. 2012: 83)

Tabelle 3-23: Item Nr. 22

Item Nr. 22	
Kategorie:	Soziale Kompetenz und Teamzusammensetzung
Frage:	Werden innerhalb Ihrer Organisationseinheit Mitarbeitern regelmäßig die Teilnahme an Fortbildungsmaßnahmen oder Tagungen zum Thema Softwarequalitätssicherung ermöglicht?
Antworten:	Nein (1) oder Ja (4)
Einordnung:	Geschlossen, Einfachauswahl
RGx	Angabe in Klammern hinter der möglichen Antwort.

Die Möglichkeit neues Wissen und Fähigkeiten in die Organisationseinheit aufzunehmen kann, ein wichtiger Teil für die Weiterentwicklung von Mitarbeitern und Organisation darstellen. Mitarbeiterentwicklung ist in der Regel unternehmensweit geregelt.

„Wenn Testmanager in die Tester investieren und ein Umfeld für kontinuierliches Lernen schaffen, dann werden die Testteammitglieder dadurch motiviert, ihre Fähigkeiten und ihr Wissen zu steigern, um für die nächste Aufgabe vorbereitet zu sein“²¹⁷

Das Thema Softwarequalitätssicherung sollte demnach hier berücksichtigt werden. Aktuelle Vorgehensweisen, Standards, Zertifizierungen, neue Werkzeuge und Vernetzung bzw. Austausch mit anderen Fachkollegen sollten sich grundsätzlich positiv auf den Reifegrad auswirken. Laut einer Umfrage²¹⁸ des ISTQB nutzen nach interner Einarbeitung 50,2% der Unternehmen Zertifizierung und 27,2% der Unternehmen Konferenzen zur weiteren Fortbildung von Mitarbeitern im Bereich QS.

3.5.5 Pre-Test des Fragebogens

Es ist eine übliche Vorgehensweise, einen Fragebogen vor dem eigentlichen Einsatz, einem sogenannten Pre-Test zu unterziehen:

*„Die Konstruktion des standardisierten Fragebogeninstruments erfolgt in zwei Schritten, zunächst als Grob-, dann als Feinkonzeption. Der konstruierte Fragebogen wird schließlich einem Fragebogen-Pretest unterzogen und bei Bedarf überarbeitet (Fragebogenre-
vision).“²¹⁹*

²¹⁷ (ISTQB u. a. 2012: 82)

²¹⁸ Siehe dazu (Linetzki u. a. 2018: 12).

²¹⁹ (Döring und Bortz 2016: 405)

Das entspricht den Schritten 9 – 10 nach Kallus²²⁰, allerdings in sehr vereinfachter Form. Nach abgeschlossener Feinkonzeptionierung des Fragebogens wurde demnach offener nicht-strukturierter Pre-Test (N=3) durchgeführt. Dieser setzte sich in seiner Stichprobe wie folgt zusammen:

- unabhängige fachfremde Personen (N=2).
- unabhängige fachkundige Personen (N=1).

Der Pre-Test stellt nach den Design-Science Prinzipien ebenfalls einen Teil des Prinzips 3 „Design Evaluation“ und 5 „Research Rigor“ dar. Ziel dieses Pre-Tests war es, den Fragebogen auf seine Plausibilität und Verständlichkeit bezüglich der Fragestellungen und Antwortmöglichkeiten hin zu überprüfen. Folgende Erkenntnisse konnten aus den Ergebnissen des Pre-Tests gezogen werden:

- Es bestehen einige Rechtschreibfehler und Grammatikfehler.
- Einige Items waren nicht vollständig ohne Hilfe verstanden worden.
- Es fehlte Items die Möglichkeit, „keine Antwort“ auszuwählen.

Die längste Ausfüllzeit innerhalb des Pre-Tests war mit 25 Minuten im geplanten Rahmen. Die Lesezeit für Einleitung und Beiblatt lag mit knapp sieben Minuten ebenfalls im zuvor definierten Rahmen. Mit also insgesamt 32 Minuten ist im Vergleich zu den ursprünglich veranschlagten 40 Minuten ein Zeitpuffer von acht Minuten zu erwarten. Die kleine Stichprobe ermöglicht allerdings nur eine begrenzte Verwertbarkeit der Ergebnisse. Weitere Schlüsse können ggf. aus der Erprobung des Verfahrens abgeleitet werden. Der vollständig konzipierte Fragebogen befindet sich in Anhang 2.

3.6 Beschreibung der Reifegradstufen

Etablierte Reifegradmessungsverfahren nutzen Beschreibungen, die die Reifegradeinstufung inhaltlich verschriftlichen. In Kapitel 2.2.1 wurde dazu bereits die Beschreibungen von TPI/TPI-Next zitiert. Im direkten Vergleich zu TPI/TPI-Next weist CMMI/TMMi deutlich längere Beschreibungen²²¹ seiner Reifegrade auf. Durch die Anlehnung an TPI/TPI-Next wird auch für diese Verfahren der Ansatz einer zusammenfassenden Kurzbeschreibung gewählt. Die Gesamtaussage der Reifegradstufe leitet sich, wie in Kapitel 3.4 beschrieben, von den Ergebnissen der einzelnen Merkmale ab. Jedes Merkmal kann gemäß Modell in einer der vier Reifegradstufen bewertet werden. Eine Merkmalsbewertung kann sich inhaltlich nur auf die dazugehörigen Items beziehen, da eine Aussage

²²⁰ Siehe (Kallus 2016: 14–15).

²²¹ Einzusehen in Kapitel 2 (van Veenendaal 2018: 9–12).

über die Inhalte der Items hinaus nicht getätigt werden kann. Daher können nur Annäherungen zur tatsächlichen Situation in der Organisationseinheit beschrieben werden. Es gilt also zu beachten, dass hier der Versuch unternommen wird, eine möglichst treffende Beschreibung der Realität zu finden. Besonders in Extremfällen, in denen die Antworten einzelner Items innerhalb eines Merkmals weit auseinander liegen, könnten die hier definierten Beschreibungen nicht vollständig zutreffen. Der Grund dafür ist die bewusste Begrenzung auf eine geringe Anzahl an Fragestellungen bzw. Items. Es empfiehlt sich daher ebenfalls eine Detailanalyse der Items unter Berücksichtigung der Freitextfelder, durchzuführen welche den konkreten Kontext der Organisationseinheit berücksichtigt und somit spezifischere Ergebnisse abgeleitet werden können.

Für die unterschiedlichen Reifegrade werden für jede Stufe annähernd die gleichen oder vergleichbaren Inhalte beschrieben, aber in unterschiedlichen Abstufungen, die sich mittels Schlüsselwörter in ihrer Wertigkeit unterscheiden. Folgende Wörter werden dazu in den jeweiligen Stufen verwendet:

- Initial: vereinzelt, wenigen, unstrukturiert, nicht, kein/keine, uneinheitlich, ohne
- Mittel: einige, ggf., ein paar,
- Gut: viele, die meisten, wahrscheinlich
- Sehr gut: sehr viele, alle, grundsätzlich, einheitlich, breite

Auf diese Weise kann über den ermittelten Reifegrad des Merkmales ein Satzbaustein entnommen werden. Die einzelnen Satzbausteine der Merkmale setzen sich letztendlich zum Gesamtreifegradergebnis zusammen.

3.6.1 Testprozess (Item 1 - 3)

1. Initial

Es finden vereinzelt unterschiedliche Testaktivitäten statt. Die wenigen Tests, die stattfinden, sind eher unstrukturiert. Nicht-Funktionale Anforderungen werden nicht im Anforderungsprozess berücksichtigt.

2. Mittel

Es finden einige unterschiedliche Testaktivitäten statt. Wenn passend dazu ein Testkonzept besteht, sind diese strukturiert. Nicht-Funktionale Anforderungen werden ggf. im Anforderungsprozess berücksichtigt.

3. Gut

Es finden viele unterschiedliche Testaktivitäten statt. Wenn passend dazu ein Testkonzept besteht, sind diese strukturiert. Nicht-Funktionale Anforderungen werden wahrscheinlich im Anforderungsprozess berücksichtigt.

4. Sehr gut

Es finden sehr viele unterschiedliche Testaktivitäten statt die strukturiert sind. Nicht-Funktionale Anforderungen werden im Anforderungsprozess berücksichtigt.

3.6.2 Testmanagement (Item 4 – 8)

1. Initial

Die Qualitätsmerkmale werden nur vereinzelt berücksichtigt. Die Anliegen und Anforderungen der Stakeholder werden im Rahmen der Qualitätssicherung grundsätzlich nicht berücksichtigt. Testmetriken werden nicht erhoben.

2. Mittel

Es werden ein paar Qualitätsmerkmale berücksichtigt. Die Anliegen und Anforderungen der Stakeholder werden im Rahmen der Qualitätssicherung vereinzelt berücksichtigt. Testmetriken werden ggf. erhoben.

3. Gut

Es werden die meisten Qualitätsmerkmale berücksichtigt. Die Anliegen und Anforderungen der Stakeholder werden im Rahmen der Qualitätssicherung berücksichtigt. Testmetriken finden meistens Anwendung.

4. Sehr gut

Es werden alle oder beinahe alle Qualitätsmerkmale berücksichtigt. Die Anliegen und Anforderungen der Stakeholder werden im Rahmen der Qualitätssicherung grundsätzlich berücksichtigt. Testmetriken finden eine grundsätzlich breite Anwendung.

3.6.3 Reviews (Item 9 - 11)

1. Initial

Es werden nur vereinzelte Reviews durchgeführt. Die wenigen, die durchgeführt werden sind uneinheitlich. Tester werden im Anforderungsprozess nicht berücksichtigt.

2. Mittel

Es werden ein paar Reviews durchgeführt. Werden dazu Checklisten eingesetzt, sind diese einheitlich. Tester werden im Anforderungsprozess ggf. berücksichtigt.

3. Gut

Es werden viele Reviews durchgeführt. Werden dazu Checklisten eingesetzt, sind diese einheitlich. Tester werden wahrscheinlich am Anforderungsprozess berücksichtigt.

4. Sehr gut

Es werden alle oder beinahe alle Reviews durchgeführt. Diese werden einheitlich durchgeführt. Tester werden am Anforderungsreviewprozess berücksichtigt.

3.6.4 Fehlermanagement (Item 12 – 15)

1. Initial

Der Prozess des Abweichungsmanagements ist unstrukturiert. Ebenfalls ist die Dokumentation der Abweichungen uneinheitlich beschrieben. Auch wird kein Fehler-Workflow-Werkzeug eingesetzt.

2. Mittel

Es findet ggf. ein strukturierter Prozess für Abweichungen statt. Möglicherweise sind die Abweichungen einheitlich beschrieben. Vielleicht wird zur Unterstützung ein Fehler-Workflow-Werkzeug eingesetzt.

3. Gut

Es ist wahrscheinlich, dass ein strukturierter Prozess für Abweichungen stattfindet, in denen die Abweichungen einheitlich beschrieben werden. Wohlmöglich wird zur Unterstützung ein Fehler-Workflow-Werkzeug eingesetzt.

4. Sehr gut

Es findet ein strukturierter Prozess für Abweichungen statt und diese sind einheitlich beschrieben. Ebenfalls wird ein Fehler-Workflow-Werkzeug eingesetzt.

3.6.5 Testwerkzeuge und Automatisierung (Item 16 – 19)

1. Initial

Testaktivitäten werden ohne Werkzeugunterstützung dokumentiert. Sollte eine Testautomatisierung bestehen, sind nur wenige der Testfälle und unterschiedliche Testaktivitäten einbezogen.

2. Mittel

Testaktivitäten werden ggf. mit Werkzeugunterstützung dokumentiert. Sollte eine Testautomatisierung bestehen, sind einige Testfälle und ein paar unterschiedliche Testaktivitäten einbezogen.

3. Gut

Testaktivitäten werden wahrscheinlich mit Werkzeugunterstützung dokumentiert. Sollte eine Testautomatisierung bestehen, sind viele Testfälle und auch unterschiedliche Testaktivitäten einbezogen.

4. Sehr gut

Testaktivitäten werden mit Werkzeugunterstützung dokumentiert. Es sind mindestens 41% der Testfälle automatisiert in denen sehr viele unterschiedliche Testaktivitäten berücksichtigt werden.

3.6.6 Soziale Kompetenz und Teamzusammensetzung (Item 20 – 22)

1. Initial

Es existiert im Testbereich kein Einarbeitungs- und Ausgliederungskonzept. Auch wird keine Teilnahme an QS-Fortbildungsmaßnahmen oder QS-Tagungen ermöglicht.

2. Mittel

Im Testbereich existiert ggf. ein Einarbeitungs- und/oder Ausgliederungskonzept. Vereinzelt könnte die Teilnahme an QS-Fortbildungsmaßnahmen oder QS-Tagungen möglich sein.

3. Gut

Im Testbereich existiert ein Einarbeitungs- und/oder Ausgliederungskonzept. Wahrscheinlich wird die Teilnahme an QS-Fortbildungsmaßnahmen oder QS-Tagungen ermöglicht.

4. Sehr gut

Im Testbereich existiert ein Einarbeitungs- und Ausgliederungskonzept. Auch wird die Teilnahme an QS-Fortbildungsmaßnahmen oder QS-Tagungen ermöglicht.

3.7 Abschluss der Artefakt-Entwicklung

Abschließend für das Kapitel 3 betrachtet wurden folgende Artefakt erzeugt:

- Reifegradmodell zur Softwarequalitätssicherung (Siehe Kapitel 3.4).
- Standardisierte Reifegradbeschreibungen auf Ebene des Merkmals (Siehe Kapitel 3.6).
- Fragebogen inkl. Detailumsetzung des Reifegradmodells. Konkret für eine Anwendung in Organisationseinheiten des öffentlichen Sektors und zur freien Verfügung für weitere Entwicklung bzw. Ergänzung (Siehe Anhang 2).
- Zum Fragebogen wurde ebenfalls eine Exceldatei entwickelt, welche die Auswertung der Datensätze unterstützt (Siehe Anhang 30).

Es folgt in den nächsten Kapiteln eine Erprobung sowie eine Validierung der Artefakte.

4 Erprobung des entwickelten Werkzeugs

Nach abgeschlossener Entwicklung, Gestaltung und Pre-Test des Verfahrens bzw. Werkzeuges steht der Einsatz des Verfahrens an. Ziel ist es, das Verfahren im Bereich des Bundesverwaltungsamtes innerhalb eines Softwareentwicklungsverfahrens zur Anwendung zu bringen und Erkenntnisse für eine zukünftige Weiterentwicklung des Werkzeuges bzw. Verfahrens zu gewinnen. Dieses Kapitel deckt die Design-Science Prinzipien 3 „Design Evaluation“, 5 „Research Rigor“ und einen Anteil von 7 „Communication of Research“ ab. Auch deckt dieses Kapitel die Schritte 9-10 nach Kallus²²² ab. Der fehlende Schritte 11 findet keine konkrete Anwendung und Schritt 12 wurde in Teilen mit dem entwickelten Auswertungsartefakt nachgekommen. Schritt 13 wird im begrenzten Rahmen mit dem folgenden Kapitel 5 abgedeckt.

4.1 Auswahl des Untersuchungsgegenstandes

Auch wenn es keine Voraussetzung darstellt, sollte ein Softwareentwicklungsverfahren im Idealfall die Zielsetzung besitzen ihre QS-Prozesse verbessern zu wollen. Im ersten Schritt wurden verschiedene Verantwortliche von Entwicklungsverfahren des BVA telefonisch kontaktiert. Zwei Verfahren zeigten daraufhin Interesse und baten um eine kurze schriftliche Beschreibung sowie die Unterlagen zur Befragung (Fragebogen und Beiblatt). Die versendete Mail mit der Beschreibung befindet sich in Anhang 3, als Dateianhang lagen das Beiblatt (siehe Anhang 1) und eine nicht ausgefüllte Ausfertigung des Fragebogens (siehe Anhang 2) bei. Das erste Verfahren entschied sich daraufhin bewusst gegen eine Teilnahme. Das zweite Verfahren stimmte einer Erprobung zu. Nach der Zustimmung zur Teilnahme wurde eine Liste mit acht E-Mailadressen möglicher Teilnehmer zur Verfügung gestellt. Allen acht möglichen Teilnehmer wurden einzeln eine Einladung (siehe Anhang 4) zur Teilnahme zugesandt. Alle versandten Fragebögen wurden mit einer eindeutige ID (z.B. MT130501) versehen. Auf diese Weise werden doppelte Einreichungen vermieden. Gleichzeitig wurde sichergestellt, dass nur die Bögen in die Auswertung einfließen, die auch ursprünglich verteilt worden sind. Zur Sicherung der Anonymität wurden die Bögen in einer beliebigen Reihenfolge an die Teilnehmer versendet. Ebenfalls wurde bewusst keine Kopie bei Versendung der E-Mail in der E-Mail Anwendung angelegt. Den Teilnehmern wurden drei Werkzeuge Zeit für eine Einsendung eingeräumt. Der Zeitraum musste bewusst auf diese Zeit verkürzt werden, damit die Ergebnisanalyse und eine Validierung zur fristgerechten Einreichung dieser Masterthesis abgeschlossen sind. Grundsätzlich sollte den Teilnehmern für das Verfahren aber mehr Zeit eingeräumt werden.

²²² (Kallus 2016: 14–15)

4.2 Abschluss und Dokumentation der Ergebnisse

Bis zum Abschluss der Frist wurden von acht Fragebögen sieben ausgefüllt und via E-Mail zurückgesendet. Nach Empfang einer Rücksendung wurde der Anhang von der E-Mail getrennt und abgespeichert. Damit keine weitere Rückverfolgbarkeit auf den Versender möglich ist, wurde nach erfolgreicher Speicherung des Anhangs die E-Mail innerhalb der Anwendung gelöscht. Damit ist die Datenerhebung abgeschlossen. Ein ausgefüllter Fragebogen gilt nun als Datensatz. Die Ergebnisse wurden in einer Exceltabelle dokumentiert und werden dort auch ausgewertet. Die Tabellenseiten befinden sich in Anhang 5 bis Anhang 27. Jedes Item erhielt eine Seite, in der alle Datensätze und die zugehörigen Antworten vermerkt wurden. Mögliche Freitextangaben wurden ebenfalls übernommen. Auch ist das zuvor definierte Mapping der Antwortmöglichkeiten auf die Reifegradskala dargestellt. Mittels der Skala ermittelt die Tabelle pro Befragten den Reifegrad eines Items/Themenpunkt und dann im Anschluss über alle Befragten hinweg im Mittel. Damit ist die Reifegradbewertung eines Items abgeschlossen. Die Ergebnisse aller Items wurden auf einer Ergebnisseite zusammengetragen. Bei der Entwicklung des Verfahrens wurde zur Sicherung der Anonymität eine Stichprobe von acht empfohlen. Leider konnte für die Erprobung nur eine Stichprobe von sieben erreicht werden. In diesem konkreten Fall ergaben sich daraus keine Einbußen in der Gewährleistung der Anonymität. Grundsätzlich gilt die Empfehlung nach Auffassung des Verfassers weiterhin und auch, wenn im Einzelfall davon abgewichen werden kann.

4.3 Ergebnisse der Reifegrad Bewertung

Nachdem hier entwickelten Reifegradmodell wird das Endergebnis mit einer Reifegradstufe 3 bewertet. Die genaue Zusammensetzung und Aufstellung befinden sich in Tabelle 4-1 und Anhang 5.

Tabelle 4-1: Ergebnis der Reifegradbewertung

Item	RG	Merkmale	RG
1	4	Testprozess	3
2	3		
3	1		
4	3	Testmanagement	3
5	Kein		
6	Kein		
7	3		
8	3		
9	3	Reviews	2
10	1		

11	3		
12	4	Fehlermanagement	4
13	4		
14	4		
15	Kein		
16	4		
17	1	Testwerkzeuge und Auto- matisierung	2
18	1		
19	Kein		
20	2	Soziale Kompetenz und Teamzusammensetzung	2
21	3		
22	1		
Gesamt RG:		3	

Das beschreibende Ergebnis mittels der entwickelten Satzbausteine lautet wie folgt:

Es finden viele unterschiedliche Testaktivitäten statt. Wenn passend dazu ein Testkonzept besteht, sind diese strukturiert. Nicht-Funktionale Anforderungen werden wahrscheinlich im Anforderungsprozess berücksichtigt. Es werden die meisten Qualitätsmerkmale berücksichtigt. Die Anliegen und Anforderungen der Stakeholder werden im Rahmen der Qualitätssicherung berücksichtigt. Testmetriken finden meistens Anwendung. Es werden ein paar Reviews durchgeführt. Werden dazu Checklisten eingesetzt, sind diese einheitlich. Tester werden im Anforderungsprozess ggf. berücksichtigt. Es findet ein strukturierter Prozess für Abweichungen statt und diese sind einheitlich beschrieben. Ebenfalls wird ein Fehler-Workflow-Werkzeug eingesetzt. Testaktivitäten werden ggf. mit Werkzeugunterstützung dokumentiert. Sollte eine Testautomatisierung bestehen, sind einige Testfälle und ein paar unterschiedliche Testaktivitäten einbezogen. Im Testbereich existiert ggf. ein Einarbeitungs- und/oder Ausgliederungskonzept. Vereinzelt könnte die Teilnahme an QS-Fortbildungsmaßnahmen oder QS-Tagungen möglich sein.

4.4 Analyse der Datenbasis

Die Reifegradbewertung stellt nach dem entwickelten Reifegradmodell eine eigenständige Aussage dar. In diesem Kapitel findet die Überprüfung auf Nachvollziehbarkeit dieser einzelnen Reifegradbewertungen statt. Grundsätzlich wird eine Detailanalyse der Ergebnisse empfohlen, da diese wichtigen Erkenntnisse hervorbringen kann, welche für die Ausprägung und das Ziel möglicher Verbesserungsmaßnahmen genutzt werden können. In diesem Kapitel wurden eine solche Analyse und die Überprüfung gemeinsam durchgeführt.

Im Weiteren folgt eine kurze Zusammenfassung der Ergebnisse eines jeden Merkmales. Daraufhin werden die Items einer Analyse im Detail unterzogen. Anschließend wird auf verschiedene mögliche Kombinationen zwischen den Items eingegangen und Erkenntnissen daraus abgeleitet.

4.4.1 Merkmal: Testprozess (Item 1 – 3)

Das Merkmal Testprozess wurde insgesamt, also über die Antworten aller Befragten hinweg, mit RG3 bewertet. Auch wenn im Bereich der nicht-funktionalen Anforderungen Verbesserungsbedarf besteht, scheinen die Struktur und die Anzahl der genutzten Testaktivitäten das Ergebnis zu bestätigen.

4.4.1.1 Item 1

(Die Datenbasis befindet sich in Anhang 6)

Alle Teilnehmer stimmten zu diesem Item mit „Ja“ ab, was zu einer Bewertung von RG4 für dieses Item führt. Das führt direkt zum Schluss, dass innerhalb der Organisationseinheit eine konkrete Beschreibung über die eingesetzten Testprozesse zur Verfügung steht und dass alle Befragten diese kennen. Die verschiedenen Freitextangaben geben aber Hinweise darauf, wie und wo eine weitere Analyse erfolgen könnte. In diesem Fall wird ein Testkonzept benannt. Da könnte für ein weiteres Vorgehen ein Abgleich dieses Testkonzepts mit den hier vorgestellten Ergebnissen erfolgen. Die RG Einstufung scheint daher soweit stimmig zu sein.

4.4.1.2 Item 2

(Die Datenbasis befindet sich in Anhang 7)

Dieses Item wurde von vielen Teilnehmern ähnlich ausgefüllt, allerdings mit einem Ausreißer. Insgesamt führte das zu einer Bewertung von RG3. Das lässt den Schluss zu, dass mit einer Ausnahme, grundsätzlich das gleiche Verständnis über die eingesetzten Methoden besteht. Auch lässt sich aus den Daten ableiten, dass eine Vielzahl von Testmethoden den Teilnehmern nicht bekannt ist. Einige Methoden, die als „Unbekannt“ angegeben worden sind, wurden von anderen Teilnehmern als „Eingesetzt“ markiert. Siehe dazu Tabelle 4-2 und die als Orange markierten Zeilen.

Tabelle 4-2: Ergebnis Item 2

	Unbekannt	Bekannt	Eingesetzt
Datensicherungstests	0	6	1
Kompatibilitätstests	4	2	1

Katastrophenbewältigungstest	5	2	0
Funktionstests	0	0	7
Integrationstests	1	2	4
Lasttests / Stresstests	0	5	2
Zugänglichkeit- / Benutzbarkeitstests	0	2	5
Ablauftests	0	0	7
Sicherheitstests	4	2	1
Abnahmetests	0	0	7
Systemtests / Systemintegrationstests	3	4	0
Komponententests	0	2	5
White-Box-Tests	5	2	0
Erfahrungsbasierte Tests	0	0	7
Explorative Tests	1	0	6
Use-Case-Tests	1	0	6
Code-Reviews	5	2	0
Akzeptanztests	4	2	1
Regressionstests	0	1	6
Session-based testing	3	0	4
Pair testing	7	0	0

Das lässt auf Wissenslücken schließen und dass aufgrund dessen ein falsches Verständnis über die eingesetzten Methoden vorliegt. Ebenfalls könnte das bedeuten, dass diese Testaktivitäten nur von einzelnen Personen innerhalb der Organisationseinheit durchgeführt werden. Da zur Einstufung die eingesetzten Testmethoden herangezogen worden sind, ist die Reifgradeinstufung nachvollziehbar.

4.4.1.3 Item 3

(Die Datenbasis befindet sich in Anhang 8)

Item 3 wurde bis auf einer Ausnahme mit „Nein“ beantwortet. Dies führt zu einer Bewertung mit RG1. Die Ergebnisse geben Grund zum Anlass, den Anforderungsprozess als ganzen und innerhalb dieses Prozesses die Berücksichtigung von nicht-funktionalen Anforderungen zu überprüfen. Die einheitliche Aussage lassen die Einstufung im Modell korrekt erscheinen.

4.4.2 Merkmal: Testmanagement (Item 4 – 8)

Das Merkmal Testmanagement wurde insgesamt mit RG3 bewertet. Diese Bewertung begründet sich auf der breiten Abdeckung und Anerkennung der QS-Merkmale, die vielseitige Einbeziehung von Stakeholdern und die Anwendung von verschiedenen Testmetriken.

4.4.2.1 Item 4

(Die Datenbasis befindet sich in Anhang 9)

Im Schnitt wurden fünf bis sechs der acht QS-Merkmale als abgedeckt ausgewählt. Dies führte zu einer Bewertung von RG3.

Tabelle 4-3: Ergebnis Item 4

	Nicht-Abgedeckt	Abgedeckt
Funktionale Eignung	0	7
Leistungsfähigkeit	2	5
Komptabilität	3	4
Benutzerfreundlichkeit	2	5
Verlässlichkeit	2	5
Sicherheit	5	2
Wartbarkeit	5	2
Portabilität	6	1

Tabelle 4-3 zeigt auf, dass vor allem bei den Merkmalen Sicherheit, Wartbarkeit und Portabilität noch Raum zur Verbesserung zu bestehen scheint. Die RG-Einstufung scheint daher soweit nachvollziehbar.

4.4.2.2 Item 5

(Die Datenbasis befindet sich in Anhang 10)

Item 5 wirkte sich nicht auf die Bewertung des Reifegrades aus, zeigt aber auf, dass unter den Teilnehmern, bis auf Portabilität die QS-Merkmale eine relevante Bedeutung haben. Auf einen Zusammenhang mit Item 4 wird im Verlauf eingegangen.

4.4.2.3 Item 6

(Die Datenbasis befindet sich in Anhang 11)

Dieses Item wurde nicht in die Reifegradbewertung mit einbezogen. Item 6 zeigt einstimmig auf, dass die Testaktivitäten von internen Mitarbeitern ausgeführt werden. Für die Abteilung, in der die Durchführung der Erhebung stattfand, stellt das keine Überraschung dar, da der Einsatz von externen Dienstleistern im Testbereich nicht stattfindet. Es bestätigt, dass der richtige Untersuchungsgegenstand für diese Befragung gewählt worden ist, da auch nur interne Mitarbeiter befragt worden sind.

4.4.2.4 Item 7

(Die Datenbasis befindet sich in Anhang 12)

Item 7 betrifft die Einbeziehung möglicher Stakeholder. Die Antworten der Befragten haben zu einer Bewertung RG3 geführt. Auffällig ist, dass alle Teilnehmer der Auffassung sind, dass weder „Entwickler“ noch „Architekten“ berücksichtigt werden. Das Freitextfeld wurde nicht verwendet, was den Schluss zulässt, dass aus Sicht der Teilnehmer keine Ergänzung der Antworten notwendig war.

4.4.2.5 Item 8

(Die Datenbasis befindet sich in Anhang 13)

In Item 8 wurde, mit einer Ausnahme, über alle Datensätze hinweg die gleiche Wahl getroffen. Es scheint für die Mehrheit der Befragten eindeutig zu sein, welche Testmetriken genutzt werden. Die Mehrheit wählte dabei alle Antwortmöglichkeiten außer „Codeüberdeckung“ und „Sonstige“. Dies führte zu einer stimmigen Bewertung von RG3. Das Freitextfeld wurde nicht verwendet, was den Schluss zulässt, dass aus Sicht der Teilnehmer keine Ergänzung der Antworten notwendig war.

4.4.3 Merkmal: Review (Item 9 – 11)

Das Merkmal Review wurde insgesamt mit RG2 bewertet. Reviews werden auf einer Vielzahl von Dokumenten bzw. Arbeitsergebnissen ausgeführt. Allerdings besteht Grund zur Annahme, dass dies nicht einheitlich geschieht. Ebenfalls besteht Unsicherheit darüber, ob Tester ausreichend im Anforderungsprozess miteinbezogen werden.

4.4.3.1 Item 9

(Die Datenbasis befindet sich in Anhang 14)

Die Auswahl der Teilnehmer für Item 9 zeigt, dass die meisten Dokumente oder Arbeitsergebnisse einem Review unterzogen werden. Das führte zu einer Bewertung mit Stufe RG3. Die Antwortverteilung fiel weitestgehend einheitlich aus, was auf ein ebenso weitgehendes Verständnis bezüglich Reviews innerhalb der Teilnehmer hinweist. Auffällig ist hier die fehlende Auswahl von technischen Reviews wie Programmcode oder Konfigurationsdateien. Ebenfalls wurden Akzeptanzkriterien nicht ausgewählt. Hier ist ohne weiteren Kontext über die tatsächliche Nutzung von Akzeptanzkriterien keine Interpretation möglich. Auch hier wurde das Freitextfeld nicht verwendet, was den Schluss zulässt, dass aus Sicht der Teilnehmer keine Ergänzung der Antworten notwendig war.

4.4.3.2 Item 10

(Die Datenbasis befindet sich in Anhang 15)

Aus den Daten für Item 10 geht einstimmig hervor, dass für Reviews grundsätzlich keine Checklisten genutzt werden. Das führt zu der Bewertung von RG1.

4.4.3.3 Item 11

(Die Datenbasis befindet sich in Anhang 16)

Die Beantwortung des Items 11 ist nicht eindeutig ausgefallen. Vier Befragte antworteten mit „Ja“ und drei mit „Nein“. Das ergab eine grenzwertige Einstufung von RG3. Es kann allerdings die Vermutung aufgestellt werden, dass das Ergebnis bei einer größeren Stichprobe eindeutiger ausgefallen wäre. Hier ist es sicher sinnvoll weiter nachzuforschen und den Ursachen genauer auf den Grund zu gehen und ggf. den Reviewprozess zu verbessern.

4.4.4 Merkmal: Fehlermanagement (Item 12 – 15)

Das Merkmal Fehlermanagement wurde insgesamt mit RG4 bewertet. Es wird ein strukturiert werkzeuggestützter Abweichungsprozess eingesetzt. Ebenfalls kommt ein einheitliches Schema für Fehlerberichte zu Einsatz. Dies führt zur Bestbewertung innerhalb des Reifegradmodells.

4.4.4.1 Item 12

(Die Datenbasis befindet sich in Anhang 17)

Dieses Item wurde von allen Teilnehmern mit einen „Ja“ beantwortet, was auch zu der Bewertung von RG4 führte. Interessant sind die Angaben im Freitextfeld, siehe dazu Tabelle 4-4²²³.

Tabelle 4-4: Freitextfelder Item 12

Testbericht
ARIS-Prozessbeschreibung
MF-ALM
MF-ALM, Testbereiche

²²³ Rechtschreibfehler wurden in den Freitextangaben nicht korrigiert.

MF-ALM, Testbericht, Statusbericht

MF-ALM, Stutusbericht, Testbericht

Eine Recherche ergab, dass es sich bei MF-ALM um ein Application-Lifecycle-Management System der Firma Micro Fokus²²⁴ handelt. Aufgrund der Anzahl der Antworten mit MF-ALM scheint es mit hoher Wahrscheinlichkeit in diesem Softwareentwicklungsverfahren eingesetzt zu werden. Weiter wurden Testbericht und oder Statusbericht genannt. All diese Antworten entsprechen nicht einer konkreten Beschreibung, sondern stellen eher Teile eines Abweichungsprozesses dar. Es kann die Vermutung aufgestellt werden, dass der Standardprozess, der mit MF-ALM geliefert wird, hier übernommen worden ist. So wird in Folge zwar ein konkreter Prozess gelebt, dieser ist aber nicht gesondert beschrieben. Dieser Fall würde die zuvor getätigte Vermutung²²⁵ bestätigen, dass zwar Prozesse innerhalb einer Organisation bekannt sind und gelebt werden, aber nicht konkret beschrieben bzw. dokumentiert. Daraus würde sich ein konkreter Vorteil für die Organisationseinheit durch die Werkzeugnutzung ableiten. In jeden Fall wäre es sinnvoll hier die weitere Nachforschung anzustreben.

4.4.4.2 Item 13

(Die Datenbasis befindet sich in Anhang 18)

Die Teilnehmer stimmten hier einstimmig für „Ja“. Das führt zu dem Schluss, dass ein einheitliches Schema für Fehlerberichte genutzt wird. Die Antworten führen zu der Bewertung RG4. Dieses Ergebnis verstärkt den Eindruck aus Item 12.

4.4.4.3 Item 14

(Die Datenbasis befindet sich in Anhang 19)

Die Teilnehmer stimmten hier einstimmig für „Ja“. Das führt zu dem Schluss, dass ein Fehler-Workflow-Werkzeug genutzt wird. Die Antworten führen zu der Bewertung RG4. Dieses Ergebnis stärkt ebenfalls den Eindruck aus Item 12. Auf diesen Umstand wird im weiteren Verlauf nicht eingegangen.

4.4.4.4 Item 15

(Die Datenbasis befindet sich in Anhang 20)

²²⁴ Weitere Informationen des Herstellers unter (Micro Focus International plc. 2019).

²²⁵ Siehe Kapitel 3.5.4.4 – Item Nr. 12.

Item 15 fließt nicht in die Bewertung über den Reifegrad mit ein. Wie in Tabelle 4-5 aufgestellt sind die Ergebnisse eher gestreut mit einer kleinen Konzentration bei der Antwort „Softwaretest“. Es scheint aber so, dass nur wenig Fehler im Produktionsbetrieb oder während einer Inbetriebnahme auftreten. Dies wäre ein Hinweis auf die Wirksamkeit der eingesetzten Qualitätssicherung.

Tabelle 4-5: Freitextfelder Item 15

	Ausgewählt
Anforderungsanalyse bzw. -prozess	1
Softwareentwicklung	0
Softwaretests	4
während einer Inbetriebnahme	0
während des Normalbetriebes	0
Andere	2

Die Auswahl der Antwort „Andere“ zeigt, dass hier die Bereitstellung eines Freitextfeldes durchaus sinnvoll gewesen wäre.

4.4.5 Merkmal: Testwerkzeuge und Automatisierung (Item 16 – 19)

Das Merkmal Testwerkzeuge und Automatisierung wurde insgesamt mit RG2 bewertet. Zur Testdokumentation wird ein digitales Managementsystem bzw. Werkzeug genutzt. Allerdings findet keine bis wenig Automatisierung von Testfällen statt.

4.4.5.1 Item 16

(Die Datenbasis befindet sich in Anhang 21)

Die Ergebnisse für Item 16 sind bis auf eine Ausnahme ebenfalls eindeutig. Allerdings musste in diesem Fall das Bewertungsschema angepasst werden. Ohne eine Anpassung läge die Einstufung bei RG2. Mit der Anpassung ergab sich eine Bewertung von RG4. Der Grund dafür ist, dass die Mehrheit der Teilnehmer die mit RG2 bewertete Antwort „Sonstige Form“ auswählte. Zusätzlich hatten alle diese Teilnehmer im Freitext „MF ALM“ eingetragen. Inhaltlich wäre für ein System wie MF ALM die mit RG4 bewertete Antwort „in einem digitalen Ticketsystem oder vergleichbares“ berechtigt gewesen. Nur einer der Teilnahme hatte dies allerdings ausgewählt. In Anbetracht der Ergebnisse wurde daher entschieden für diesen Lauf des Verfahrens die Bewertung der Antwort „Sonstige Form“ auf RG4 anzuheben. Es liegt allerdings auf der Hand, dass die Ausprägung „MF ALM“ der vorgesehenen Antwortmöglichkeit von den Teilnehmern nicht zugeordnet werden konnte. Das führt zu dem Schluss, dass die Antwortausprägungen für das Item 16 überdacht werden müssen um in Zukunft ein eindeutiges Antwortverhalten zu

erreichen. Zur Vollständigkeit findet sich in Anhang 21 für Item 16 beide Auswertungsvarianten. In das Endergebnis der Reifegradbewertung fließt allerdings nur die angepasste Bewertung ein.

4.4.5.2 Item 17

(Die Datenbasis befindet sich in Anhang 22)

In Item 17 gaben sechs von sieben Befragten einen Automatisierungsgrad von 0% an. Ein Befragter gab einen Automatisierungsgrad von 1%-10% an. Das Ergebnis weist daher eindeutig darauf hin, dass, wenn überhaupt, nur wenig Testautomatisierung zum Einsatz kommt. Der niedrige Prozentsatz führt zu einer Bewertung von RG1. Es ist allerdings möglich, dass sich vereinzelt erste Ansätze in der Erprobung befinden. Weitere Nachforschung oder Forcierung dieses Ansatzes könnten einen nächsten Schritt darstellen.

4.4.5.3 Item 18

(Die Datenbasis befindet sich in Anhang 23)

Der Eindruck aus Item 17 bestätigt sich hier, da die Mehrzahl der Antwort auf „Nicht-Automatisiert“ fällt, was ebenfalls zu einer Bewertung von RG1 führt. Es verstärkt aber auch die Vermutung erster Automatisierungsansätze, weil in einem Fall „Use-Case-Tests“ als automatisiert gewählt worden ist.

4.4.5.4 Item 19

(Die Datenbasis befindet sich in Anhang 24)

Item 19 fließt nicht in die Reifegradbewertung ein. Zwei von sieben Teilnehmern sind der Meinung, dass es Pläne zum Aufbau einer Testautomatisierung gibt. Der Rest der Teilnehmer ist dort allerdings anderer Meinung. Das Ergebnis bestätigt die Daten aus Item 17 und 18.

4.4.6 Merkmal: Soziale Kompetenz und Teamzusammensetzung (Item 20 – 22)

Das Merkmal Soziale Kompetenz und Teamzusammensetzung wurde insgesamt mit RG2 bewertet. Die unschlüssigen Ergebnisse über die Existenz von Einarbeitungs- und Ausgliederungskonzepten sowie die anscheinend wenigen Möglichkeiten der Befragten an QS-Fortbildungen teilzunehmen führten zu dieser Bewertung.

4.4.6.1 Item 20

(Die Datenbasis befindet sich in Anhang 25)

Nach den Ergebnissen von Item 20 scheint für die Mehrzahl der Teilnehmer kein Einarbeitungskonzept im Bereich Test zu bestehen. Zwei von sieben gaben an, dass ein solches Konzept existiert. Das ungleiche Antwortverhalten führt zu einer Einstufung von RG2. Die Angaben lassen keinen eindeutigen Schluss zu. Es ist denkbar, dass in der nahen Vergangenheit ein Konzept eingeführt worden ist und ein Großteil der Teilnehmer von der Nutzung des Konzeptes nicht mehr betroffen war. Wenn das der Fall wäre, ist es aber dennoch ungewöhnlich, dass die Existenz eines solchen Konzeptes nicht bekannt ist. Sollte es ein Konzept geben, muss dies den Mitarbeitern vorgestellt werden, falls nicht, wäre es einzuführen.

4.4.6.2 Item 21

(Die Datenbasis befindet sich in Anhang 26)

Item 21 lässt mit einem Antwortverhältnis von 4/3 keinen klaren Schluss zu. Es besteht die Möglichkeit, dass die Stichprobe in diesem konkreten Item zu klein war. Da sich die Mehrheit für die Existenz eines Ausgliederungskonzeptes ausgesprochen hat, führt das zu einer grenzwertigen Bewertung von RG3. Grundsätzlich ist die Bewertung aber mit der der vorherigen Items vergleichbar.

4.4.6.3 Item 22

(Die Datenbasis befindet sich in Anhang 27)

In Item 22 gab die Mehrheit an, dass ihnen keine Möglichkeit für QS-Fortbildungen oder QS-Tagungen angeboten wird. Das führt zu dem Schluss, dass ggf. der Bedarf für diese gegeben ist. Die daraus folgende Bewertung liegt bei RG1.

4.5 Kombination der Erkenntnisse einzelner Items

Nachdem nun jedes Item für sich einzeln betrachtet wurde wird in diesem Kapitel die Erkenntnisse der Items über die Merkmale kombiniert um damit konkrete Widersprüche oder Verstärkungen aufzudecken. Ziel ist ebenfalls ein Probelauf zur Darstellung beispielhafter Ableitungen auf Basis der Datenerhebung, wie sie im Anschluss einer Reifegradmessung erfolgen könnte. Es gilt hierbei zu beachten, dass die Ableitungen nur auf Grund der Datenbasis getätigt werden. Personen mit weiterem detailliertem Hintergrundwissen bezüglich der Organisationseinheit könnten detaillierte oder gänzliche andere Schlüsse erreichen. Zur besseren Nachvollziehbarkeit wird jeder Punkt mit folgender

Referenz versehen: „KEx“ für kombinatorische Erkenntnis das x steht dabei für eine Nummerierung.

- **KE1:**

Die Ergebnisse von Item 1, 8, 12, 13, 14 und 16 zeigen auf, dass die Organisationseinheit stark von einem werkzeuggestützten Test- und Abweichungsprozess profitiert. Es lässt sich ebenfalls die Annahme aufstellen, dass die Prozesse um das Tool herum entwickelt worden sind. Der Reifegrad dieser Prozesse könnte sich daher auf einem guten Niveau befinden, da gerade MF ALM (ehemals HP ALM) von Haus aus viele Funktionalitäten und Standards mitbringt wie zum Beispiel die eingesetzten Testmetriken.
- **KE2:**

Die Ergebnisse von Item 2, 6, 20 und 22 lassen den Schluss zu, dass ein ungleicher Wissenstand und ein erhöhter Bedarf an Fortbildungsmaßnahmen im Bereich Softwarequalitätssicherung in dieser Organisationseinheit bestehen. Es kann daraus die Annahme aufgestellt werden, dass die bestehenden Testprozesse zwar greifen und wirksam sind, aber beinahe kein „neues“ Wissen von außen in die Organisationseinheit dringt und nur wenig Weiterentwicklung der bestehenden Vorgehensweisen stattfindet.
- **KE3:**

Die Ergebnisse von Item 2, 3, 4 und 5 zeigen auf, dass bis auf die QS-Merkmale Sicherheit, Wartbarkeit und Portabilität alles abgedeckt ist. Die genutzten Testaktivitäten bestätigen diesen Schluss. Gleichzeitig werden aber nicht-funktionale Anforderungen im Anforderungsprozess nicht berücksichtigt. Diese stellen allerdings die Basis für einen Großteil der QS-Merkmale dar. Es kann daher die Annahme getroffen werden, dass die angenommene erreichte Abdeckung der QS-Merkmale eher auf Erfahrungswerten der Mitarbeiter basiert, als auf konkreten gestellten Anforderungen.
- **KE4:**

Die Ergebnisse von Item 2, 4, 7, 8, 9, 17, 18 lassen den Schluss zu, dass technische Aspekte in den verschiedenen Prozessen und Stakeholder keine Betrachtung finden. Im Umkehrschluss scheint sich das bestehende Vorgehen auf den fachlichen Aspekt einzuschränken.
- **KE5:**

Die Ergebnisse von Item 2, 15, 17, 18, 19 und 22 lassen den Schluss zu, dass eine Vielzahl von Testaktivitäten durchgeführt werden, die allerdings nur manuell zur Ausführung gebracht werden. Auch wenn dem Anschein nach vereinzelt Ansätze erprobt werden, scheint es keine konkrete Planung für eine Automatisierung zu geben. In Anbetracht des geringen Technikbezugs und der wenigen Fortbildungsmaßnahmen führt das zum Schluss, dass für eine Einführung von Testautomatisierung die Grundlagen fehlen und diese in einem ersten Schritt geschaffen werden müssten.

- **KE6:**

Die Ergebnisse von Item 11,15,16, 21 geben Hinweise darauf, dass ggf. die Fragestellung oder die Antwortmöglichkeiten überarbeitet werden sollten. In diesen Items wurden Freitextfelder genutzt, Ausweichantworten wie „Sonstige“ gewählt oder aber das Antwortverhalten der Befragten war nicht eindeutig. Das kann unterschiedliche Gründe haben, die sich nicht auf das Item beziehen, allerdings kann dies nach aktuellem Kenntnisstand auch nicht ausgeschlossen werden.

5 Validierung des Reifegradverfahrens

Die bisherigen Analysen und Erkenntnisse basierten auf den Ergebnissen der Datenerhebung sowie der subjektiven Wahrnehmung des Verfassers. Ziel dieses Verfahrens ist es die Realität in komprimierter Form so weit wie möglich wiederzugeben. Angelehnt an den Ansatz der Design Science²²⁶ stellt dies ein wichtiges Kriterium dar:

„If the artifact does not solve the problem (search, implementability), it has no utility. If utility is not demonstrated (evaluation), then there is no basis upon which to accept the claims that it provides any contribution (contribution).“²²⁷

Da im Kern des Verfahrens ein Fragebogen steht, bezieht sich das ebenfalls auf die Validität dieses Erhebungsinstrumentes:

„Der ermittelte Testwert muss valide sein (7 Abschn. 1.3.1). Das Kriterium der Validität betrifft die Gültigkeit des Tests: Ein Test ist valide, wenn er auch das erfasst, was er erfassen soll.“²²⁸

Das bedeutet: Nur wenn das Ergebnis diese Eigenschaft besitzt, kann das hier entwickelte Verfahren einen Mehrwert darstellen und als Basis für ein weiteres Vorgehen herangezogen werden. Ein wichtiger erster Schritt hierzu war es, die Mitarbeiter der Organisationseinheit direkt mit in die Datenerhebung einzubeziehen. Angelehnt an diesen Ansatz und dessen Prinzip Sieben²²⁹ zur Artefakt-Entwicklung wäre der nächste logische Schritt die Validierung bzw. Evaluation dieser Ergebnisse. Daraus abgeleitet sollte eine Validierung durch einen Personenkreis, der das notwendige Wissen über das Qualitätssicherungsvorgehen der zuvor bewerteten Organisationseinheit besitzt, erfolgen. Es liegt der Gedanke nahe, dass dazu die Befragten selbst als Mitarbeiter der Organisationseinheit in Form eines Interviews mit den Ergebnissen als Gruppe oder einzeln konfrontiert werden. Aus den Ergebnissen einer solchen Diskussion könnten konkrete Rückschlüsse über die Validität des Verfahrens gewonnen werden. Mit Blick auf den zeitlichen Rahmen und den Abgabetermin dieser Masterthesis ist allerdings ein vollständiges Interview oder vergleichbare Vorgehen mit allen Befragten aus zeitlichen Gründen nicht mehr durchführbar. Es wird daher ein verschlankter experimenteller Ansatz in Form eines halb-strukturierten Interviews mit dem Referatsleiter der Organisationseinheit durchgeführt. Aus Sicht des Verfassers ist der Referatsleiter in diesem Szenario nach den Befragten, die einzige Person, auf die die zuvor genannte Voraussetzung

²²⁶ Siehe dazu (Hevner u. a. 2004) und (Simon 1996).

²²⁷ (Hevner u. a. 2004: 91)

²²⁸ (Hussy u. a. 2013: 86)

²²⁹ Vgl. (Hevner u. a. 2004: 90–91).

ebenfalls zutrifft. Diese Vorgehensweise wird in dem Bewusstsein gewählt, obwohl sie keine vollständige Validierung und Evaluation des Verfahrens darstellt, aber dennoch erste Erkenntnisse für die weitere Forschung in Hinblick auf das Verfahren hervorbringen kann. Im Weiteren folgen nun die Gestaltung und Strukturierung des Interviews. Im Anschluss des Interviews wird auf die Ergebnisse eingegangen.

5.1 Interviewgestaltung

Es wird ein Interview explorativer Natur angestrebt, in diesem mittels offener Fragestellungen der Befragte sich mit den Ergebnissen des Verfahrens auseinandersetzen kann. Als Zeitraum wird eine Stunde angesetzt. Aus praktikablen Gründen und mit Blick auf die derzeitige Situation der aktuellen Pandemielage in Deutschland wird ein Telefoninterview veranschlagt. Das Interview wird mit Zustimmung des Befragten als Tonaufnahmen auf der zugehörigen CD der Masterthesis abgelegt (Siehe dazu Anhang 31 und auf der CD im Ordner „2_Anhänge“). Eine Transkription des Interviews befindet sich in Anhang 32. Das Interview wird in drei Abschnitte eingeteilt, welche im Weiteren beschrieben werden.

5.1.1 Interviewabschnitt Eins

In Abschnitt Eins ist das es Ziel, eine unabhängige Einschätzung des Befragten über die von ihm erwarteten Ergebnisse auf Ebene der Merkmale des Verfahrens einzuholen. Dazu wird dem Befragten vorab eine leere Ergebnistabelle, inkl. der Fragestellung zur Verfügung gestellt. Der Befragte erhält damit die Möglichkeit aus seiner Sicht zu jedem Merkmal den Reifegrad zu bestimmen. Diese Tabelle erhält er mit der Bitte, diese bis zum Interviewtermin auszufüllen. Die leere Tabelle kann in Anhang 28 eingesehen werden.

5.1.2 Interviewabschnitt Zwei

In Abschnitt Zwei des Interviews ist es das Ziel den Befragten direkt mit den Ergebnissen des Verfahrens zu konfrontieren und dazu seine persönliche Einschätzung und Sichtweise in Erfahrung zu bringen. Dazu werden die von den Befragten ausgefüllte Tabelle und die Ergebnisse des Verfahrens gegenübergestellt. Als Ergebnis werden dem Befragten für jedes Merkmal Folgendes vorgestellt: (Siehe die Tabelle dazu in Anhang 29)

- Das Reifegradergebnis
- Die zugehörige Reifegradbeschreibung aus Kapitel 3.6

Die Ebene der Merkmale wird aus zwei Gründen gewählt. Zum einen soll die Anonymität der Teilnehmer weiterhin sichergestellt werden. Zum anderen würde ein Interview

auf Itemebene die Dauer einer Stunde überschreiten und wahrscheinlich in einer Detailanalyse ufern. Würde aber auf der anderen Seite nur das Gesamtreifeergebnis in einem solchem Interview genutzt werden, ist es wahrscheinlich, dass die Erkenntnisse keinen Mehrwert zur Validierungsüberprüfung beitragen werden. Daher wird der Befragte zu jedem Merkmal mit folgender offener Fragestellung aufgefordert, seine persönliche Einschätzung über die ggf. abweichenden und oder übereinstimmenden Ergebnisse zu äußern:

„Wie schätzen Sie die erreichten Ergebnisse der Merkmale und des Verfahrens im Vergleich zu Ihrer eigenen Einschätzung ein?“

Bei Bedarf werden im Verlauf des Gespräches noch Anstöße gegeben auf mögliche Unterschiede und auf den möglichen Grund auf diese einzugehen. Abschließend wird das Gesamtreifegradergebnis des Verfahrens genannt und auch dazu die konkrete Meinung des Befragten eingeholt. Abschließend wird folgende Fragestellung dem Befragten gestellt: „Welche Ergebnisse überraschten Sie am meisten?“

5.1.3 Interviewabschnitt Drei

In Abschnitt Drei ist es Ziel im Sinne der Testprozessverbesserung in Erfahrung zu bringen, welche möglichen bzw. theoretischen Maßnahmen er abschließend aus den Ergebnissen dieses Verfahren ableiten würde. Dazu wird folgende Fragestellung genutzt: „Welche Schlüsse und möglichen weiteren Schritte ziehen Sie aus den Ergebnissen dieses Verfahrens?“

5.2 Ergebnisse des Interviews

In diesem Kapitel werden die Ergebnisse der zuvor definierten Interviewabschnitte beschrieben. Das Interview wurde nach Zustimmung des Befragten aufgezeichnet. Die Tonaufzeichnung²³⁰ findet sich auf der beiliegenden CD der Masterthesis (Siehe dazu Anhang 31 und auf der CD im Ordner „2_Anhänge“). Die Inhalte der Aufnahme werden hier so wiedergegeben, wie der Verfasser sie verstanden hat. Eine Transkription des Interviews befindet sich in Anhang 32.

In Abschnitt Eins wurde vom Befragten eine Selbsteinschätzung auf Basis des Reifegradmodells auf Ebene der Merkmale getätigt. Das Ergebnis der Selbsteinschätzung und der Erprobung des Verfahrens befinden sich in Tabelle 5-1.

²³⁰ (Wittsiepe und Kahlert 2020) Anm.: Die weiteren Quellenangaben in den Fußnoten zur Tonaufnahme beziehen sich immer auf die Minuten innerhalb der Aufnahme.

Tabelle 5-1: Selbsteinschätzung vs. Erprobung

Merkmale	Reifegrad		Abweichung zur Erprobung
	Selbsteinschätzung	Verfahrenserprobung	
Testprozess	3	3	0
Testmanagement	2	3	+1
Reviews	4	2	-2
Fehlermanagement	4	4	0
Testwerkzeuge und Automatisierung	3	2	-1
Soziale Kompetenz und Teamzusammensetzung	2	2	0
Gesamtreifegrad	3	3	0

Stellt man die Ergebnisse direkt gegenüber, ist zu erkennen, dass der Gesamtreifegrad übereinstimmt. Vergleicht man die Bewertungen der einzelnen Merkmale, ergeben sich dennoch Differenzen in den Bewertungen der Merkmale Testmanagement, Reviews sowie Testwerkzeuge und Automatisierung. Das gleiche Endergebnis lässt sich rechnerisch durch das Rundungsverfahren auf ganze Zahlen erklären. Ermittelt man den Mittelwert ohne diese, spiegelt sich auch hier der Unterschied wieder: Für die Selbsteinschätzung 3 und für das Verfahrensergebnis 2,666666667.

Abschnitt zwei behandelte die direkte Konfrontation und Gegenüberstellung beider Ergebnisse mit dem Gesprächspartner. Merkmale, bei denen es eine vollständige Übereinstimmung gab, wurden vom Befragten gar nicht bis wenig hinterfragt und mündlich bestätigt. Dies traf für die Merkmale Testprozess²³¹, Fehlermanagement²³², Soziale Kompetenz und Teamzusammensetzung²³³ zu. Darüber hinaus wurden folgende Punkte während der Besprechung vom Befragten geäußert:

- Das eingesetzte Test-Werkzeug wird als sehr wertvoll für die dortige QS gesehen.
- Es existiert bisher kein Einarbeitungskonzept und der bisherige Prozess ist verbesserungswürdig. Allerdings werden hier Hürden für eine Änderung befürchtet.

²³¹ Vgl. (Wittsiepe und Kahlert 2020: 02:22-04:11 Min.).

²³² Vgl. (Wittsiepe und Kahlert 2020: 12:32-14:05 Min.).

²³³ Vgl. (Wittsiepe und Kahlert 2020: 17:08-20:46 Min.).

- Das Qualifizierungsniveau ist nach Ansicht des Befragten aufgrund der vielen Erfahrungswerte auf einem hohen Stand.

Über diese Merkmale hinaus lagen auch Merkmale mit einer leichten Abweichung vor über die etwas detaillierter gesprochen worden ist. Eine leichte Abweichung ist eine positive oder negative Differenz in der Reifegradbewertung von einer Stufe im Reifegradmodell. Folgende Merkmale betrifft dies: Testmanagement²³⁴ sowie Testwerkzeuge und Automatisierung²³⁵. Das Merkmal Testmanagement wurde vom Befragten um eine Stufe niedriger eingeschätzt als das Verfahren ermittelte. Das Gespräch ergab, dass wohl besonders im Bereich der Testmetriken für Testaktivitäten mit fachlichem Kontext noch Defizite bestehen und daraus im Kern die Bewertung entstand. Das Merkmal Testwerkzeuge und Automatisierung wurde vom Befragten um eine Stufe höher bewertet als in der Erprobung. Kommentiert wurde dieser Punkt wie folgt:

- Es bestehen anfänglich erste Bemühungen für eine Testautomatisierung.
- Ggf. wird Testautomatisierung in den nächsten Jahren mit einer neuen Anwendung eingeführt, ist aber aktuell kein Thema.
- Die starke Testwerkzeugnutzung wurde nochmal hervorgehoben.

Am Ende wurde aber vom Gesprächspartner bestätigt, dass zurzeit keine Testautomatisierung durchgeführt wird. Er kam zum dem Schluss, dass die Bewertung von RG2 nachvollziehbar ist. Das verbleibende und einzige Merkmal mit einer stärkeren Abweichung war Review²³⁶. Dieses Merkmal wurde mit einer Abweichung von 2 Stufen in der Selbsteinschätzung höher bewertet. Der Gesprächspartner war grundsätzlich von den eingesetzten Maßnahmen im Bereich der Reviews überzeugt und hätte hier ein besseres Ergebnis erwartet. Es wurde die Vermutung geäußert, dass es hier ggf. ein Kommunikations- oder Verständnisproblem über das Thema Review im Referat besteht. Im Gespräch kam auch zutage, dass Checklisten für Reviews regelmäßig eingesetzt werden. Es wurde allerdings erwähnt, dass bisher wohl immer die gleichen Personen die Reviews durchführten und erst seit kurzen hier eine Veränderung stattfand. Das Gesamtreifegradergebnis lag in beiden Fällen auf RG3. Der Befragte wirkte zufrieden mit diesem Gesamtergebnis. Die Frage²³⁷ „Welche Ergebnisse überraschten Sie am meisten?“ wurde mit dem Ergebnis des Merkmales Reviews beantwortet. Es wurde dabei

²³⁴ Vgl. (Wittsiepe und Kahlert 2020: 04:12-08:08 Min.).

²³⁵ Vgl. (Wittsiepe und Kahlert 2020: 14:45-17:07 Min.).

²³⁶ Vgl. (Wittsiepe und Kahlert 2020: 08:09-12:31 Min.).

²³⁷ Vgl. (Wittsiepe und Kahlert 2020: 22:20-23:50 Min.).

angemerkt, dass der Befragte sich die Ergebnisse durchaus erklären könnte. Auf Nachfrage hin wurden keine weiteren Ergebnisse als überraschend empfunden.

Im dritten Abschnitt²³⁸ wurde über die möglichen Schlüsse und weitere Schritte gesprochen. Dabei wurden vom Befragten drei Themenpunkte genannt: Verbesserungen im Bereich Testautomatisierung, Kommunikation zum Thema Reviews und die Verbesserung des Onboardings- bzw. Einführungskonzept.

5.3 Analyse der Interviewergebnisse und Validierung der Reifegradmessung

Im Weiteren folgt die Analyse der Interviewergebnisse in direktem Bezug zu den zuvor beschriebenen Reifegradergebnissen²³⁹, der kombinatorischen Analyse²⁴⁰ und ggf. einzelner Items²⁴¹. Die Kernergebnisse des Interviews lassen sich wie folgt zusammenfassen:

- Bis auf ein paar Abweichungen im Detail scheint das Reifegradergebnis die Sicht der Teilnehmer und in weiten Teilen die Realität der Organisationseinheit abzubilden.
- Es wird das Testmanagementwerkzeug MF-ALM genutzt.
- Es gibt vereinzelt leichte Abweichungen, die sich aber im nachvollziehbaren Bereich bewegen.
- Es gibt Hinweise darauf, dass die Art der Fragestellung im Bereich Testmetrik nicht den Kern des Themas treffen konnte.
- Die starke Abweichung im Bereich der Reviews wurde mit Überraschung wahrgenommen. Eine mögliche Erklärung wird im Bereich eines Verständnisproblems über die realen Abläufe bei den Teilnehmern vermutet.
- Die Ergebnisse regten den Befragten dazu an, erste Ansätze für zukünftige Maßnahmen zu entwickeln.
- Im Gegensatz zu den einheitlichen Angaben der Teilnehmer im Verfahren existieren einheitliche Checklisten für Reviews.
- Es gibt kein konkretes Einführungskonzept. Auf ein fehlendes Ausgliederungskonzept wurde nicht eingegangen.
- Das Wissensniveau im Bereich QS wurde als ausreichend beschrieben.

²³⁸ Vgl. (Wittsiepe und Kahlert 2020: 23:50-26:27 Min.).

²³⁹ Siehe Kapitel 4.3.

²⁴⁰ Siehe Kapitel 4.5.

²⁴¹ Siehe Kapitel 4.4.

5.3.1 Validierung Reifegradmessung

Das Interview lässt den Schluss zu, dass das Reifegradmodell und die daraus abgeleitete Bewertung greifen und die Realität durchaus in einem begrenzten Rahmen darstellen können. Besonders die zuvor durchgeführte Selbsteinschätzung im Rahmen des Interviews bestätigt dies. Auch wenn es Abweichungen im Detail gab, so wurden diese bewertungstechnisch durch das Rundungsverfahren im Gesamtergebnis ausgeglichen.

Für die Abweichungen könnten viele Erklärungen in Frage kommen. Um diese einzugrenzen, wären weitere Nachforschungen und Gespräche notwendig. Der zeitliche Rahmen für die Abgabe der Masterthesis ermöglicht dies allerdings nicht. Aus diesem Grund können die Gründe für die Abweichungen nicht eindeutig zugeordnet werden. Nutzt man das Mittel der Ableitung auf Basis der Ergebnisse des Interviews und des Verfahrens, so wären mindestens zwei Erklärungen möglich. Zu einem könnten die Teilnehmer in den abweichenden Punkten nicht die Erwartungshaltung des Referatsleiters erfüllen. Gegensätzliche Aussagen im Bereich der Fortbildung und Reviews lassen diesen Schluss zu. Weiter könnte zwischen den beiden Themen eine Kausalität bestehen. Auf der anderen Seite ist es ebenso möglich, dass die Items die Ansichten der Teilnehmer nicht vollständig abbilden konnten. Geht man letzterem Ansatz weiter nach, so existieren Hinweise dafür, dass in den Themenbereichen der Testmetrik, Review, Ausgliederungskonzept und ggf. Fortbildung die Fragestellung weiter verfeinert oder verändert werden müssen, um in Folge eine bessere Abbildung zu erreichen. Darüber hinaus besteht immer Grund zur Annahme, dass weitere Faktoren für die Abweichungen existieren, auf die es allerdings mit dem bestehenden Wissenstand keine Hinweise gibt oder die zum aktuellen Zeitpunkt nicht erkannt worden sind. Unter diesen Gesichtspunkten ist die Tatsache interessant, dass trotz der Abweichungen der Referatsleiter sich und seine Organisationseinheit in den Ergebnissen wiederfinden konnte.

Ein weiterer Punkt, der in die Betrachtung mit einfließen muss, ist, dass unter der Annahme, dass die Reifegradmessung des Verfahrens zu 100% korrekt gewesen wäre, aus Sicht der Organisationseinheit in manchen Bereichen unabhängig von der Bewertung eine konkrete Verbesserung vorerst und bewusst nicht vorgesehen ist. Die Erkenntnisse im Bereich Automatisierung in Kombination mit den Interviewaussagen stützen dies. Das bedeutet, dass auch eine niedrige Reifegradebewertung seine strategische Berechtigung haben kann und sollte.

Das schlussendliche Ziel dieses Verfahrens der Reifegradmessung konnte trotz der Abweichungen erreicht werden: Der Befragte konnte für seine Organisationseinheit einen ersten Eindruck erhalten und Ansätze für mögliche Verbesserungen aus den Ergebnissen ableiten. Ob diese weiterverfolgt und verwirklicht werden, und wie sich diese Verbesserungen am Ende gestalten, liegt außerhalb des Rahmens dieses Verfahrens

und der Masterthesis. Dennoch wäre ein wichtiger nächster Schritt für die Verfeinerung und weitere Entwicklung des Reifegradverfahrens über die Masterthesis hinaus im Anschluss dieser Verbesserung eine erneute Reifegradmessung mit exakt demselben Verfahren durchzuführen. Daraus könnten aus Sicht des Verfassers unter anderem folgende Erkenntnisse entstehen:

- Wird die Erfassung des Reifegrades im gleichen Maße möglich wie bei der ersten Messung, ist das Instrument also reliabel?
- Wird sich eine mögliche Reifegradsteigerung in der erneuten Messung wieder spiegeln und somit die Vergleichbarkeit beider Messungen bestätigen?
- Wiederholen sich oder ergeben sich vergleichbare Abweichungen auch in der zweiten Messung?
- Wirkt sich eine erneute Reifegradmessung auf die Anzahl der Teilnehmer aus?

5.3.2 Validierung der kombinatorischen Ergebnisse

Im Weiteren wird sich auf die in Kapitel 4.5 entwickelten Erkenntnisse bezogen. Ziel ist es, Rückschlüsse auf die Ergebnisse von den Aussagen der Interviews abzuleiten um ggf. eine Verstärkung und Minderung der Erkenntnisse festzustellen.

- **KE1:**
Mehrere Antworten über verschiedene Items²⁴² hinweg ließen die Nutzung eines Testmanagementwerkzeuges vermuten. Diese Vermutung bestätigt sich im Rahmen des Interviews. Auch die Annahme, dass die Organisationseinheit von der Nutzung profitiert verstärkt sich hierdurch.
- **KE2:**
Die Äußerung, dass ein hoher Wissensstand innerhalb des Referates herrscht, bestätigt den Schluss, dass wenig neues Wissen in die genutzten Methoden einfließt. Das Interview bestätigt damit indirekt das Ergebnis, dass QS-Fortbildungsmaßnahmen innerhalb der Organisationseinheiten weniger angeboten bzw. genutzt werden. Die Daten aus dem Verfahren und diese Aussagen wirken auf den ersten Blick gegensätzlich, zeigen aber auf, dass unter den Teilnehmern ein berechtigtes Interesse an Fortbildungsmaßnahmen bestehen könnte. Die im Interview erwähnte geplante neue Anwendungsentwicklung stützt die Annahme für den Bedarf weiterer QS-Fortbildungsmaßnahmen unter den Teilnehmern.
- **KE3:**
Zu KE3 wurden innerhalb des Interviews weder Aussagen getroffen, die den Schluss stützen oder verneinen. Hier wäre eine weitere Nachforschung sinnvoll und notwendig, um dem weiter auf den Grund zu gehen.

²⁴² Siehe die Items 1, 12, 14 sowie 16.

- **KE4:**
Die Aussage, dass gerade bei den fachlich getriebenen Testmetriken noch Verbesserungsbedarf besteht, verstärkt den Eindruck aus KE4, dass technische Aspekte weniger Betrachtung finden. Weitere Hinweise gab es hingegen nicht.
- **KE5:**
Das Interview bestätigte eindeutig den Schluss, dass keine Testautomatisierung innerhalb der Organisationseinheit genutzt wird. Ebenfalls wurde bestätigt, dass es in der Vergangenheit erste Ansätze für eine Testautomatisierung vorlag. Die Distanz zu den Technikthemen konnten hingegen weder bestätigt noch verneint werden.
- **KE6:**
Das Interview zeigte im Bereich Reviews ergänzende Hinweise für ein mögliches Verbesserungspotential des Verfahrens auf.

6 Abschlussbewertung

Zu Anfang wurde eine Problemstellung sowie ein konkretes Forschungsziel beschrieben und aufgezeigt. Das Ziel lautete daraufhin:

Konzeption eines begründeten Verfahrens zur Reifegradmessung der Softwarequalitätssicherung im öffentlichen Sektor.

Zur Vorbereitung dessen wurde der aktuelle Stand dieser Thematik zusammenfassend aufgearbeitet. Darauf folgten in Kapitel 3, nach dem Vorgehen der Design-Science, die Entwicklung eines Reifegradmodells inkl. Messinstrument und Auswertungsmatrix. In beiden darauffolgenden Kapiteln wurde die Erprobung dieser Artefakte durchgeführt und die Ergebnisse im möglichen Rahmen einer Validierung unterzogen.

Mit Blick auf die Ergebnisse scheint das hier entwickelte Verfahren seinen ursprünglich angestrebten Zweck der QS- Reifegradmessung zu erfüllen. Auch wenn diese Aussage sicherlich nur mit Einschränkungen gültig ist, wurde ein nachvollziehbares Ergebnis für die Beteiligten erreicht. Die Konzeption einer Reifegradmessung der Software-Qualitätssicherung im öffentlichen Sektor wurde hiermit abgeschlossen. Im Weiteren folgt eine kritische Würdigung der Ergebnisse dieser Thesis sowie ein abschließendes Fazit.

6.1 Kritische Würdigung und Entwicklungsräume

Das Verfahren zeigte im Verlauf seiner Entwicklung einige Schwächen auf. Auch wenn wahrscheinlich nicht alle Schwächen hier eine Benennung finden können, wird im Weiteren versucht auf mögliche Bereiche einzugehen:

6.1.1 Gewählte Themen der Softwarequalitätssicherung

Eine offensichtliche Schwäche wäre der Bereich Themenauswahl die das Verfahren für eine Reifegradmessung heranzieht. Im direkten Vergleich mit Methoden wie TPI-Next oder TMMi ist schnell zu erkennen, dass dieses Verfahren nicht den gleichen Umfang abdeckt. Und auch weil Überschneidungen sich in einigen Punkten auf natürliche Art ergeben, wirft das unweigerlich auch einen Schatten auf die Bereiche, die keine Überschneidungen besitzen. Es war nie das Ziel eine vergleichbare Abdeckung zu erreichen, aber es kann sicherlich darüber debattiert werden warum bestimmte Themenbereiche hier eine Abdeckung erführen und andere eben nicht. Bereits die initiale Einschränkung auf die Inhalte des ISTQB Syllabus könnte so ein Punkt sein. Daraus schließt sich ein möglicher Raum für Verbesserung, der in der Forschung weiter aufgegriffen werden kann. Zum Beispiel wäre eine Kollaboration von mehreren QS-Fachexperten möglich, die die Themenpunkte und deren Bewertung auf Basis ihres Erfahrungsschatzes allein schon bereichern könnten. Ein anderer Weg könnte eine intensive Literaturrecherche

sein, die bewusst auf der Suche nach Punkten für eine Reifegradmessung, die über den ISTQB Syllabus hinausgeht.

6.1.2 Detailtiefe der Items

Die bestehenden Erkenntnisse zeigen Hinweise darauf, dass die angesetzte Detailtiefe einiger Items ggf. nicht die nötige Darstellung der realen Verhältnisse erreichen konnte oder die Ergebnisse keinen eindeutigen Schluss auf diese ermöglichten. Konkrete Beispiele gab es hier im Bereich der „Reviews“, „Fehlermanagement“, „Testwerkzeuge und Automatisierung“ und „Soziale Kompetenz und Teamzusammensetzung“. Würde das Verfahren zu einer erneuten Ausführung gebracht werden, sollten diese Punkte mit mehr Fokus betrachtet werden. Konkrete Verbesserungsvorschläge wurden auch bereits in Kapitel 4 genannt. Idealerweise erfolgt eine erneute Durchführung mit einer größeren Stichprobe.

6.1.3 Eingeschränkte Evaluation

Da der zeitliche Rahmen der Verfahrens- und Ergebnisevaluation nur kurz ausfiel, wäre eine umfangreichere Evaluation in Zukunft angebracht. Ein Vorschlag für das weitere Vorgehen wäre, dass mit allen Teilnehmern des Verfahrens die Ergebnisse gemeinsam diskutiert werden. Ebenfalls könnten ergänzende Dokumentanalysen weitere Rückschlüsse auf die Validität des Verfahrens bzw. der Ergebnisse ermöglichen.

6.1.4 Struktureller Aufbau der Befragung bzw. Fragebogens

Die Entwicklung und der Aufbau der Befragung und des Fragebogens erfolgten mit einem starken Fokus auf die Wirtschaftsinformatik und vor allem die Softwarequalitätssicherung. Es wurden zwar zur Entwicklung Werke speziell für die Erstellung von Fragebögen eingesetzt und zur Rate gezogen. Es ist aber klar, dass nicht alle Aspekte und Feinheiten dieser berücksichtigt werden konnten, da die Entwicklung solcher Instrumente viele Herausforderungen mit sich bringen kann, denen hier nicht nachgekommen werden konnte.

„Die Entwicklung eines Fragebogens ist vielmehr eine außerordentlich komplizierte Angelegenheit und kann nur dann zu einem befriedigenden Ergebnis führen, wenn dabei neben Intuition, Sprachgefühl und Erfahrung auch und vor allem wissenschaftliche Erkenntnisse über die bei einer Befragung ablaufenden Prozesse Berücksichtigung finden. Selbst hochkarätige Umfrageforscher gestehen sich und anderen ein, dass man einen

guten Fragebogen nicht am „grünen Tisch“ entwickeln kann: „Even after years of experience, no expert can write a perfect questionnaire“ (Sudman und Bradburn 1982: 283).²⁴³

Daher wäre es sicherlich eine Bereicherung für das Verfahren, wenn dieses durch fachkundige Experten auf Herz und Nieren geprüft und ggf. angepasst würde.

6.2 Fazit

Der zeitliche Rahmen einer Masterthesis gilt immer als begrenzt und daher ist die Entwicklung mit dieser Abgabe auch abgeschlossen. Und dennoch ist es immer der Anspruch ein abgerundetes Paket abzuliefern. So bleibt ein wenig das Gefühl zurück, dass eine größere und umfangreichere Erprobung weitere spannende Ergebnisse aufgezeigt hätte. Es ist in jeden Fall der Wunsch, dass eine weitere Entwicklung auch in einem größeren Rahmen stattfinden könnte, um dort konkret die benannten Entwicklungsspielräume anzugehen. Außerhalb einer Masterthesis und in einer idealen Welt ist es sicherlich auch sinnvoll für die Entwicklung solcher Instrumente unterschiedliche Disziplinen der Forschung mit einzubeziehen, um damit in Summe ein besseres Instrument hervor zu bringen. Eine Weiterentwicklung und Forschung durch andere für eine mögliche „Version 1.x oder 2.0“ ist erwünscht.

Mit Blick auf eine Kultur des „Keine negative Kritik ist des Lobes genug“ ist besonders erwähnenswert, dass das Referat, in der die Erprobung erfolgte, sich mit einem sehr positiven Feedback bedankte und die Erkenntnisse konkret nutzen möchte. Es könnte in Zukunft auch eine Replizierung der Erprobung mit demselben Verfahren möglich sein. Das zeigt aus Sicht des Verfassers, dass alles in allem ein zufriedenstellendes Ergebnis erreicht werden konnte. Das Thema der Softwarequalitätssicherung ist nach Ansicht des Verfassers ein Bereich, der gerade im öffentlichen Sektor mehr Betrachtung finden sollte und Instrumente wie dieses, können auch bereits im kleinen Rahmen direkten und positiven Mehrwert beitragen. Im Umfeld der Tätigkeit des Verfassers finden sich in jedem Fall die Möglichkeit einer weiteren Nutzung und Entwicklung dieses Verfahrens und dies hoffentlich auch außerhalb des Bundesverwaltungsamtes.

²⁴³ (Porst 2014: 14)

Abbildungsverzeichnis

Abbildung 1-1: Herleitung Forschungsantwort.....	7
Abbildung 2-1: Teufelsquadrat	13
Abbildung 2-2: The evolution of TMA and TPI models and their relationships	16
Abbildung 2-3: TPI-Bewertungsmatrix.....	18
Abbildung 2-4: CMM/CMMI Entwicklung.....	20
Abbildung 2-5: CMMI/TMMi Reifegradstufen	22
Abbildung 3-1: Reifegradmodell	36
Abbildung 3-2: Themenpunkt Reifegrad	37

Tabellenverzeichnis

Tabelle 1-1: Design-Science Research Guidelines	11
Tabelle 3-1: Checkliste der Leitlinien zur Itemformulierung.....	45
Tabelle 3-2: Item Nr. 1	47
Tabelle 3-3: Item Nr. 2	48
Tabelle 3-4: Item Nr. 3	50
Tabelle 3-5: Item Nr. 4	51
Tabelle 3-6: Item Nr. 5	52
Tabelle 3-7: Item Nr. 6	53
Tabelle 3-8: Item Nr. 7	54
Tabelle 3-9: Item Nr.8	54
Tabelle 3-10: Item Nr. 9.....	56
Tabelle 3-11: Item Nr. 10	56
Tabelle 3-12: Item Nr. 11	57
Tabelle 3-13: Item Nr. 12.....	59
Tabelle 3-14: Item Nr. 13.....	60
Tabelle 3-15: Item Nr. 14	61
Tabelle 3-16: Item Nr. 15.....	62
Tabelle 3-17: Item Nr. 16.....	64
Tabelle 3-18: Item Nr. 17.....	65
Tabelle 3-19: Item Nr. 18.....	65
Tabelle 3-20: Item Nr. 19.....	66
Tabelle 3-21: Item Nr. 20.....	67
Tabelle 3-22: Item Nr. 21	68
Tabelle 3-23: Item Nr. 22	69
Tabelle 4-1: Ergebnis der Reifegradbewertung	76
Tabelle 4-2: Ergebnis Item 2	78
Tabelle 4-3: Ergebnis Item 4	80
Tabelle 4-4: Freitextfelder Item 12	82
Tabelle 4-5: Freitextfelder Item 15	84
Tabelle 5-1: Selbsteinschätzung vs. Erprobung	92

Anhangsverzeichnis

Anhang 1: Beiblatt für Frage 4 und 5	104
Anhang 2: Fragebogen	108
Anhang 3: E-Mail zur Beschreibung des Verfahrens	117
Anhang 4: E-Mail Einladung des Verfahrens	118
Anhang 5: Ergebnistabelle Reifegrad	119
Anhang 6: Item 1	120
Anhang 7: Item 2	121
Anhang 8: Item 3	122
Anhang 9: Item 4	123
Anhang 10: Item 5	124
Anhang 11: Item 6	124
Anhang 12: Item 7	125
Anhang 13: Item 8	126
Anhang 14: Item 9	127
Anhang 15: Item 10	128
Anhang 16: Item 11	128
Anhang 17: Item 12	129
Anhang 18: Item 13	129
Anhang 19: Item 14	130
Anhang 20: Item 15	130
Anhang 21: Item 16	131
Anhang 22: Item 17	132
Anhang 23: Item 18	133
Anhang 24: Item 19	134
Anhang 25: Item 20	134
Anhang 26: Item 21	135
Anhang 27: Item 22	135
Anhang 28: Interview Selbsteinschätzung Leer	136
Anhang 29: Selbsteinschätzung Vergleich	138
Anhang 30: Gesamtergebnistabelle Auswertungen_Leer.xlsx	141
Anhang 31: Interview Referatsleiter 04.06.20	141
Anhang 32: Transkription Interview Referatsleiter 04.06.20	142
Anhang 33: Lebenslauf	152

Anhang 1: Beiblatt für Frage 4 und 5

Dieses Beiblatt wird für die Fragestellung Nummer 4 und 5 benötigt. Es dient dazu die genutzten Begrifflichkeiten kurz zu erläutern. Sie können es jetzt lesen oder aber auch später, wenn Sie bei den betreffenden Fragen angekommen sind.

Funktionale Eignung

Diese Eigenschaft gibt an, inwieweit ein Produkt oder System Funktionen bereitstellt, die den angegebenen und implizierten Anforderungen entsprechen, wenn sie unter bestimmten Bedingungen verwendet werden. Dieses Merkmal setzt sich aus Folgenden Untermerkmalen zusammen:

- Funktionale Vollständigkeit - Grad, in dem der Funktionsumfang alle festgelegten Aufgaben und Benutzerziele abdeckt.
- Funktionskorrektheit - Grad, in dem ein Produkt oder System die richtigen Ergebnisse mit der erforderlichen Präzision liefert.
- Funktionale Angemessenheit - Grad, in dem die Funktionen die Erfüllung bestimmter Aufgaben und Ziele erleichtern.

Leistungsfähigkeit

Dieses Merkmal repräsentiert die Leistung im Verhältnis zur Menge der unter den angegebenen Bedingungen verwendeten Ressourcen. Dieses Merkmal setzt sich aus Folgenden Untermerkmalen zusammen:

- Zeitverhalten - Grad, in dem die Antwort- und Verarbeitungszeiten sowie die Durchsatzraten eines Produkts oder Systems bei der Ausführung seiner Funktionen den Anforderungen entsprechen.
- Ressourcennutzung - Grad, in dem die Mengen und Arten von Ressourcen, die von einem Produkt oder System bei der Ausführung seiner Funktionen verwendet werden, den Anforderungen entsprechen.
- Kapazität - Grad, in dem die Höchstgrenzen eines Produkts oder Systemparameters den Anforderungen entsprechen.

Kompatibilität

Grad, in dem ein Produkt, System oder eine Komponente Informationen mit anderen Produkten, Systemen oder Komponenten austauschen und / oder die erforderlichen Funktionen ausführen kann, während dieselbe Hardware- oder Softwareumgebung gemeinsam genutzt wird. Dieses Merkmal setzt sich aus Folgenden Untermerkmalen zusammen:

- Koexistenz - Grad, in dem ein Produkt seine erforderlichen Funktionen effizient ausführen kann, während eine gemeinsame Umgebung und Ressourcen mit anderen Produkten geteilt werden, ohne dass sich dies nachteilig auf andere Produkte auswirkt.

- Interoperabilität - Grad, in dem zwei oder mehr Systeme, Produkte oder Komponenten Informationen austauschen und die ausgetauschten Informationen verwenden können

Benutzerfreundlichkeit

Grad, in dem ein Produkt oder System von bestimmten Benutzern verwendet werden kann, um bestimmte Ziele mit Effektivität, Effizienz und Zufriedenheit in einem bestimmten Nutzungskontext zu erreichen. Dieses Merkmal setzt sich aus Folgenden Untermerkmalen zusammen:

- Erkennbarkeit der Angemessenheit - Grad, in dem Benutzer erkennen können, ob ein Produkt oder ein System für ihre Anforderungen geeignet sind.
- Lernfähigkeit - Grad, in dem ein Produkt oder System von bestimmten Benutzern verwendet werden kann, um bestimmte Ziele zu erreichen: Lernen, das Produkt oder System in einem bestimmten Anwendungskontext mit Effektivität, Effizienz, Risikofreiheit und Zufriedenheit zu verwenden.
- Bedienbarkeit - Grad, in dem ein Produkt oder System Attribute aufweist, die die Bedienung und Steuerung vereinfachen.
- Benutzerfehlerschutz. Inwieweit ein System Benutzer vor Fehlern schützt.
- Ästhetik der Benutzeroberfläche - Grad, in dem eine Benutzeroberfläche eine angenehme und zufriedenstellende Interaktion für den Benutzer ermöglicht.
- Zugänglichkeit - Grad, in dem ein Produkt oder System von Personen mit den unterschiedlichsten Merkmalen und Fähigkeiten verwendet werden kann, um ein bestimmtes Ziel in einem bestimmten Verwendungskontext zu erreichen.

Verlässlichkeit

Grad, in dem ein System, Produkt oder eine Komponente bestimmte Funktionen unter bestimmten Bedingungen für einen bestimmten Zeitraum ausführt. Dieses Merkmal setzt sich aus Folgenden Untermerkmalen zusammen:

- Fälligkeit - Grad, in dem ein System, Produkt oder eine Komponente die Anforderungen an die Zuverlässigkeit im normalen Betrieb erfüllt.
- Verfügbarkeit - Grad, in dem ein System, Produkt oder eine Komponente betriebsbereit und bei Bedarf zugänglich ist.
- Fehlertoleranz - Grad, in dem ein System, Produkt oder eine Komponente trotz vorhandener Hardware- oder Softwarefehler wie vorgesehen funktioniert.
- Wiederherstellbarkeit - Grad, in dem ein Produkt oder System im Falle einer Unterbrechung oder eines Fehlers die direkt betroffenen Daten wiederherstellen und den gewünschten Status des Systems wiederherstellen kann.

Sicherheit

Grad, in dem ein Produkt oder System Informationen und Daten schützt, so dass Personen oder andere Produkte oder Systeme den Grad des Datenzugriffs haben, der ihren Arten und Berechtigungsstufen entspricht. Dieses Merkmal setzt sich aus Folgenden Untermerkmalen zusammen:

- Vertraulichkeit - Grad, in dem ein Produkt oder System sicherstellt, dass Daten nur für diejenigen zugänglich sind, die zum Zugriff berechtigt sind.
- Integrität - Grad, in dem ein System, Produkt oder eine Komponente den unbefugten Zugriff auf oder die Änderung von Computerprogrammen oder -daten verhindert.
- Nicht-Zurückweisung - Grad, in dem nachgewiesen werden kann, dass Handlungen oder Ereignisse stattgefunden haben, so dass die Ereignisse oder Handlungen später nicht zurückgewiesen werden können.
- Rechenschaftspflicht - Grad, in dem die Aktionen eines Unternehmens eindeutig auf das Unternehmen zurückgeführt werden können.
- Authentizität - Grad, in dem nachgewiesen werden kann, dass die Identität eines Subjekts oder einer Ressource die behauptete ist.

Wartbarkeit

Diese Eigenschaft gibt den Grad an Effektivität und Effizienz an, mit dem ein Produkt oder System modifiziert werden kann, um es zu verbessern, zu korrigieren oder an Änderungen der Umgebung und der Anforderungen anzupassen. Dieses Merkmal setzt sich aus Folgenden Untermerkmalen zusammen:

- Modularität - Grad, in dem ein System oder Computerprogramm aus diskreten Komponenten besteht, sodass eine Änderung an einer Komponente nur minimale Auswirkungen auf andere Komponenten hat.
- Wiederverwendbarkeit - Grad, in dem ein Asset in mehr als einem System oder beim Aufbau anderer Assets verwendet werden kann.
- Analysierbarkeit - Grad der Wirksamkeit und Effizienz, mit dem es möglich ist, die Auswirkungen einer beabsichtigten Änderung an einem oder mehreren seiner Teile auf ein Produkt oder System zu bewerten oder ein Produkt auf Mängel oder Fehlerursachen zu diagnostizieren oder Teile zu identifizieren modifiziert sein.
- Modifizierbarkeit - Grad, in dem ein Produkt oder System effektiv und effizient modifiziert werden kann, ohne dass Fehler auftreten oder die vorhandene Produktqualität beeinträchtigt wird.
- Testbarkeit - Grad der Wirksamkeit und Effizienz, mit dem Testkriterien für ein System, ein Produkt oder eine Komponente festgelegt und Tests durchgeführt werden können, um festzustellen, ob diese Kriterien erfüllt wurden.

Portabilität

Grad der Effektivität und Effizienz, mit der ein System, Produkt oder eine Komponente von einer Hardware, Software oder einer anderen Betriebs- oder Nutzungsumgebung auf eine andere übertragen werden kann. Dieses Merkmal setzt sich aus Folgenden Untermerkmalen zusammen:

- Anpassungsfähigkeit - Grad, in dem ein Produkt oder System effektiv und effizient an unterschiedliche oder sich entwickelnde Hardware, Software oder andere Betriebs- oder Nutzungsumgebungen angepasst werden kann.

- Installierbarkeit - Grad der Effektivität und Effizienz, mit der ein Produkt oder System in einer bestimmten Umgebung erfolgreich installiert und / oder deinstalliert werden kann.
- Austauschbarkeit - Grad, in dem ein Produkt ein anderes angegebenes Softwareprodukt für denselben Zweck in derselben Umgebung ersetzen kann.

Anhang 2: Fragebogen

(Hinweis: Die exakte Formatierung unterscheidet sich von original Dokument)

Sehr geehrte Damen und Herren,

ich würde mich Ihnen gerne vorstellen: Mein Name ist Mike Wittsiepe. Vielleicht kennen Sie mich bereits aus dem „Forum für Test und Qualitätssicherung (ForTeQ)“ im BVA. Erstmal Vielen Dank, dass Sie sich die Zeit nehmen an dieser Befragung zum Thema „Reifegradmessung in der Softwarequalitätssicherung“ teilzunehmen. Im Rahmen meiner Master-Thesis entwickle ich dieses Werkzeug. Für Ihre Teilnahme bin ich Ihnen daher sehr dankbar, da die Erprobung dieses Werkzeugs einen wichtigen Beitrag zu meiner Abschlussarbeit leistet.

Im Folgenden werden Ihnen kurz die Hintergründe, der Ablauf des Verfahrens und die Zielsetzung dargelegt. Bitte lesen Sie diese aufmerksam und in Ruhe durch.

Hintergründe

Die Erwartungen an die Qualität und Leistungsfähigkeit von Software steigen seit der Einführung des Internets und der damit einhergehenden Vernetzung immer weiter an. Gleichzeitig erhöht sich die Komplexität von Software immer mehr. Während man früher eine Anwendung mit verhältnismäßig wenig Aufwand vollständig verstehen und bedienen konnte, werden dazu heutzutage mitunter ganze Teams beschäftigt. Dadurch wachsen die Herausforderungen an die Qualitätssicherung immens. Vor einigen Jahren konnte man vielleicht noch sagen „Das wird schon irgendwie hinhauen“ und wenn mal im Betrieb was in Schieflage geraten war, konnte dies wahrscheinlich ohne große Auswirkung behoben werden. Heutzutage können Fehleranalysen und -korrekturen Wochen in Anspruch nehmen Die mögliche Folge daraus ist: Digitale Prozesse und Dienstleistungen deutschlandweit ggf. sogar europaweit, werden im schlimmsten (Fehler-) Fall blockiert und laufen nicht. Menschen oder Unternehmen warten vielleicht dringend auf Geld oder Zahlungen. Und wenn Software im Bereich der öffentlichen Sicherheit verwendet wird, könnten sogar Menschenleben auf dem Spiel stehen. Es ist allgemein anerkannt, dass in jeder Software Fehler stecken. Idealweise werden vor einer Inbetriebnahme so viele Fehler wie möglich entdeckt. Die Mittel dazu sind allerdings begrenzt. Es kommt also immer darauf an mit den gegebenen Mitteln mindestens die kritischen und gefährlichen Fehler frühzeitig aufzudecken, obwohl man vollkommen fehlerfreie Software vermutlich niemals herstellen kann.

Zielsetzung

Ziel ist es also innerhalb eines Softwareentwicklungsverfahrens oder -wartungsverfahrens die gegebenen Mittel so kostengünstig und effizient wie möglich zu nutzen. Vor dem Hintergrund der steigenden Erwartungen ist dies häufig eine Gratwanderung und etwas wo jede Organisation sich konstant weiterentwickeln kann. Der Zweck dieser Befragung ist es diesen Entwicklungsprozess im Rahmen der Qualitätssicherung zu unterstützen. Dazu wird ein sogenannter Reifegrad erhoben. Damit ist es möglich Entwicklungspotentiale, Herausforderungen und mögliche Erleichterungen einfach aufzudecken. Und wer könnte besser bei der Aufdeckung helfen

als Mitarbeiterinnen und Mitarbeiter die in ihrer Arbeit täglich diese Software entwickeln bzw. warten. Bitte beachten Sie, es existieren weder „richtige“ noch „falsche“ Antworten. Bitte lesen Sie sich die Fragen aufmerksam durch und treffen, wenn möglich, spontan eine Entscheidung.

Ablauf des Verfahrens

Nachdem Sie das Dokument digital ausgefüllt haben, speichern Sie es bitte im Docx- oder PDF-Format ab und senden es bitte an die folgende E-Mailadresse: „mike.wittsiepe@bva.bund.de“.

Das Verfahren verläuft vollständig anonymisiert. Ihr ausgefüllter Bogen wird nach Empfang von der E-Mail getrennt und gespeichert. Die E-Mail selbst wird daraufhin gelöscht. Jeder Bogen erhält eine einmalige ID, diese dient allein dazu, dass ein Fragebogen nicht mehrfach eingereicht werden kann, da dies sonst die Ergebnisse verfälschen würde. Ihre mögliche Teilnahme an dieser Befragung wurde im Vorfeld mit Ihrem Vorgesetzten abgestimmt und wird von diesem unterstützt. Es ist ebenfalls abgestimmt, dass Ihre Vorgesetzten keinen Einblick in die ausgefüllten Fragebogen erhalten. Die Teilnahme ist selbstverständlich freiwillig und kann jederzeit abgebrochen werden.

*Vielen Dank und viele Grüße
Mike Wittsiepe*

Frage 1

Existiert für Ihre Organisationseinheit eine konkrete Beschreibung (z.B. Teststrategie, Testkonzept) über die eingesetzten Testprozesse? Wenn Ja, in welcher Form liegen diese vor?

(Bitte wählen Sie nur Ja oder Nein aus)

Nein

Ja, in folgender Form: [Klicken oder tippen Sie hier, um Text einzugeben.](#)

Frage 2

Welche Testaktivitäten sind Ihnen bekannt und werden Ihres Wissens nach in Ihrer Organisationseinheit regelmäßig eingesetzt?

(Bitte wählen Sie pro Testart nur einer der drei gegebenen Möglichkeiten)

Testarten	Unbekannt	Bekannt	Eingesetzt
Datensicherungstests	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Kompatibilitätstests	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Katastrophenbewältigungstest (z.B. bei Ausfall)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Funktionstests	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Integrationstests	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Lasttests / Stresstests	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Zugänglichkeit- / Benutzbarkeitstests	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Ablauftests	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Sicherheitstests	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Abnahmetests	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Systemtests / Systemintegrationstests	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Komponententests	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
White-Box-Tests (z.B. Pfad-, Entscheidungs-, zustandsbasiert)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Erfahrungsbasierte Tests	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Explorative Tests	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Use-Case-Tests	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Code-Reviews	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Akzeptanztests	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Regressionstests	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Session-based testing	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Pair testing	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Frage 3

Werden Ihrer Meinung nach nicht-funktionale Anforderungen im Rahmen des aktuellen Anforderungsprozesses ausreichend berücksichtigt?

(Bitte wählen Sie nur Ja oder Nein aus)

Ja

Nein

Frage 4

Welche Qualitätsmerkmale decken die genutzten Testaktivitäten innerhalb Ihrer Organisationseinheit ab? Siehe Erläuterungen zu den Qualitätsmerkmalen auf dem Beiblatt.

(Bitte wählen Sie pro Qualitätsmerkmale nur einer der zwei gegebenen Möglichkeiten aus)

Qualitätsmerkmale	Nicht-Abgedeckt	Abgedeckt
Funktionale Eignung	<input type="checkbox"/>	<input type="checkbox"/>
Leistungsfähigkeit	<input type="checkbox"/>	<input type="checkbox"/>
Komptabilität	<input type="checkbox"/>	<input type="checkbox"/>
Benutzerfreundlichkeit	<input type="checkbox"/>	<input type="checkbox"/>
Verlässlichkeit	<input type="checkbox"/>	<input type="checkbox"/>
Sicherheit	<input type="checkbox"/>	<input type="checkbox"/>
Wartbarkeit	<input type="checkbox"/>	<input type="checkbox"/>
Portabilität	<input type="checkbox"/>	<input type="checkbox"/>

Frage 5

Welche Relevanz haben die Qualitätsmerkmale Ihrer Meinung nach innerhalb Ihrer Organisationseinheit? Siehe Erläuterungen zu den Qualitätsmerkmalen auf dem Beiblatt.

(Von „Wichtig“ bis „Nicht Wichtig“.)

Qualitätsmerkmale	Reihenfolge
Funktionale Eignung	Wählen Sie ein Element aus.
Leistungsfähigkeit	Wählen Sie ein Element aus.
Komptabilität	Wählen Sie ein Element aus.
Benutzerfreundlichkeit	Wählen Sie ein Element aus.
Verlässlichkeit	Wählen Sie ein Element aus.
Sicherheit	Wählen Sie ein Element aus.
Wartbarkeit	Wählen Sie ein Element aus.
Portabilität	Wählen Sie ein Element aus.

Frage 6

Von welcher Personengruppen werden Testaktivitäten Ihrer Meinung hauptsächlich innerhalb Ihrer Organisationseinheit ausgeführt?

(Bitte wählen Sie nur einer der beiden Möglichkeiten aus)

Wählen Sie ein Element aus.

Frage 7

Welche Stakeholder wurden Ihrer Meinung nach für die bestehenden Qualitätssicherungsmaßnahmen Ihrer Organisationseinheit berücksichtigt?

(Die Mehrfachauswahl ist hier erlaubt)

- Entwickler; Architekten; Endnutzer; Fachbereich; IT-Abteilung;
 - Support bzw. Helpdesk; Sachgebietsleiter; Sicherheitsbeauftragter / Abteilung;
 - Datenschutzbeauftragter / Abteilung; Releasemanagement;
 - Sonstige: Klicken oder tippen Sie hier, um Text einzugeben.
-

Frage 8

Welche Testmetriken finden Ihrer Ansicht nach innerhalb Ihrer Organisationseinheit Verwendung?

(Die Mehrfachauswahl ist hier erlaubt)

- Anforderungsabdeckung; Abweichungsübersicht (Offen, Geschlossen, ...);
 - Codeüberdeckung; Testfallstatus (Offen, Ausgeführt, Blockiert, in Arbeit, ...);
 - Testumgebungsverfügbarkeit;
 - Sonstige: Klicken oder tippen Sie hier, um Text einzugeben.
-

Frage 9

Auf welchen Dokumenten oder Arbeitsergebnissen werden Reviews in Ihrer Organisationseinheit durchgeführt?

(Die Mehrfachauswahl ist hier erlaubt)

- Anforderungen; Akzeptanzkriterien; Testfälle; Spezifikationsdokumente;
 - Schnittstellenbeschreibungen; Testergebnisse; Fehlertickets, Programmcode;
 - Konfigurationsdateien; Testdaten; Benutzerhandbücher;
 - Sonstige Dokumente: Klicken oder tippen Sie hier, um Text einzugeben.
-

Frage 10

Werden zur Unterstützung der Reviews jeweils einheitliche Checklisten eingesetzt?

(Bitte wählen Sie nur Ja oder Nein aus)

Nein

Ja, für folgende Dokumente: [Klicken oder tippen Sie hier, um Text einzugeben.](#)

Frage 11

Werden innerhalb Ihrer Organisationseinheit Tester an einem Reviewprozess der zu entwickelten Anforderungen beteiligt?

(Bitte wählen Sie nur Ja oder Nein aus)

Ja

Nein

Frage 12

Existiert für Ihre Organisationseinheit eine konkrete Beschreibung über die eingesetzten Fehler- bzw. Abweichungsprozesse?

(Bitte wählen Sie nur Ja oder Nein aus)

Nein

Ja, in folgender Form: [Klicken oder tippen Sie hier, um Text einzugeben.](#)

Frage 13

Wird Ihrer Meinung nach innerhalb Ihrer Organisationseinheit für die Fehler- bzw. Abweichungsdokumentation (dem sogenannten Fehlerbericht) ein einheitliches Schema genutzt?

(Bitte wählen Sie nur Ja oder Nein aus)

Ja

Nein

Frage 14

Wird innerhalb Ihrer Organisationseinheit ein Fehler-Workflow-Werkzeug eingesetzt?

(Bitte wählen Sie nur Ja oder Nein aus)

Nein

Ja

Frage 15

Wann wird Ihrer Erfahrung nach innerhalb Ihrer Organisationseinheit Fehler bzw. Abweichungen oder Spezifikationslücken am ehesten entdeckt?

(Bitte wählen Sie einer der Möglichkeiten aus)

Wählen Sie ein Element aus.

Frage 16

Wo und in welcher Form werden in Ihrer Organisationseinheit die Testfälle und Ergebnisse dokumentiert?

(Bitte wählen Sie einer der Möglichkeiten aus)

Wählen Sie ein Element aus.

Nur bei Angabe von Sonstige Form Klicken oder tippen Sie hier, um Text einzugeben.

Frage 17

Wie hoch ist Ihrer Meinung nach aktuell der Anteil an automatisierten Testfällen innerhalb Ihrer Organisationseinheit? (Bitte schätzen Sie)

(Bitte wählen Sie einer der Möglichkeiten aus)

Wählen Sie ein Element aus.

Frage 18

Welche Testaktivitäten sind innerhalb Ihrer Organisationseinheit Ihres Wissens nach bereits automatisiert?

(Bitte wählen Sie pro Qualitätsmerkmale nur einer der zwei gegebenen Möglichkeiten aus. Sollten Sie keinen der genannten Testaktivitäten nutzen, so wählen Sie das bitte das im Anschluss gegebene Feld dazu aus.)

Testarten	Nicht-Auto- matisiert	Automa- tisiert
Datensicherungstests	<input type="checkbox"/>	<input type="checkbox"/>
Kompatibilitätstests	<input type="checkbox"/>	<input type="checkbox"/>
Katastrophenbewältigungstest (z.B. bei Ausfall)	<input type="checkbox"/>	<input type="checkbox"/>
Funktionstests	<input type="checkbox"/>	<input type="checkbox"/>
Integrationstests	<input type="checkbox"/>	<input type="checkbox"/>
Lasttests / Stresstests	<input type="checkbox"/>	<input type="checkbox"/>
Zugänglichkeit- / Benutzbarkeitstests	<input type="checkbox"/>	<input type="checkbox"/>
Ablauftests	<input type="checkbox"/>	<input type="checkbox"/>
Sicherheitstests	<input type="checkbox"/>	<input type="checkbox"/>
Abnahmetests	<input type="checkbox"/>	<input type="checkbox"/>
Systemtests / Systemintegrationstests	<input type="checkbox"/>	<input type="checkbox"/>
Komponententests	<input type="checkbox"/>	<input type="checkbox"/>
White-Box-Tests (z.B. Pfad-, Entscheidungs-, zustandsbasiert)	<input type="checkbox"/>	<input type="checkbox"/>
Erfahrungsbasierte Tests	<input type="checkbox"/>	<input type="checkbox"/>
Explorative Tests	<input type="checkbox"/>	<input type="checkbox"/>
Use-Case-Tests	<input type="checkbox"/>	<input type="checkbox"/>
Code-Reviews	<input type="checkbox"/>	<input type="checkbox"/>
Akzeptanztests	<input type="checkbox"/>	<input type="checkbox"/>
Regressionstests	<input type="checkbox"/>	<input type="checkbox"/>
Session-based testing	<input type="checkbox"/>	<input type="checkbox"/>
Pair testing	<input type="checkbox"/>	<input type="checkbox"/>

Keine der genannten Testaktivitäten wird für Tests verwendet und daher auch nicht für die Automatisierung berücksichtigt.

Frage 19

Bestehen Pläne dazu, über das aktuelle Maß hinaus, Testautomatisierungsaktivitäten einzuführen?

(Bitte wählen Sie nur Ja oder Nein aus)

Ja

Nein

Frage 20

Existiert für Ihre Organisationseinheit ein Einarbeitungskonzept bzw. -plan für neue Mitarbeiter mit Testaufgaben?

(Bitte wählen Sie nur Ja oder Nein aus)

Nein

Ja

Frage 21

Existiert ein Ausgliederungskonzept (z.B. Übergabeprotokolle, Wissensdokumentation) für Mitarbeiter mit Testaufgaben die Ihre Organisationseinheit geplant verlassen?

(Bitte wählen Sie nur Ja oder Nein aus)

Ja

Nein

Frage 22

Werden innerhalb Ihrer Organisationseinheit Mitarbeitern regelmäßig die Teilnahme an Fortbildungsmaßnahmen oder Tagungen zum Thema Softwarequalitätssicherung ermöglicht?

(Bitte wählen Sie nur Ja oder Nein aus)

Nein

Ja

Vielen Dank für Ihre Teilnahme an dieser Befragung.

Anhang 3: E-Mail zur Beschreibung des Verfahrens

Hallo Herr XXX,

vielen Dank das Sie Interesse an der Erprobung dieses Werkzeuges zeigen. Im Anhang finden Sie zwei Dateien. Die "Fragebogen.docx" Datei stellt den eigentlichen Fragebogen dar. Dort sind auch Hintergrund, Zielsetzung und Ablauf beschrieben. "Fragebogen_Beiblatt.pdf" ist ein informatives Beiblatt um bestimmte Begrifflichkeiten zu erläutern. Natürlich wäre die Teilnahme freiwillig, aber es macht Sinn das so viele wie möglich teilnehmen. Im Fokus steht allein die Methodik der Qualitätssicherung. Um ein möglichst unabhängiges Ergebnis zu erreichen, würde ich sie bitte die Inhalte und genauen Fragestellung des Fragebogens mit ihrem Team nicht konkret zu besprechen. Da alles via Mail-Kommunikation läuft, stellte die Ausfüllung aus dem Home-Office kein Problem dar. Die Bearbeitung des Bogens ist bewusst eingeschränkt, so dass nur bestimmte Felder bearbeitet werden dürfen.

Teilnehmervoraussetzung ist:

- Muss direkt oder indirekt beteiligt sein an dem Softwareentwicklungsprojekt.
- Muss die deutsche Sprache beherrschen
- Muss mindestens 6 Monate bei Ihnen im Verfahren tätig sein.

Konkret könnten das Personen sein, die folgenden Rolle oder ähnliches besitzen: Tester, Testmanager, Sachgebietsleiter, Referatsleiter, Fachbereichskollegen, Entwickler, ggf. weitere Mitarbeiter die am Entwicklungsverfahren beteiligt sind.

Sollten Sie die Befragung erlauben, würde ich um eine E-Mailliste der möglichen Teilnehmer bitten, so dass ich jeden direkt und einzeln anschreiben kann. Das soll zusätzlich die Anonymität sicherstellen. Es wird auch auf personenbezogene Daten im Fragebogen verzichtet. Ich würde den Teilnehmern 3-4 Tage zur Ausfüllung geben.

Ich würde mich sehr freuen, wenn das klappen könnte.

Sie können mich jederzeit unter folgender Nummer erreichen: XXX

Mit freundlichen Grüßen

Mike Wittsiepe

Anhang 4: E-Mail Einladung des Verfahrens

Sehr geehrte Damen und Herren,

Ihr Verfahren hat sich bereiterklärt an einer Befragung zum Thema Reifegrad der Softwarequalitätssicherung teilzunehmen.

Alle weiteren Informationen und Anweisungen finden Sie in der angehangenen Word-Datei "Befragung.docx". Die PDF Datei "Befragung_Beiblatt.pdf" dient zur Erläuterung bestimmter Begrifflichkeiten. Falls eine Meldung über Markos in Word erscheint, bestätigen Sie diese bitte mit "Inhalt aktivieren". Nachdem Sie das Dokument digital ausgefüllt haben, speichern Sie es bitte im Docx- oder PDF-Format ab und senden es bitte an die folgende

E-Mailadresse: "<E-Mailadresse>"

Sie können dafür auch die Antwortfunktion verwenden.

Die Frist zur Einreichung ist der <Datum>.

Sollten Sie Probleme bei der Ausfüllung des Fragebogens haben, so können Sie gerne jederzeit auf mich zu kommen.

Vielen Dank

Mit freundlichen Grüßen

Mike Wittsiepe

Anhang 5: Ergebnistabelle Reifegrad

Die Anhänge 5 – 29 stellen die Gesamtergebnistabelle „Auswertungen.xlsx“ dar in der jedes Item ein eigenes Tabellenblatt ist.

(Hinweis: Aus der Datei „Auswertungen.xlsx“ – Tabellenblatt „Ergebnis RG“ entnommen)

Übersicht der Einreichungen							
Anzahl der Bögen Einreichungen:	7		Anzahl der Bögen verteilt:		8		
Datensatz:	1	2	3	4	5	6	7
Bogen ID:	MT1305 01	MT1305 03	MT1305 04	MT1605 06	MT1305 07	MT1305 08	MT1305 05
Doppelcheck durchgeführt:	X	X	X	X	X	X	X

Item	RG	Merkmale	RG
1	4	Testprozess	3
2	3		
3	1		
4	3	Testmanagement	3
5	Kein		
6	Kein		
7	3		
8	3		
9	3	Reviews	2
10	1		
11	3		
12	4	Fehlermanagement	4
13	4		
14	4		
15	Kein		
16	4	Testwerkzeuge und Automatisierung	2
17	1		
18	1		
19	Kein		
20	2	Soziale Kompetenz und Teamzusammensetzung	2
21	3		
22	1		
Gesamt RG:		3	

Anhang 6: Item 1

(Hinweis: Aus der Datei „Auswertungen.xlsx“ – Tabellenblatt „Item 1“ entnommen)

Fragestellung: Existiert für Ihre Organisationseinheit eine konkrete Beschreibung (z.B. Teststrategie, Testkonzept) über die eingesetzten Testprozesse? Wenn Ja, in welcher Form liegen diese vor?

Anzahl: 7

RG Ergebnis: 4

Antworten:	Ja	Nein	Freitext	RG
Datensatz 1	1	0	Testkonzept	4
Datensatz 2	1	0	Word- und Excel-Dateien	4
Datensatz 3	1	0	Testkonzept	4
Datensatz 4	1	0		4
Datensatz 5	1	0	Testkonzept in Textform	4
Datensatz 6	1	0	MF-ALM und Testkonzept (pdf-Form)	4
Datensatz 7	1	0	Testkonzept	4

Mapping	
Antwort	RG
Ja	4
Nein	1

Anhang 7: Item 2

(Hinweis: Aus der Datei „Auswertungen.xlsx“ – Tabellenblatt „Item 2“ entnommen, Formatierung angepasst und Legende eingefügt)

Fragestellung:

Welche Testaktivitäten sind Ihnen bekannt und werden Ihres Wissens nach in Ihrer Organisationseinheit regelmäßig eingesetzt?

Legende: U = Unbekannt; B = Bekannt; E = Eingesetzt

Anzahl: 7

RG Ergebnis: 3

Datensatz	1			2			3			4			5			6			7			Mapping		
	U	B	E	U	B	E	U	B	E	U	B	E	U	B	E	U	B	E	U	B	E	Von	Bis	RG
Datensicherungstests			1			1			1			1			1			1			1	0	5	1
Kompatibilitätstests			1			1			1			1			1			1			1	6	10	2
Katastrophenbewältigungstest (z.B. bei Ausfall)			1			1			1			1			1			1			1	11	15	3
Funktionstests			1			1			1			1			1			1			1	16	21	4
Integrationstests			1			1			1			1			1			1			1			1
Lasttests / Stresstests			1			1			1			1			1			1			1			1
Zugänglichkeit- / Benutzbarkeitstests			1			1			1			1			1			1			1			1
Ablauftests			1			1			1			1			1			1			1			1
Sicherheitstests			1			1			1			1			1			1			1			1
Abnahmetests			1			1			1			1			1			1			1			1
Systemtests / Systemintegrationstests			1			1			1			1			1			1			1			1
Komponententests			1			1			1			1			1			1			1			1
White-Box-Tests (z.B. Pfad-, Entscheidungs-, zustandsbasiert)			1			1			1			1			1			1			1			1
Erfahrungsbasierte Tests			1			1			1			1			1			1			1			1
Explorative Tests			1			1			1			1			1			1			1			1
Use-Case-Tests			1			1			1			1			1			1			1			1
Code-Reviews			1			1			1			1			1			1			1			1
Akzeptanztests			1			1			1			1			1			1			1			1
Regressionstests			1			1			1			1			1			1			1			1
Session-based testing			1			1			1			1			1			1			1			1
Pair testing			1			1			1			1			1			1			1			1
Summe			11			10			11			5			11			11			11			
RG			3			2			3			1			3			3			3			

Anhang 8: Item 3

(Hinweis: Aus der Datei „Auswertungen.xlsx“ – Tabellenblatt „Item 3“ entnommen)

Fragestellung: Werden Ihrer Meinung nach nicht-funktionalen Anforderungen im Rahmen des aktuellen Anforderungsprozesses ausreichend berücksichtigt?

Anzahl:		RG Ergebnis:	
	7		1
Antworten:	Ja	Nein	RG
Datensatz 1	1		4
Datensatz 2		1	1
Datensatz 3		1	1
Datensatz 4		1	1
Datensatz 5		1	1
Datensatz 6		1	1
Datensatz 7		1	1

Mapping	
Antwort	RG
Ja	4
Nein	1

Anhang 10: Item 5

(Hinweis: Aus der Datei „Auswertungen.xlsx“ – Tabellenblatt „Item 5“ entnommen)

Fragestellung: Welche Relevanz haben die Qualitätsmerkmale Ihrer Meinung nach innerhalb Ihrer Organisationseinheit?

Anzahl:	RG Ergebnis: Kein						
Datensatz	1	2	3	4	5	6	7
Antworten	Rele- vanz	Rele- vanz	Rele- vanz	Rele- vanz	Rele- vanz	Rele- vanz	Rele- vanz
Funktionale Eignung	3	3	3	3	3	3	3
Leistungsfähigkeit	3	3	3	3	3	3	3
Komptabilität	2	2	3	3	3	3	3
Benutzerfreundlichkeit	3	3	3	3	3	3	3
Verlässlichkeit	3	3	3	3	3	3	3
Sicherheit	3	3	3	3	3	3	3
Wartbarkeit	2	3	3	3	3	3	3
Portabilität	1	2	2	3	2	2	2
Skala							
Wichtig	3						
Mittel	2						
Nicht Wichtig	1						

Anhang 11: Item 6

(Hinweis: Aus der Datei „Auswertungen.xlsx“ – Tabellenblatt „Item 6“ entnommen)

Fragestellung: Von welcher Personengruppen werden Testaktivitäten Ihrer Meinung hauptsächlich innerhalb Ihrer Organisationseinheit ausgeführt?

Anzahl:	RG Ergebnis: Kein	
Antworten:	Int. Ma	Ext. Ma
Datensatz 1	1	
Datensatz 2	1	
Datensatz 3	1	
Datensatz 4	1	
Datensatz 5	1	
Datensatz 6	1	
Datensatz 7	1	

Anhang 12: Item 7

(Hinweis: Aus der Datei „Auswertungen.xlsx“ – Tabellenblatt „Item 7“ entnommen, Formatierung angepasst und Legende eingefügt)

Fragestellung:

Welche Stakeholder wurden Ihrer Meinung nach für die bestehenden Qualitätssicherungsmaßnahmen Ihrer Organisationseinheit berücksichtigt?

Legende:

A. = Ausgewählt

Anzahl:

7

RG Ergebnis: 3

Datensatz	1	2	3	4	5	6	7
Antworten	A.	A.	A.	A.	A.	A.	A.
Entwickler							
Architekten							
Endnutzer	1	1	1		1	1	1
Fachbereich	1	1	1	1	1	1	1
IT			1	1	1	1	1
Support bzw. Helpdesk		1	1		1	1	1
Sachgebietsleiter	1	1	1			1	1
Sicherheitsbeauftragter / Abteilung		1	1		1	1	1
Datenschutzbeauftragter / Abteilung	1	1	1		1	1	1
Releasemanagement	1	1	1	1	1	1	1
Sonstige							
Freitext							
Summe	5	7	8	3	7	8	8
RG	2	3	3	1	3	3	3
Mapping							
Von	Bis	RG					
0	3	1					
4	6	2					
7	9	3					
10	11	4					

Anhang 13: Item 8

(Hinweis: Aus der Datei „Auswertungen.xlsx“ – Tabellenblatt „Item 8“ entnommen und Formatierung angepasst)

Fragestellung: Welche Testmetriken finden Ihrer Ansicht nach innerhalb Ihrer Organisationseinheit Verwendung?

Anzahl: 7 **RG Ergebnis:** 3

Datensatz	1	2	3	4	5	6	7
Antworten	Ausgewählt	Ausgewählt	Ausgewählt	Ausgewählt	Ausgewählt	Ausgewählt	Ausgewählt
Anforderungsabdeckung	1	1	1	1	1	1	1
Abweichungsübersicht		1	1	1	1	1	1
Codeüberdeckung							
Testfallstatus	1	1	1	1	1	1	1
Testumgebungsverfügbarkeit	1	1	1	1	1	1	1
Sonstige							
Freitext							
Summe	3	4	4	4	4	4	4
RG	3	3	3	3	3	3	3
Mapping							
Von	Bis	RG					
0	1	1					
2	2	2					
3	4	3					
5	6	4					

Anhang 14: Item 9

(Hinweis: Aus der Datei „Auswertungen.xlsx“ – Tabellenblatt „Item 9“ entnommen, Formatierung angepasst und Legende eingefügt)

Fragestellung: Auf welchen Dokumenten oder Arbeitsergebnissen werden Reviews in Ihrer Organisationseinheit durchgeführt?

Legende: A. = Ausgewählt

Anzahl: 7 **RG Ergebnis:** 3

Datensatz	1	2	3	4	5	6	7
Antworten	A.	A.	A.	A.	A.	A.	A.
Anforderungen	1	1	1	1	1	1	1
Akzeptanzkriterien							
Testfälle	1	1	1	1	1	1	1
Spezifikationsdokumente	1	1	1		1	1	1
Schnittstellenbeschreibungen	1	1					
Testergebnisse	1	1	1	1	1	1	1
Fehlertickets	1	1	1	1		1	1
Programmcode							
Konfigurationsdateien							
Testdaten	1	1	1	1	1	1	1
Benutzerhandbücher	1	1	1	1	1	1	1
Sonstige							
Freitext							
Summe	8	8	7	6	6	7	7
RG	3	3	3	2	2	3	3
Mapping							
Von	Bis	RG					
0	3	1					
4	6	2					
7	9	3					
10	12	4					

Anhang 15: Item 10

(Hinweis: Aus der Datei „Auswertungen.xlsx“ – Tabellenblatt „Item 10“ entnommen und Formatierung angepasst)

Fragestellung: Werden zur Unterstützung der Reviews jeweils einheitliche Checklisten eingesetzt?

Anzahl: 7 **RG Ergebnis:** 1

Antworten:	Ja	Nein	Freitext	RG
Datensatz 1		1		1
Datensatz 2		1		1
Datensatz 3		1		1
Datensatz 4		1		1
Datensatz 5		1		1
Datensatz 6		1		1
Datensatz 7		1		1
Mapping				
Antwort	RG			
Ja	4			
Nein	1			

Anhang 16: Item 11

(Hinweis: Aus der Datei „Auswertungen.xlsx“ – Tabellenblatt „Item 11“ entnommen und Formatierung angepasst)

Fragestellung: Werden innerhalb Ihrer Organisationseinheit Tester an einem Reviewprozess der zu entwickelten Anforderungen beteiligt?

Anzahl: 7 **RG Ergebnis:** 3

Antworten:	Ja	Nein	RG
Datensatz 1		1	1
Datensatz 2		1	1
Datensatz 3	1		4
Datensatz 4		1	1
Datensatz 5	1		4
Datensatz 6	1		4
Datensatz 7	1		4
Mapping			
Antwort	RG		
Ja	4		
Nein	1		

Anhang 17: Item 12

(Hinweis: Aus der Datei „Auswertungen.xlsx“ – Tabellenblatt „Item 12“ entnommen und Formatierung angepasst)

Fragestellung: Existiert für Ihre Organisationseinheit eine konkrete Beschreibung über die eingesetzten Fehler- bzw. Abweichungsprozesse?

Anzahl: 7 **RG Ergebnis:** 4

Antworten:	Ja	Nein	Freitext	RG
Datensatz 1	1		Testbericht	4
Datensatz 2	1		ARIS-Prozessbeschreibung	4
Datensatz 3	1		MF-ALM	4
Datensatz 4	1			4
Datensatz 5	1		MF-ALM, Testbereiche	4
Datensatz 6	1		MF-ALM, Testbericht, Statusbericht	4
Datensatz 7	1		MF-ALM, Stutusbericht, Testbericht	4

Mapping	
Antwort	RG
Ja	4
Nein	1

Anhang 18: Item 13

(Hinweis: Aus der Datei „Auswertungen.xlsx“ – Tabellenblatt „Item 13“ entnommen und Formatierung angepasst)

Fragestellung: Wird Ihrer Meinung nach innerhalb Ihrer Organisationseinheit für die Fehler- bzw. Abweichungsdokumentation (dem sogenannten Fehlerbericht) ein einheitliches Schema genutzt?

Anzahl: 7 **RG Ergebnis:** 4

Antworten:	Ja	Nein	RG
Datensatz 1	1		4
Datensatz 2	1		4
Datensatz 3	1		4
Datensatz 4	1		4
Datensatz 5	1		4
Datensatz 6	1		4
Datensatz 7	1		4

Mapping	
Antwort	RG
Ja	4
Nein	1

Anhang 19: Item 14

(Hinweis: Aus der Datei „Auswertungen.xlsx“ – Tabellenblatt „Item 14“ entnommen und Formatierung angepasst)

Fragestellung: Wird innerhalb Ihrer Organisationseinheit ein Fehler-Workflow-Werkzeug eingesetzt?

Anzahl: 7 **RG Ergebnis:** 4

Antworten:	Ja	Nein	RG
Datensatz 1	1		4
Datensatz 2	1		4
Datensatz 3	1		4
Datensatz 4	1		4
Datensatz 5	1		4
Datensatz 6	1		4
Datensatz 7	1		4

Mapping	
Antwort	RG
Ja	4
Nein	1

Anhang 20: Item 15

(Hinweis: Aus der Datei „Auswertungen.xlsx“ – Tabellenblatt „Item 15“ entnommen und Formatierung angepasst)

Fragestellung: Wann wird Ihrer Erfahrung nach innerhalb Ihrer Organisationseinheit Fehler bzw. Abweichungen oder Spezifikationslücken am ehesten entdeckt?

Anzahl: 7 **RG Ergebnis:** Kein

Datensatz	1	2	3	4	5	6	7
Antworten	Ausgewählt	Ausgewählt	Ausgewählt	Ausgewählt	Ausgewählt	Ausgewählt	Ausgewählt
Anforderungsanalyse bzw. -prozess		1					
Softwareentwicklung							
Softwaretests	1			1	1	1	
während einer Inbetriebnahme							
während des Normalbetriebes							
Andere			1				1

Anhang 21: Item 16

(Hinweis: Aus der Datei „Auswertungen.xlsx“ – Tabellenblatt „Item 16“ entnommen und Formatierung angepasst)

Fragestellung:	Wo und in welcher Form werden in Ihrer Organisationseinheit die Testfälle und Ergebnisse dokumentiert?						
Anzahl:	7	RG Ergebnis: 2			RG Ergebnis angepasst: 4		
Datensatz	1	2	3	4	5	6	7
Antworten	Ausgewählt	Ausgewählt	Ausgewählt	Ausgewählt	Ausgewählt	Ausgewählt	Ausgewählt
Stift und Papier							
strukturiert - lokal auf verschiedenen Computern				1			
strukturiert - im Netzwerkspeicher							
in einem digitalen Ticketsystem oder vergleichbares		1					
Unstrukturierte Form oder keine Dokumentation							
Sonstige Form	1		1		1	1	1
Freitext	MF ALM		MF ALM		MF ALM	MF ALM	MF ALM
RG	2	4	2	2	2	2	2
RG angepasst	4	4	4	2	4	4	4
Mapping							
Auswahl		RG	RG angepasst				
Stift und Papier		1	1				
strukturiert - lokal auf verschiedenen Computern		2	2				
strukturiert - im Netzwerkspeicher		3	3				
in einem digitalen Ticketsystem oder vergleichbares		4	4				
Unstrukturierte Form oder keine Dokumentation		1	1				
Sonstige Form		2	4				

Anhang 22: Item 17

(Hinweis: Aus der Datei „Auswertungen.xlsx“ – Tabellenblatt „Item 17“ entnommen und Formatierung angepasst)

Fragestellung: Wie hoch ist Ihrer Meinung nach aktuell der Anteil an automatisierten Testfällen innerhalb Ihrer Organisationseinheit?

Anzahl: 7 **RG Ergebnis:** 1

Datensatz	1	2	3	4	5	6	7
Antworten	Ausgewählt	Ausgewählt	Ausgewählt	Ausgewählt	Ausgewählt	Ausgewählt	Ausgewählt
0%	1		1	1	1	1	1
1% - 10 %		1					
11% - 20%							
21% - 30%							
31% - 40%							
41% - 50%							
Mehr als 50%							
RG	1	1	1	1	1	1	1
Mapping							
Auswahl	RG						
0%	1						
1% - 10 %	1						
11% - 20%	2						
21% - 30%	3						
31% - 40%	3						
41% - 50%	4						
Mehr als 50%	4						

Anhang 23: Item 18

(Hinweis: Aus der Datei „Auswertungen.xlsx“ – Tabellenblatt „Item 18“ entnommen, Formatierung angepasst und Legende eingefügt)

Fragestellung: Welche Testaktivitäten sind innerhalb Ihrer Organisationseinheit Ihres Wissens nach bereits automatisiert?

Legende: Ni. A = Nicht-Automatisiert; A = Automatisiert

Anzahl: 7 **RG Ergebnis:** 1

Datensatz	1		2		3		4		5		6		7		Mapping		
	Ni. A	A	Ni. A	A	Ni. A	A	Ni. A	A	Ni. A	A	Ni. A	A	Ni. A	A	Von	Bis	RG
Datensicherungstests	1				1		1		1		1		1		0	5	1
Kompatibilitätstests					1		1		1		1		1		6	10	2
Katastrophenbewältigungstest (z.B. bei Ausfall)					1		1		1		1		1		11	15	3
Funktionstests	1		1		1		1		1		1		1		16	21	4
Integrationstests	1				1		1		1		1		1		Keine Auswahl		1
Lasttests / Stresstests	1				1		1		1		1		1				
Zugänglichkeit- / Benutzbarkeits- tests			1		1		1		1		1		1				
Ablauf- tests	1		1		1		1		1		1		1				
Sicherheitstests			1		1		1		1		1		1				
Abnahmetests	1		1		1		1		1		1		1				
Systemtests / Systemintegrationstests					1		1		1		1		1				
Komponententests	1				1				1		1		1				
White-Box-Tests (z.B. Pfad-, Entscheidungs-, zustandsbasiert)					1				1		1		1				
Erfahrungsbasierte Tests	1		1		1		1		1		1		1				
Explorative Tests	1		1		1		1		1		1		1				
Use-Case-Tests	1			1	1				1		1		1				
Code-Reviews					1				1		1		1				
Akzeptanztests	1		1		1		1		1		1		1				
Regressionstests	1		1		1		1		1		1		1				
Session-based testing					1				1		1		1				
Pair testing					1				1		1		1				
Summe	0		1		0		0		0		0		0				
Alternativ Keine	-		-		-		-		1		-		-				
RG	1		1		1		1		1		1		1				

Anhang 24: Item 19

(Hinweis: Aus der Datei „Auswertungen.xlsx“ – Tabellenblatt „Item 19“ entnommen und Formatierung angepasst)

Fragestellung: Bestehen Pläne dazu, über das aktuelle Maß hinaus, Testautomatisierungsaktivitäten einzuführen?

Anzahl: 7 **RG Ergebnis:** Kein

Antworten:	Ja	Nein
Datensatz 1	1	
Datensatz 2	1	
Datensatz 3		1
Datensatz 4		1
Datensatz 5		1
Datensatz 6		1
Datensatz 7		1

Mapping	
Antwort	RG
Ja	4
Nein	1

Anhang 25: Item 20

(Hinweis: Aus der Datei „Auswertungen.xlsx“ – Tabellenblatt „Item 20“ entnommen und Formatierung angepasst)

Fragestellung: Existiert für Ihre Organisationseinheit ein Einarbeitungskonzept bzw. -plan für neue Mitarbeiter mit Testaufgaben?

Anzahl: 7 **RG Ergebnis:** 2

Antworten:	Ja	Nein	RG
Datensatz 1		1	1
Datensatz 2	1		4
Datensatz 3		1	1
Datensatz 4	1		4
Datensatz 5		1	1
Datensatz 6		1	1
Datensatz 7		1	1

Mapping	
Antwort	RG
Ja	4
Nein	1

Anhang 26: Item 21

(Hinweis: Aus der Datei „Auswertungen.xlsx“ – Tabellenblatt „Item 20“ entnommen und Formatierung angepasst)

Fragestellung: Existiert ein Ausgliederungskonzept (z.B. Übergabeprotokolle, Wissensdokumentation) für Mitarbeiter mit Testaufgaben die Ihre Organisationseinheit geplant verlassen?

Anzahl:	7	RG Ergebnis:		3
Antworten:	Ja	Nein		RG
Datensatz 1		1		1
Datensatz 2		1		1
Datensatz 3	1			4
Datensatz 4		1		1
Datensatz 5	1			4
Datensatz 6	1			4
Datensatz 7	1			4
Mapping				
Antwort	RG			
Ja	4			
Nein	1			

Anhang 27: Item 22

(Hinweis: Aus der Datei „Auswertungen.xlsx“ – Tabellenblatt „Item 20“ entnommen und Formatierung angepasst)

Fragestellung: Werden innerhalb Ihrer Organisationseinheit Mitarbeitern regelmäßig die Teilnahme an Fortbildungsmaßnahmen oder Tagungen zum Thema Softwarequalitätssicherung ermöglicht?

Anzahl:	7	RG Ergebnis:		1
Antworten:	Ja	Nein		RG
Datensatz 1	1			4
Datensatz 2		1		1
Datensatz 3		1		1
Datensatz 4		1		1
Datensatz 5		1		1
Datensatz 6		1		1
Datensatz 7		1		1
Mapping				
Antwort	RG			
Ja	4			
Nein	1			

Anhang 28: Interview Selbsteinschätzung Leer

(Hinweis: Aus der Datei „Interview.xlsx“ – Tabellenblatt „Selbsteinschätzung“ entnommen und Formatierung angepasst)

Bitte schätzen Sie den Reifegrad der Softwarequalitätssicherung bezogen auf die gegebenen Merkmale Ihrer Meinung nach ein Es sind dafür 4 verschiedene Einstufungen möglich: 1 - Initial; 2 - Mittel; 3 - Gut; 4 - Sehr gut Verwenden Sie dazu die mit "Reifegrad" gekennzeichnete Spalte F			
Item	Fragestellungen	Merkmale	Reifegrad
1	Existiert für Ihre Organisationseinheit eine konkrete Beschreibung (z.B. Teststrategie, Testkonzept) über die eingesetzten Testprozesse?	Testprozess	
2	Welche Testaktivitäten sind Ihnen bekannt und werden Ihres Wissens nach in Ihrer Organisationseinheit regelmäßig eingesetzt?		
3	Werden Ihrer Meinung nach nicht-funktionale Anforderungen im Rahmen des aktuellen Anforderungsprozesses ausreichend berücksichtigt?		
4	Welche Qualitätsmerkmale decken die genutzten Testaktivitäten innerhalb Ihrer Organisationseinheit ab?	Testmanagement	
5	Es erfolgte keine Reifegradbewertung mit dieser Frage		
6	Es erfolgte keine Reifegradbewertung mit dieser Frage		
7	Welche Stakeholder wurden Ihrer Meinung nach für die bestehenden Qualitätssicherungsmaßnahmen Ihrer Organisationseinheit berücksichtigt?		
8	Welche Testmetriken finden Ihrer Ansicht nach innerhalb Ihrer Organisationseinheit Verwendung?		
9	Auf welchen Dokumenten oder Arbeitsergebnissen werden Reviews in Ihrer Organisationseinheit durchgeführt?	Reviews	
10	Werden zur Unterstützung der Reviews jeweils einheitliche Checklisten eingesetzt?		
11	Werden innerhalb Ihrer Organisationseinheit Tester an einem Reviewprozess der zu entwickelten Anforderungen beteiligt?		

12	Existiert für Ihre Organisationseinheit eine konkrete Beschreibung über die eingesetzten Fehler- bzw. Abweichungsprozesse?	Fehlermanagement	
13	Wird Ihrer Meinung nach innerhalb Ihrer Organisationseinheit für die Fehler- bzw. Abweichungsdokumentation (dem sogenannten Fehlerbericht) ein einheitliches Schema genutzt?		
14	Wird innerhalb Ihrer Organisationseinheit ein Fehler-Workflow-Werkzeug eingesetzt?		
15	Es erfolgte keine Reifegradbewertung mit dieser Frage		
16	Wo und in welcher Form werden in Ihrer Organisationseinheit die Test-fälle und Ergebnisse dokumentiert?	Testwerkzeuge und Automatisierung	
17	Wie hoch ist Ihrer Meinung nach aktuell der Anteil an automatisierten Testfällen innerhalb Ihrer Organisationseinheit? (Bitte schätzen Sie)		
18	Welche Testaktivitäten sind innerhalb Ihrer Organisationseinheit Ihres Wissens nach bereits automatisiert?		
19	Es erfolgte keine Reifegradbewertung mit dieser Frage		
20	Existiert für Ihre Organisationseinheit ein Einarbeitungskonzept bzw. -plan für neue Mitarbeiter mit Testaufgaben?	Soziale Kompetenz und Teamzusammensetzung	
21	Existiert ein Ausgliederungskonzept (z.B. Übergabeprotokolle, Wissensdokumentation) für Mitarbeiter mit Testaufgaben die Ihre Organisationseinheit geplant verlassen?		
22	Werden innerhalb Ihrer Organisationseinheit Mitarbeitern regelmäßig die Teilnahme an Fortbildungsmaßnahmen oder Tagungen zum Thema Softwarequalitätssicherung ermöglicht?		
Gesamt RG:			

Anhang 29: Selbsteinschätzung Vergleich

(Hinweis: Aus der Datei „Interview.xlsx“ – Tabellenblatt „Selbsteinschätzung Vergleich“ entnommen und Formatierung angepasst)

Item	Fragestellungen	Merkmale	Ergebnis Reifegradbeschreibung Verfahren	Reifegrad	
				Selbsteinschätzung	Verfahren
1	Existiert für Ihre Organisationseinheit eine konkrete Beschreibung (z.B. Teststrategie, Testkonzept) über die eingesetzten Testprozesse?	Testprozess	Es finden viele unterschiedliche Testaktivitäten statt, wenn passend dazu ein Testkonzept besteht sind diese strukturiert. Nicht-Funktionale Anforderungen werden wahrscheinlich im Anforderungsprozess berücksichtigt.	3	3
2	Welche Testaktivitäten sind Ihnen bekannt und werden Ihres Wissens nach in Ihrer Organisationseinheit regelmäßig eingesetzt?				
3	Werden Ihrer Meinung nach nicht-funktionale Anforderungen im Rahmen des aktuellen Anforderungsprozesses ausreichend berücksichtigt?				
4	Welche Qualitätsmerkmale decken die genutzten Testaktivitäten innerhalb Ihrer Organisationseinheit ab?	Testmanagement	Es werden die meisten Qualitätsmerkmale berücksichtigt. Die Anliegen und Anforderungen der Stakeholder werden im Rahmen der Qualitätssicherung berücksichtigt. Testmetriken finden meistens Anwendung.	2	3
5	Es erfolgte keine Reifegradbewertung mit dieser Frage				
6	Es erfolgte keine Reifegradbewertung mit dieser Frage				
7	Welche Stakeholder wurden Ihrer Meinung nach für die bestehenden Qualitätssicherungsmaßnahmen Ihrer Organisationseinheit berücksichtigt?				

8	Welche Testmetriken finden Ihrer Ansicht nach innerhalb Ihrer Organisationseinheit Verwendung?				
9	Auf welchen Dokumenten oder Arbeitsergebnissen werden Reviews in Ihrer Organisationseinheit durchgeführt?	Reviews	Es werden ein paar Reviews durchgeführt. Werden dazu Checklisten eingesetzt, sind diese einheitlich. Tester werden im Anforderungsprozess ggf. berücksichtigt.	4	2
10	Werden zur Unterstützung der Reviews jeweils einheitliche Checklisten eingesetzt?				
11	Werden innerhalb Ihrer Organisationseinheit Tester an einem Reviewprozess der zu entwickelten Anforderungen beteiligt?				
12	Existiert für Ihre Organisationseinheit eine konkrete Beschreibung über die eingesetzten Fehler- bzw. Abweichungsprozesse?	Fehlermanagement	Es findet ein strukturierter Prozess für Abweichungen statt und diese sind einheitlich beschrieben. Ebenfalls wird ein Fehler-Workflow-Werkzeug eingesetzt.	4	4
13	Wird Ihrer Meinung nach innerhalb Ihrer Organisationseinheit für die Fehler bzw. Abweichungsdokumentation (dem sogenannten Fehlerbericht) ein einheitliches Schema genutzt?				
14	Wird innerhalb Ihrer Organisationseinheit ein Fehler-Workflow-Werkzeug eingesetzt?				
15	Es erfolgte keine Reifegradbewertung mit dieser Frage				
16	Wo und in welcher Form werden in Ihrer Organisationseinheit die Testfälle und Ergebnisse dokumentiert?	Testwerkzeuge und Automatisierung	Testaktivitäten werden ggf. mit Werkzeugunterstützung dokumentiert. Sollte eine Testautomatisierung bestehen, sind einige Testfälle	3	2

17	Wie hoch ist Ihrer Meinung nach aktuell der Anteil an automatisierten Testfällen innerhalb Ihrer Organisationseinheit? (Bitte schätzen Sie)		und ein paar unterschiedliche Testaktivitäten einbezogen.		
18	Welche Testaktivitäten sind innerhalb Ihrer Organisationseinheit Ihres Wissens nach bereits automatisiert?				
19	Es erfolgte keine Reifegradbewertung mit dieser Frage				
20	Existiert für Ihre Organisationseinheit ein Einarbeitungskonzept bzw. -plan für neue Mitarbeiter mit Testaufgaben?	Soziale Kompetenz und Teamzusammensetzung	Im Testbereich existiert ggf. ein Einarbeitungs- und oder Ausgliederungskonzept. Vereinzelt könnte die Teilnahme an QS-Fortbildungsmaßnahmen oder QS-Tagungen möglich sein.	2	2
21	Existiert ein Ausgliederungskonzept (z.B. Übergabeprotokolle, Wissensdokumentation) für Mitarbeiter mit Testaufgaben die Ihre Organisationseinheit geplant verlassen?				
22	Werden innerhalb Ihrer Organisationseinheit Mitarbeitern regelmäßig die Teilnahme an Fortbildungsmaßnahmen oder Tagungen zum Thema Softwarequalitätssicherung ermöglicht?				
Gesamt RG:				3	3

Anhang 30: Gesamtergebnistabelle Auswertungen_Leer.xlsx

Es handelt sich hierbei um dieselbe Ergebnistabelle wie sie in Anhang 5 bis Anhang 27 referenziert wurde, allerdings um eine leere Variante. Diese steht zur weiteren Nutzung zur Verfügung. Diese Exceldatei befindet sich auf der beiliegenden CD-ROM unter „2_Anhänge\Anhang 30_Auswertungen_Leer.xlsx“

Anhang 31: Interview Referatsleiter 04.06.20

Das Interview wurde am 04.06.2020 mit Herrn Kahlert durchgeführt. Er ist der zugehörige Referatsleiter in der die Verfahrenserprobung durchgeführt worden war.

Die Tonaufnahme befindet sich auf der beiliegenden CD-ROM unter „2_Anhänge\Anhang 31_Interview Referatsleiter 04.06.20.mp3“. Die Gesamtlänge beträgt: 26 Minuten und 42 Sekunden.

Anhang 32: Transkription Interview Referatsleiter

04.06.20

Dies ist die Transkription des im Anhang 31 durchgeführten Interview vom 04.06.2020. Die Transkription wurde nach den Regeln von Kuckartz angefertigt:

„Für das vorliegende Projekt haben wir bewusst einfache und schnell erlernbare Transkriptionsregeln⁸ verwendet:

1. *Es wird wörtlich transkribiert, also nicht lautsprachlich oder zusammenfassend. Vorhandene Dialekte werden nicht mit transkribiert.*
2. *Die Sprache und Interpunktion wird leicht geglättet, d. h. an das Schriftdeutsch angenähert. Beispielsweise wird aus „Er hatte noch so'n Buch genannt“ -> „Er hatte noch so ein Buch genannt“.*
3. *Alle Angaben, die einen Rückschluss auf eine befragte Person erlauben, werden anonymisiert.*
4. *Deutliche, längere Pausen werden durch Auslassungspunkte (...) markiert.*
5. *Besonders betonte Begriffe werden durch Unterstreichungen gekennzeichnet.*
6. *Zustimmende bzw. bestätigende Lautäußerungen der Interviewer (Mhm, Aha etc.) werden nicht mit transkribiert, sofern sie den Redefluss der befragten Person nicht unterbrechen.*
7. *Einwürfe der jeweils anderen Person werden in Klammern gesetzt.*
8. *Lautäußerungen der befragten Person, die die Aussage unterstützen oder verdeutlichen (etwa lachen oder seufzen), werden in Klammern notiert.*
9. *Die interviewende Person wird durch ein „I“, die befragte Person durch ein „B“, gefolgt von ihrer Kennnummer, gekennzeichnet (etwa „B4:“).*
10. *Jeder Sprecherwechsel wird durch zweimaliges Drücken der Enter-Taste, also einer Leerzeile zwischen den Sprechern, deutlich gemacht, um die Lesbarkeit zu erhöhen.²⁴⁴*

²⁴⁴ (Kuckartz u. a. 2008: 27–28). Anm.: Fußnote ⁸ aus dem Original: „Weitere Hinweise zu komplexeren Transkriptionsregeln finden sich in Kuckartz (2007: 37 ff.).“

Weiteren gilt: „I“ steht für Interview und „B“ für interviewende Person, in diesem Fall Herr Kahlert. Abgebrochene Sätze enden mit einem „\“ und Wörter die akustisch nicht zugeordnet werden konnten, sind wie folgt dargestellt: (Wort?). Hinweise sind des Verfassers sind in <Hinweis> eingeklammert.

Es folgt die Transkription:

I: So Herr Kahlert die Aufnahme ist gestartet ich würde Sie nochmal bitten, ihre Zustimmung zur Audioaufnahme einfach mündlich zu wiederholen.

B: ja genau ich stimme der Audioaufnahme sicherlich zu.

I: Super. (B: Gerne.)

I: Vielen Dank. Im groben wird es darum gehen, dass wir die Ergebnisse der Interviews, nicht der Interviews der Befragung durchgehen und Ihre Selbsteinschätzung vielleicht auch da gegenüberstellen. Und damit fange ich jetzt quasi an ich habe ihre Zusendung hier vor mir liegen, was Sie als Selbsteinschätzung gesendet hatten. Erstmal (B: Ja.) danke für die Ausfüllung an der Stelle. Und jetzt kann ich, jetzt würde ich ihnen quasi per Mail, die echten Ergebnisse zusenden. Und dann gehen wir quasi Merkmal für Merkmal mal darauf ein, wie quasi die Ergebnisse aus, auf Basis der Antworten der Befragten quasi entstanden waren. Und dann können wir quasi über jedes Merkmal einzugehen, wo sie glauben, dass Sie Unterschiede sehen oder sowas.

B: ja gut können wir so.

Wittsiepe: ich schicke ihn Das jetzt einfach per Mail rüber. Da das ganze Ja nicht persönlich stattfindet/

01:21 <wegen technischer Probleme entstand hier ein murmeln und es sind vereinzelt laute Gedanken des Interviewers zu hören. Diese wurde nicht transkribiert.> 02:08

B: Irgendwie klappt es grad mit der E-Mail nicht. Ich lese Ihnen die einfach vor. Ich glaube, dass es am einfachsten. Dann haben wir auch den Audioteil. Also, ich kann Ihnen das im Nachgang natürlich auch noch einmal optisch zukommen lassen.

I: Ja, ne dann machen wir das so, es ist mir gleich.

B: Okay, also für Merkmal Testprozesse hatte ich ihn ja die Items nochmal aufgelistet, wie sie im Fragebogen zum Einsatz gekommen sind. Also Die Fragestellungen konkret. (I: (Genau?)). Sie hatten dort als Selbsteinschätzung drei, Reifegrad drei angegeben. Und passend dazu ist das Ergebnis auch drei gewesen, an dieser Stelle.

B: OK.

I: Das, also was die Befragten dort wieder\ also im Grunde trifft sich da anscheinend Ihre (B: Meine Einschätzung\ Genau (B: Ergebnisse.) Genau. Zu jedem Reifegrad gibt es quasi ein, eine Beschreibung, die würde ich Ihnen jetzt eben laut vorlesen. Und dann können wir auch darüber reden, wo sie glauben, ob es da auch Unterschiede gibt am Rahmen dieser Beschreibung. (B: Ja). Die Beschreibung ist so ein bisschen standardisierte kann die kann quasi nicht, jede Einzelheit im Grunde wiedergeben, aber sollte ungefähr treffen. Also für Reifegrad drei im Bereich Merkmal Testprozesse ist die Beschreibung. Es finden viele unterschiedliche Testaktivitäten statt. Wenn passend dazu ein Testkonzept besteht, sind diese strukturiert. Nicht-funktionale Anforderungen werden wahrscheinlich im Anforderungsprozess berücksichtigt. Das war im Grunde die Beschreibung. Und jetzt werden, wäre meine konkrete Frage an Sie. Genau, wie schätzen Sie die erreichten Ergebnisse der Merkmale und des Verfahrens im Vergleich zu ihrer eigenen Einschätzung ein? Glauben Sie, das trifft ihre Selbsteinschätzung auch in der Beschreibung?

B: Würde ich jetzt erst mal sagen Ja. (I: (leise) OK).

I: Ja gut, dadurch das Ergebnis sich getroffen haben. Ergibt sich hier sicherlich auch nicht so viel Hinterfragungsbedarf an der Stelle. Okay, Super, das freut mich. Dann würden wir auf das nächste Merkmal eingehen Testmanagement. Auch hier habe ich die Items Ihnen ja noch einmal aufgelistet. Zwei sind nicht in die Reifegradbewertung mit eingeflossen. Vielleicht für Sie zum Hintergrund die dienen dazu, dass wenn man von den Ergebnissen später Maßnahmen ableiten will, dann so ein bisschen mehr Interpretationsraum hat und ein bisschen mehr Futter hat für möglich Maßnahmen einfach, deswegen waren da zwei Fragen, die nicht direkt in die Reifegradbewertung eingeflossen sind. (B: ok, ja verstehe). Genau sie hatten in der Selbsteinschätzung, einen Reifegrad von zwei angegeben. Und das Ergebnis war auch hier drei an der Stelle.

B: (Fragendes?) Okay.

I: Genau, also ein wenig besser zu dem, was Sie eingeschätzt haben, was ein gutes Zeichen ist würde ich behaupten.

B: (Lachend) andersrum wäre es schlechter.

I: Ja. Die Reifegradbeschreibungen, die ich hier herausgearbeitet habe, lautet wie folgt. Es werden die meisten Qualitätsmerkmale berücksichtigt. Die Anliegen und Anforderungen der Stakeholder werden im Rahmen der Qualitätssicherung berücksichtigt. Und Testmetriken finden meistens Anwendung.

B: Ja. Ja, da habe ich das war meinem Testmetriken, warum ich wahrscheinlich da niedriger bewertet habe. Weil ich denke bei\ wenn ich jetzt das Testmetriken mich beziehe, dann habe ich ja wirklich messbare Ergebnisse, die ich auch wirklich ganz sauber, sauberes Messergebnisse, sondern es ist eine messbare Testung. Das trifft jetzt nicht unbedingt bei allen Tests zu, also wenn (ich?) jetzt zum Beispiel mal einen Performance Test hätte, dann habe ich eine Antwort-Zeit-Verhalten, Das kann ich genau messen. Diese Schwierigkeit habe ich jetzt in der Testdurchführung, dieser fachliche Test die wir hier teilweise durchführen gesehen. Und deshalb habe ich ja die Messbarkeit ist da nicht so einfach zugreifen. Deshalb bin ich da, mit der Einstufung runtergegangen. Wir haben ja keine automatisierte Testung und von daher hatte ich gedacht, da sind wir vielleicht noch nicht ganz bei gut. Das war meine Einschätzung dazu. Die anderen Punkte. Stakeholder, denke ich, da berücksichtigen wir alle in der Organisationseinheit. Da sind wir, glaube ich schon auf einem guten Weg. Da achten wir (immer?) sehr stark darauf, dass wir auch die richtigen Steakholder beteiligt haben. Alle mit einbinden in irgendeiner Art und Weise. Das wir wirklich auf die betroffenen dann auch wirklich beziehen der Testung und nicht irgendwas testen, das nachher keiner, wirklich, dann, ja (nur?) die Dinge testen, die dann möglicherweise gar nicht gebraucht werden. Ja

I: Ja, super. Aber das, also ihr braucht, wenn ich das richtig verstanden habe. Sie glauben auf jeden Fall, dass bei den Testmetriken eigentlich noch Aufholbedarf wäre an der Stelle. So habe ich das\

B: Ja, wenn man jetzt sauberer Messgrößen hat und weiß genau in das erwartete Ergebnis, oder nein bei erwarteten Ergebnissen muss ja nicht sein. Aber dass ich wirklich genaue Messung, dass der Test quasi wie eine Messung auch durchgeführt werden kann. Das ich den messbar mein Ergebnis habe. Das, bei mir ist da immer sehr viel Variabilität ist bei uns aber drin. Man kann nicht immer ganz genau sagen, das. Gut man sagt schon das erwarte ich, dass ist das Ergebnis sicher, ist nicht erzielt worden. Ich mache, ich messe auch die Anzahl der Defekte. Ich messe auch, wie viel Abweichung habe ich usw. Aber es noch nicht so konkret, dass ich als sagen würde, das wäre schon optimal. An der Stelle.

I: okay, gut, ja, das ist ja nachvollziehbar. An der Stelle. Okay, dann würde ich zum nächsten Merkmal übergehen. Reviews. Hier waren es drei Items die dazu geführt haben zu der Bewertung. Sie hatten eine Selbsteinschätzung vom Reifegrad vier eingetragen und im Ergebnis da zwei rausgekommen. (B: oh). Das ist erst mal das Ergebnis, so die Beschreibung, die da, die sich quasi von der Stufe zwei ableiten würde, wäre. Es werden ein paar Reviews durchgeführt, werden dazu Checklisten eingesetzt. Sind diese einheitlich. Tester werden im Anpassungsprozess gegebenenfalls berücksichtigt. Wie erklären Sie, wie würden Sie sich den Unterschied erklären an der Stelle?

B: Naja, da ist vielleicht dann doch noch mal ein bisschen Aufarbeitungsarbeit in der Testgruppe notwendig. Also wir haben, die Testergebnisse werden bei uns sehr ordentlich und sauber dokumentiert. Und wir führen nach den Testungen auch Reviews durch. Wir führen nicht nur nach den Testungen Reviews lessons learned durch, sondern auch wöchentlich zu jeder Testung führen wir Statusreviews durch wo dann die Themen besprochen werden. Vielleicht liegt es daran, dass immer die gleichen diejenigen sind, die sich hier daran beteiligen, die dann diese dieses Thema Testung da vertreten in diesen Reviews. Das haben wir jetzt schon mal ein bisschen geändert und haben da auch neue Kollegen. Es war aber noch bevor wir die Umfrage gestartet haben, mal getan, weil das wäre dann auch gleich, ja, nach

folgender Verbesserungsvorschlag Verbesserungsidee. Erklären kann ich mir wirklich nur, dass die Kollegen nur Meinung sind, sie testen und es wird nicht entsprechend gereviewt. Oder sie haben es falsch verstanden und denken, dass, weil wenn die Änderungen dann hier folgen, dass jetzt mal ein wegen der Dienstleister. Bei uns ITZ-Bund. Dann nicht reagiert, dass das vielleicht weil es, oft werden Defekte einfach nicht behoben. So muss man es sagen, dass dann so verstanden haben, dass das vielleicht zum Ärgernis führt. Ich teste hier, ich stelle Fehler fest, und ich gebe sie ja halt auch in quasi entsprechend weiter, aber es wird nicht korrigiert. Es kommt nicht zur Verbesserung. Vielleicht, dass das, so ein bisschen da das Bauchgefühl mitgespielt hat, dass dadurch die, die Bewertung bei Reviews schlechter ausgefallen ist. Aber von unserer Sache her führen wir wirklich wie gesagt Reviews durch. Glaube ich auch im Moment in ausreichender Zahl. Wir müssen uns vielleicht noch ein bisschen, bessere in der Kommunikation in den Testbereich bringen.

I: Das kann, das kann gut sein. Also ich kann, wenn ich so in die Auswertungen die Detailauswertung gucke, kann ich bestätigen, dass quasi die Anzahl der Reviews durchaus gerechtfertigt ist, was die meisten angegeben haben tatsächlich. Ein kleiner, also was ein wenig gemacht wird, sind so technische Reviews tatsächlich oder Reviews technischer Dokumente. Um es genau zu sagen. Da gab es eine Einschränkung und im Bereich der einheitlichen Prozesse für Reviews. Also zum Beispiel das Reviews einheitlich durchgeführt werden mit Checklisten beispielsweise. So wirklich immer die gleichen Reviews\

B: Ok, Ja, Nein, die gibt es aber. Die führen wir anhand einer Checkliste durch. Das sind auch immer die Gleichen, die wir da einheitlich nutzen. Wir machen die auch im Anforderungsmanagement, im Testen, da es sind also auch alle beteiligt, selbst Schulung und Support sind mit an diesen Reviews beteiligt. Also das müsste, theoretisch hätte das besser ausfallen müssen. Aber vielleicht ist das ein bisschen, mangelnde, Kommunikation, bei den Testers (teils?).

I: Das kann gut sein. Ich kann konkret sagen, dass keiner der Teilnehmer quasi mit Ja gestimmt hat. Auf die Frage wären zur Unterstützung der Reviews jeweils einheitliche Checklisten eingesetzt.

B: Ok.

I: Aber das kann verschiedene Gründe haben. Interessant ist halt das hier wirklich. Eine Abweichung ist, an dieser Stelle. Okay, gehen wir weiter, wenn sie, wenn es für sie okay ist. (B: ja) Das Merkmal Fehlermanagement. Hier hatten sie eine Selbsteinschätzung von Reifegrad vier gewählt, und das ist auch das Ergebnis gewesen, an der Stelle. Die Items. Also ich lese Ihnen mal die Beschreibung vor. Es findet ein strukturierter Prozess für die Abweichung statt und diese sind einheitlich beschrieben. Ebenfalls wird einen Fehler-Workflow-Werkzeug eingesetzt.

B: Wir hatten mal vorgestellt dieses HP ALM oder Microsoft Focus ALM heißt jetzt MF ALM, nein Micro Focus ALM. Damit sind wir, damit deckt man natürlich einen großen Teil ab. Das ist ja so teil\ ist zwar teilautomatisiert, es ist nicht wirklich ein automatisiertes Verfahren der Testung. Aber, was jetzt Dokumentationen betrifft, Fehlermanagement, das (Ford?)-Management mit dem Dienstleister da, da sind wir sehr, sehr gut Tool-Unterstützt. Und da haben wir auch eine ziemlich lückenlose fehlerfreie Abweichungs-Dokumentation. Und deshalb glaube ich, da haben auch alle wahrscheinlich eher das Ganze positiv gesehen. Das ist natürlich eine große Stütze, jetzt für so einen Tester, wenn er sich dann wirklich zu tun viel Arbeit erspart. Ja auch dann, weil das die Dokumentation komplett darüber folgen kann.

I: Ja, das war ein in, so ziemlich auch meine Schlüsse die ich daraus gezogen hat tatsächlich. Deswegen freut mich auch die Bestätigung ihrerseits. Also soweit passte das dann. Man muss dazu sagen, dass sie Reifgrade ja nicht unbedingt heißen, dass da etwas gut oder schlecht ist, sondern manchmal kann ein niedriger Reifegrad ja auch gewollt sein. Oder man sagt ja, dass es dieses Level, was wir erreichen wollen. Deswegen muss man sich jetzt auch nicht immer an den einzelnen Stufen aufhängen und sagen, dass es jetzt schlecht oder so. Das sollte, sollte man vielleicht einfach Bewusst sein. Sondern man hat einfach Luft nach oben an der Stelle, wenn eine Bewertung niedriger ausgefallen ist. Das heißt nicht, dass man schlecht in seiner Vorgehensweise ist oder sowas in der Form. Okay. Nächstes Merkmal. Testwerkzeuge und Automatisierung. Hier haben Sie eine Selbsteinschätzung von Reifegrad drei gewählt. Das tatsächliche Ergebnis des Verfahrens war Reifegrads zwei. Ich lese wieder die Beschreibung kurz vor. Testaktivitäten werden gegebenenfalls mit Werkzeug-Unterstützung dokumentiert. Sollte eine Testautomatisierung bestehen, sind einige Testfälle und ein paar unterschiedliche Test Aktivitäten einbezogen. (...) Genau. (B: Ja). Wie haben Sie sich\

B: Also Test\, wenn ich jetzt auf Testautomatisierung schaue. Weil wir haben zwar mal, gut das ist aber eigenprogrammiert. Anfänglich mal so kleine Schritte, sind wir in die Testautomatisierung schonmal gegangen. Also baut der Kollege B. Aber ansonsten ist Testautomatisierung natürlich für alle noch völlig Fremdwort. Testwerkzeugs natürlich, wie gesagt unser Tool unser HP ALM oder MF ALM-Tool auch was jetzt Dokumentationen und so weiter betrifft. Aber halt noch nicht in Richtung Automatisierung. Da sind wir natürlich an der Stelle noch nicht soweit. Vielleicht wird sich das normal mit dem neuen Kernsystem was wir demnächst, ja in ein zwei Jahren vielleicht dann einführen, dann verändern. Aber aktuell wie gesagt kann ich mir deshalb vorstellen, dass wir da bei einer zwei eher gelandet sind, aus Sicht der Beteiligten.

I: Genau, ja. Ihre Aussage bestätigt im Grunde auch die Ergebnisse an der Stelle, dass die Tester Automatisierung ist definitiv hier, im Grunde also der Anteil daraus ist eher niedrig ausgefallen an der Stelle. Es gab tatsächlich vereinzelt Teilnehmer, die von einem bisschen Testautomatisierung berichtet haben. Deswegen würde das auch Ihre Einschätzung bestätigen und die Tool-Unterstützung definitiv ist das was, was das hier noch quasi ein Punkt nach oben gebracht hat an der Stelle. Bei den, weil dieses Merkmal der Testwerkzeuge und Automatisierung ab einfach. Okay (B: Ja) Das passt ja gut zusammen. Letzter Punkt letztes, Merkmal zumindest wäre Sozialkompetenz und Teamzusammensetzung. Dort haben Sie sich mit der Selbsteinschätzung auf Reifegrad zwei eingewählt. Und weil das tatsächliche Ergebnis des auch Reifegrad zwei. (B: Ja). Konkrete Beschreibungen dazu wäre im Testbereich existiert gegebenenfalls seien Einarbeitungs- und Ausgliederungskonzept. (B: Ja). Vereinzelt, könnte die Teilnahme an QS Fortbildungsmaßnahmen oder QS Tagungen möglich sein.

B: Ja, also da wie gesagt auch. Die Einschätzung zwei sich auch wirklich so. Wir haben also, das Team an sich, ist eigentlich schon recht weit, so auch von der Qualifizierung her. Weil sie halt schon sehr lange in diesem Geschäft sind und auch dieses Tool mit dem wir arbeiten, also diese Anwendungen die wir in der Testung kriegen alle\ zweimal im Jahr sehr gut kennen und (dadurch?) auch einen großen Erfahrungsschatz haben. Aber das Thema, hier ging es ja darum, ob wir ein Einarbeitungskonzept haben. Und ich habe jetzt gerade aktuell den Fall. Ist natürlich ist auch durch Corona ein bisschen schwierig, jemand einzuarbeiten. Aber wir arbeiten als immer mal neue Kollegen ein. Und, da müsste man daran arbeiten. Das ist also wirklich noch, da gibt es halt keine Möglichkeit. Ist manchmal auch ein Zeitproblem, oder auch Schulungstechnisch jemanden wirklich sauber einzuführen. Das ist Prozess, der dann irgendwann mal läuft, weil sie die Leute so learning-by-doing irgendwann mal so weit sind. Aber man könnte das Ganze verbessern. Auf jeden Fall also, das nicht aktuell, so dass wir da wirklich sagen, wir haben hier ein Szenario von einer Art Schulung oder Einarbeitungsphase. Für jemanden, der jetzt bei uns aktuell neu hinzukommt und weitere Tests durchführen will, der kann dann sicherlich schon irgendwie arbeiten. Aber bis man bei uns mit den Tools dann umgehen kann, das dauert noch meistens. Und das ist auf jeden Fall auch

Verbesserungswürdig würde ich sagen, da haben wir noch nicht so einen hohen Reifegrad erreicht. Ist aber auch nicht ganz leicht zu erreichen. Wie gesagt haben müssen für verschiedene Rahmenbedingungen stimmen. Wir arbeiten auch mittlerweile auch viel in Telearbeit. Das macht es auch dann nicht leichter, jemand einzuarbeiten. Ja, deshalb denke ich, diese zwei an der Stelle. Also die Stimmung in unserem Team selbst dies sicherlich gut, da ist glaube ich, keiner, das ist eher positiv. Aber wie gesagt, was jetzt neu, neue Kollegen trifft, die haben dann erst noch eine größere Hürde.

I: ja, das würde, ja das deckt ja sehr gut auch das ab, was sich ungefähr auch aus den Ergebnissen geschlossen hätte. An der Stelle, Stimmung war ja ein Punkt, den ich nicht im Fragebogen mit aufgenommen habe, bewusst. Weil ich würde mich sehr schwer damit tun, auch persönlich, von der Stimmung des Teams auf den tatsächlichen Reifegrad von Qualitätssicherung zu schließen.

B: Nein, das ist auch einfach nur, das ist Sicherlich kein richtiges Kriterium, um gute Testreifegrad festzustellen, dass (I: Genau). denke ich auch, das sehe ich auch so.

I: Ja, deswegen habe ich mich gerade bei diesem Thema auf die Sachen, auf die prozessgestützten Sachen eher konzentriert, weil das ja (B: Genau) mit, weil das hat den Reifegrad in der Hinsicht ja dann unterstützen würde, quasi. (B: Ja). Okay. Gut, das Gesamtergebnis vielleicht, sicherlich für Sie auch interessant wäre. Also ich, im Grunde wird das Ergebnis jetzt so berechnet und das auch für Sie transparent zu gestalten. Ich nehme alle Reifegrade, auch hier, also das, was sie selbst eingeschätzt haben, alle sechs Punkte und bildet den Mittelwert daraus. Und dann Runde ich halt nach den üblichen Regelungen, also ab 0,5 auf, unter 0,5 ab, sodass kennt man ja. Und daraus würde dann quasi ein Gesamtreifegrad entstehen.

Das ist bei ihrer Selbsteinschätzung drei gewesen und bei den Gesamt\ bei dem Ergebnis des Verfahrens auch drei. (B: Ah ja). Da haben sich die Zahlen dann im Grunde so ausgeglichen. Also auch wenn wir Abweichungen hatten, vielleicht in einzelnen Punkten, hat sich das mehr oder weniger wieder ausgeglichen.

B: Ja, erfreulich.

I: Ja, ich freue mich auch darüber. Die Beschreibung\ Es gibt jetzt quasi für die Gesamteinschätzung nicht nochmal eine eigene Beschreibung, sondern die Teile, die ich Ihnen gerade vorgelesen habe, bilden die Gesamteinschätzung, dass es so (B: Ok) das Ziel meines Verfahrens, weil ich, dass einfach dann auf viel näher dran ist an der ganzen Thematik als wie wenn ich jetzt nochmal eine High Level Einschätzungen an der Stelle gemacht hätte, die vielleicht nichts mehr mit der Realität zu tun hat. Interessant auf jeden Fall. Super, dann wäre meine nächste Frage an Sie. Außer Sie haben jetzt noch Fragen bis zu diesem Punkt oder Anmerkungen.

B: Nein, das ist soweit alles klar für mich. Keine Fragen.

I: Okay. Welche Ergebnisse überraschten sie am meisten? Von denen, was wir gerade besprochen haben.

B: Ja, das Thema Reviews hat mich jetzt halt überrascht, dass da die Differenz doch größer war. Ich kann mir jetzt im Nachgang vorstellen, woran es liegt. Aber überrascht hat es ich mich halt zunächst dann schon, weil wir eben da eigentlich den Wünschen was das Review betrifft, entsprechend aber anscheinend nicht überall so gesehen, vielleicht auch falsch verstanden wurde, kann auch schon ein, Ja.

I: Das würde ja auch Hinweise dafür geben, dass man vielleicht da auch noch mal Nachsteuern kann an der Stelle, an diesem Prozess (B: Genau) einfach. Also, das Werkzeug liefert ja nur das ab, was die Leute denken. Das ist einer der großen Vorteile, an der Stelle. Gibt es weitere Punkte, die Sie überrascht haben?

B: Nein, die anderen haben mich alle nicht überrascht. Da bin ich, glaub ich soweit hätte ich. Das konnte ich jetzt auch so nachvollziehen, da war ich jetzt nicht überrascht.

I: Danke dafür. Der letzte Teile des Interviews stellt quasi, würde jetzt so ein bisschen in den explorativen Bereich gehen, dass wir jetzt, dass Sie jetzt die Ergebnisse im Grunde für sich mal so vor Augen halten und dann überlegen, welche Schlüsse oder möglichen weitere Schritte können Sie aus den Ergebnissen dieses Verfahren ziehen? Also im Grunde haben Sie ja schon ein zwei Punkte angedeutet, in dem Verbesserungsbedarf besteht und meinten. Die Idee setzt im Grunde, würden, können Sie mit den Ergebnissen daraus im Grunde weitere Schritte oder Schlüsse ziehen und welche wären das?

B: Ja, also das wäre, auf jeden Fall einmal das Thema Testwerkzeug und Automatisierung, da vielleicht hinzukommen, was natürlich auch immer von technischen Umgebungen und Gegebenheiten abhängt, aber da in Richtung automatisierte Tests auch vielleicht weiterzukommen. Da hoffen wir, dass der neue Dienstleister uns da auch was mit anbietet. Das wäre ein Schluss, denn ich auch hier rausziehen würde. Dann wie gesagt nochmal eine Kommunikation Richtung was, wie sieht es aus mit unseren regelmäßigen Reviews führen. Da vielleicht einfach nochmal eine größere Bandbreite reinzubringen, an den Mitarbeitern an diesem Thema. Ja das wären so für mich Schlüsse die ich daraus, glaube ich erst mal ziehen würde. Ansonsten, schau jetzt gerade noch einmal über die Liste drüber, die anderen Punkte sind. Ja, also man kann natürlich an dem Thema. Ja, sicher das habe ich jetzt vergessen, das Thema, hatte ich auch gerade gesagt. Entwicklung einer Ausbildungsinstanz, dass man so eine Art Einarbeitungskonzept nochmal für den Bereich Testung erstellt. Das man wirklich dann anders die Kollegen heranführen kann, wenn jetzt jemand zum Testbereich wechselt.

Sich dafür qualifizieren möchte. Also, da ist man, da ist auch noch ein Handlungsbedarf. Das würde ich auch als Schluss daraus mitnehmen.

I: Okay, es waren jetzt genau die drei Punkte, in denen quasi eine zweier Bewertung erfolgte. Ein logischer Schluss und nachvollziehbar auch von meiner Seite aus. Super. Okay, das war's mit meinen Fragestellungen. Wir sind ziemlich gut durchgekommen. Ich würde die Audioaufnahme jetzt an diesem Punkt beenden.

B: Ja.

I: Ok. Moment.

Anhang 33: Lebenslauf

Persönliche Angaben			
Geburtsdatum:	28.07.1988	Geburtsort:	Mühlheim a. d. Ruhr
Familienstand:	Verheiratet	Staatsangehörigkeit:	Deutsch
Studium			
RFH Köln	2018 – 2020: Master of Science, Wirtschaftsinformatik		
RFH Köln	2011 - 2015: Bachelor of Arts, Business and Informations Systems (Wirtschaftsinformatik), in Abendform		
Beruflicher Werdegang			
Bundesverwaltungsamt	2017/08 – dato: IT-Projektmanager/Testmanager		
EVO Payments International GmbH	2017/02 – 2017/07: Test-Experte		
Andagon GmbH	2016 – 2017: Berater für Qualitätssicherung und Testautomatisierung		
DB Kommunikationstechnik GmbH	2010 – 2015: IT-Systemelektroniker im Fachbereich der Einbruch- und Brandmeldetechnik		
Besondere Kenntnisse			
Sehr gute Kenntnisse in:			
<ul style="list-style-type: none"> • Projektmanagement, Testmanagement, ISTQB, Scrum • Tools: HP ALM / QC, Jira, Aqua, Confluence, Zephyr • Anlagen- und Sicherheitstechnik, Telekommunikationsanlagen • Computerhardware, IT-Netze, Servicelandschaften • MS-Office, MS-Projekt, MS-Windows, Lotus Notes, Sharepoint, ThinkCell • Java, PHP, Datenbanken (SQL) 			
Gute Kenntnisse in:			
<ul style="list-style-type: none"> • Elektrotechnik, Informatik, Internet of Things • Testautomatisierung (ReadAPI, JMeter, HP LoadRunner, HP UFT) • Versionsmanagementtools (z.b. Perforce, Subversion, GitHub) • Linux Debian Derivate, SLES, Linux/Unix Shell Scripting • Android, iOS, Windows Phone • Kalkulationen, Angebotserstellung, Aufmaß-Ermittlungen 			
Weitere Kenntnisse in: HTML, JavaScript, Photoshop, Graylog			
Qualifikationen			
<ul style="list-style-type: none"> • Basiszertifikat im Projektmanagement der GPM – Deutsche Gesellschaft für Projektmanagement e. V. • ISTQB Certified Tester – Foundation Level • DAAD B2 – Englisch • Diverse Qualifikationen im Bereich der EMA/BMA-Sicherheitstechnik (Honeywell / Esser) 			

Literaturverzeichnis

- Afzal, Wasif; Alone, Snehal; Glocksien, Kerstin; u. a. (2016): „Software test process improvement approaches: A systematic literature review and an industrial case study“. In: *Journal of Systems and Software*. 111, S. 1–33, doi: 10.1016/j.jss.2015.08.048.
- Amaral, Luís; Bazzana, Gualtiero; Pe'er, Tal; u. a. (2017): „ISTQB® Effectiveness Survey 2016-17“. Abgerufen am 11.06.2020 von https://www.istqb.org/documents/istqb_survey_2016-2017.pdf.
- Avci, Oral (2008): „Aus Fehlern in der Softwareentwicklung lernen Wie durch Fehleranalysen die Prozesse der Anforderungsanalyse und der Qualitätssicherung verbessert werden können“. Köln: Köln.
- Bath, Graham; Veenendaal, Erik van (2014): *Improving the test process: implementing improvement and change--a study guide for the ISTQB expert level module*. 1st edition. Santa Barbara, CA: Rocky Nook Inc.
- Bedenk, Birgit; Geißler, Holger; Hennecke, Volker; u. a. (2004): „Leitfaden für Online-Mitarbeiterbefragungen“. Abgerufen am 11.06.2020 von https://www.bvm.org/fileadmin/user_upload/Verbandsdokumente/Standesregeln/LF_2004_Online-MAB.pdf.
- Bitkom e.V. (2018): „Blockchain wird zu einem Top-Thema in der Digitalwirtschaft“. *Blockchain wird zu einem Top-Thema in der Digitalwirtschaft*. Abgerufen am 11.06.2020 von <https://www.bitkom.org/Presse/Presseinformation/Blockchain-wird-zu-einem-Top-Thema-in-der-Digitalwirtschaft.html>.
- Bitkom e.V.; EITO (Hrsg.) (2020): „Umsatz mit Software in Deutschland in den Jahren 2007 bis 2020“.
- Booth, Wayne C.; Colomb, Gregory G.; Williams, Joseph M.; u. a. (2016): *The craft of research*. Fourth edition. Chicago: The University of Chicago Press (Chicago guides to writing, editing, and publishing).
- Bundesrechnungshof (2010): *Bemerkungen - Weitere Prüfungsergebnisse Nr. 12 „Nachteilige Vertragsbedingungen und schleppende Rückabwicklung einer IT-Beschaffung erschweren Rückforderung von 42 Mio. Euro“*. Bonn: Bundesrechnungshof.
- Bundesrechnungshof (2008): *Bericht an den Beauftragten der Bundesregierung für Informationstechnik und an den Rat der IT-Beauftragten nach § 88 Abs. 2 BHO über die Querschnittsprüfung zu der Sicherung der Softwarequalität in IT-Vorhaben der Bundesverwaltung*. (Nr. VII 2-2006-0162) Bonn: Bundesrechnungshof.
- Bundesrechnungshof (2018): *Bericht an den Haushaltsausschuss des Deutschen Bundestages nach § 88 Abs. 2 BHO über die Prüfung des Projekts „Rollenbasierte Oberflächen (ROBASO)“ und weiterer Projekte der Bundesagentur für Arbeit*. (Nr. VI 1-2017-1087 / VI 1-2016-0560) Bonn: Bundesrechnungshof.

Bundesrechnungshof (2019): *Bericht an den Haushaltsausschuss des Deutschen Bundestages nach § 88 Abs. 2 BHO zum aktuellen Stand und zu den Fortschritten des Zusammenwirkens von Bund und Ländern im Vorhaben KONSENS*. (Nr. VIII 1-2017-1195/3) Bonn: Bundesrechnungshof.

Bundesverwaltungsamt (2018): „Infoblatt zum Bundesverwaltungsamt“. Abgerufen am 11.06.2020 von https://www.bva.bund.de/SharedDocs/Downloads/DE/Presse/Informationsschriften/Infoblaetter/infoblatt_bva.pdf?__blob=publicationFile&v=4.

Bundesverwaltungsamt (o. J.): „Lebenslauf Christoph Verenkotte“. Abgerufen am 11.06.2020 von https://www.bva.bund.de/DE/Das-BVA/Organisation/Prasident-Verenkotte/Lebenslauf_Praesident_Verenkotte.pdf;jsessionid=9016EC74B7E634B37F02F8355397943C.intranet672?__blob=publicationFile&v=8.

Bundesverwaltungsamt (2020): „Organigramm“. Abgerufen am 11.06.2020 von https://www.bva.bund.de/SharedDocs/Downloads/DE/Organigramm.pdf;jsessionid=046BC99A66337ABAE45C325044527031.intranet251?__blob=publicationFile&v=11.

Cap Gemini Sogeti (Firm) (2009): *TPI® next: business driven test process improvement*. 's-Hertogenbosch, Netherlands: UTN Publishers.

Carnegie Mellon University (2011): „CMMI® für Entwicklung, Version 1.3“. Abgerufen am 11.06.2020 von https://resources.sei.cmu.edu/asset_files/WhitePaper/2011_019_001_28795.pdf.

Daigl, Matthias; Glunz, Rolf (2016): *ISO 29119: die Softwaretest-Normen verstehen und anwenden*. Heidelberg: dpunkt.verlag.

Döring, Nicola; Bortz, Jürgen (2016): *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften*. 5. vollständig überarbeitete, aktualisierte und erweiterte Auflage. Berlin Heidelberg: Springer (Springer-Lehrbuch).

Ewijk, Alexander van; Neumann, Julia; Sogeti Nederland BV (Hrsg.) (2011): *TPI Next: geschäftsbasierte Verbesserung des Testprozesses*. 1. Aufl. Heidelberg: dpunkt.verl.

Ewusi-Mensah, Kweku (1997): „Critical issues in abandoned information systems development projects“. In: *Communications of the ACM*. 40 (9), S. 74–80, doi: 10.1145/260750.260775.

Garousi, Vahid; Felderer, Michael; Hacaloğlu, Tuna (2017): „Software test maturity assessment and test process improvement: A multivocal literature review“. In: *Information and Software Technology*. 85, S. 16–42, doi: 10.1016/j.infsof.2017.01.001.

Garousi, Vahid; Felderer, Michael; Hacaloglu, Tuna (2018): „What We Know about Software Test Maturity and Test Process Improvement“. In: *IEEE Software*. 35 (1), S. 84–92, doi: 10.1109/MS.2017.4541043.

- German Testing Board e.V. (2018): „Jahresbericht 2018“. Abgerufen am 11.06.2020 von https://www.german-testing-board.info/wp-content/uploads/2019/07/GTB_Jahresbericht2018_190726-1.pdf.
- German Testing Board e.V. (2019): „Organisation“. *Organisation*. Abgerufen am 11.06.2020 von <https://www.german-testing-board.info/gtb/german-testing-board/organisation/>.
- Google Ireland Limited (o. J.): „Google Übersetzer“. *Google Übersetzer*. Abgerufen am 11.06.2020 von <https://translate.google.com/>.
- Haberl, Peter; Vosseberg, Karin; Spillner, Andreas; u. a. (2012): „Umfrage 2011: Softwaretest In Praxis“. Abgerufen am 03.04.2020 von http://www.softwaretest-umfrage.de/2011/Broschuere_Softwaretest_in_der_Praxis.pdf.
- Hevner; March; Park; u. a. (2004): „Design Science in Information Systems Research“. In: *MIS Quarterly*. 28 (1), S. 75, doi: 10.2307/25148625.
- Hoffmann, Dirk W. (2013): *Software-Qualität*. 2., aktualisierte und korrigierte Aufl. Berlin: Springer Vieweg (eXamen.press).
- Hussy, Walter; Schreier, Margrit; Echterhoff, Gerald (2013): *Forschungsmethoden in Psychologie und Sozialwissenschaften für Bachelor: mit 23 Tabellen*. 2., überarbeitete Auflage. Berlin Heidelberg: Springer (Springer-Lehrbuch).
- ISO 29119-1 (2013): *Software and systems engineering Software testing Part 1: Concepts and definitions*. IEEE doi: 10.1109/IEEESTD.2013.6588537.
- ISO 29119-4 (2015): *29119-4-2015 - ISO/IEC/IEEE International Standard - Software and systems engineering--Software testing--Part 4: Test techniques*. Place of publication not identified: IEEE.
- iso25000.com (2019): „ISO/IEC 25010“. *ISO/IEC 25010*. Abgerufen am 11.06.2020 von <https://iso25000.com/index.php/en/iso-25000-standards/iso-25010?limit=3&limitstart=0>.
- ISTQB (2019): „Facts & Figures“. *Facts & Figures*. Abgerufen am 11.06.2020 von <https://www.istqb.org/about-us/facts-figures.html>.
- ISTQB (2020): „Vision & Mission“. *Vision & Mission*. Abgerufen am 11.06.2020 von <https://www.istqb.org/about-us/vision-mission.html>.
- ISTQB; Black, Rex; McKay, Judy; u. a. (2012): „Certified Tester Advanced Level Syllabus Test Manager - Version 2012“. International Software Testing Qualifications Board (ISTQB).
- ISTQB; Olsen, Klaus; Parveen, Tauhida; u. a. (2018): „Certified Tester Foundation Level Syllabus - Version 2018“. International Software Testing Qualifications Board (ISTQB).
- Kallus, Konrad Wolfgang (2016): *Erstellung von Fragebogen*. 2., aktualisierte und überarbeitete Auflage. Wien: Facultas (UTB Schlüsselkompetenzen).
- Kannenbergh, Axel (2017): „60 Millionen Euro versenkt: Bundesagentur für Arbeit stoppt IT-Projekt ROBASO“. *60 Millionen Euro versenkt: Bundesagentur für Arbeit stoppt IT-*

Projekt ROBASO. Abgerufen am 11.06.2020 von <https://www.heise.de/newsticker/meldung/60-Millionen-Euro-versenkt-Bundesagentur-fuer-Arbeit-stoppt-IT-Projekt-ROBASO-3627866.html>.

Koomen, Tim; Pol, Martin; Sogeti Nederland B.V. (1998): „IMPROVEMENT OF THE TEST PROCESS using TPI®“. Abgerufen am 11.06.2020 von https://itq.ch/pdf/tpi/tpi_uk.PDF.

Kuckartz, Udo; Dresing, Thorsten; Rädiker, Stefan; u. a. (2008): *Qualitative Evaluation: der Einstieg in die Praxis*. 2., aktualisierte Auflage. Wiesbaden: VS, Verlag für Sozialwissenschaften.

Linetzki, Alon; Banerjee, Ritendra; Archer, Debbie (2018): „ISTQB® Worldwide Software Testing Practices Survey 2017-18“. Abgerufen am 11.06.2020 von <https://www.istqb.org/references/surveys/istqb%C2%AE-worldwide-software-testing-practices-survey-2017-18.html>.

Metzger, Armin; Münzel, Jörg; Simon, Frank; u. a. (2013): „Der ISTQB-Expert-Level: Testmanagement in modernen Systemen“. *OBJEKTSpektrum*. S. 58–61.

Micro Focus International plc. (2019): „Micro Focus ALM/Quality Center - Datasheet“. Abgerufen am 11.06.2020 von https://www.microfocus.com/media/datasheet/micro_focus_alm_quality_center_ds.pdf.

Moosbrugger, Helfried; Kelava, Augustin (Hrsg.) (2012): *Testtheorie und Fragebogenkonstruktion: mit 66 Abbildungen und 41 Tabellen*. 2., aktualisierte und überarbeitete Auflage. Berlin Heidelberg: Springer (Springer-Lehrbuch).

Myers, Glenford J.; Sandler, Corey; Badgett, Tom (2012): *The art of software testing*. 3rd ed. Hoboken, N.J: John Wiley & Sons.

Otter, Reinder; Margré, Wouter; Langebroek, Tom; u. a. (2014a): „TPI NEXT® Test maturity - Cluster oriented (xls) Version 2.1.1“. Abgerufen am 11.06.2020 von https://www.tmap.net/sites/tmap/files/2019-08/TPI_NEXT_tool_v2.1.2_0%20clusters_0_1.xls.

Otter, Reinder; Margré, Wouter; Langebroek, Tom; u. a. (2014b): „TPI NEXT® Test maturity - Full, not oriented, list (xls) Version 2.1.1“. Abgerufen am 11.06.2020 von https://www.tmap.net/sites/tmap/files/2019-08/TPI_NEXT_tool_v2.1.2_0%20one%20list_0.xls.

Pierenkemper, Sarah; Institut der deutschen Wirtschaft Köln e.V. (Hrsg.) (2016): „Handlungsempfehlung Mitarbeiterbefragung“.

Porst, Rolf (2014): *Fragebogen: ein Arbeitsbuch*. 4., erweiterte Auflage. Wiesbaden: Springer VS (Studienskripten zur Soziologie).

Reichertz, Jo (2016): *Qualitative und interpretative Sozialforschung: eine Einladung*. Wiesbaden: Springer VS (Studentexte zur Soziologie).

Ron Swinkels (2000): *A comparison of TMM and other Test Process Improvement Models*. Frits Philips Institute.

- Schubert, Klaus (Hrsg.) (2005): *Handwörterbuch des ökonomischen Systems der Bundesrepublik Deutschland*. 1. Aufl. Wiesbaden: VS, Verl. für Sozialwiss.
- Simon, Frank; Fischer, Manuel; Vosseberg, Karin; u. a. (2016): „Management-Summary der GTB Softwaretestumfrage 2015/2016“. Abgerufen am 11.06.2020 von http://www.softwaretest-umfrage.de/2016/GTB_Survey_Management_Summary_KVAS_Neu.pdf.
- Simon, Herbert A. (1996): *The sciences of the artificial*. 3rd ed. Cambridge, Mass: MIT Press.
- Sneed, Harry M. (1987): *Software-Management*. Köln: Müller (Online-DV-Praxis).
- Sogeti Deutschland GmbH (2020): „TPI NEXT® - Testprozessverbesserung“. *TPI NEXT® - Testprozessverbesserung*. Abgerufen am 11.06.2020 von <https://www.sogeti.de/dienstleistungen/software-qualitaessicherung-und--test/Testprozessberatung/TPINEXT/>.
- sogeti.com (2020): „Test Process Improvement (TPI)“. *Test Process Improvement (TPI)*. Abgerufen am 11.06.2020 von <https://www.tmap.net/building-blocks/test-process-improvement-tpi>.
- Spillner, Andreas; Linz, Tilo (2019): *Basiswissen Softwaretest: Aus- und Weiterbildung zum Certified Tester Foundation Level nach ISTQB-Standard*.
- Spillner, Andreas; Roßner, Thomas; Winter, Mario; u. a. (2014): *Praxiswissen Softwaretest - Testmanagement: Aus- und Weiterbildung zum Certified Tester - Advanced Level nach ISTQB-Standard*. 4., überarbeitete und erweiterte Auflage. Heidelberg: dpunkt.-Verlag.
- Spillner, Andreas; Vosseberg, Karin; Winter, Mario (2016): „Forschungs-Survey Softwaretestumfrage 2015/2016“. Abgerufen am 11.06.2020 von http://www.softwaretest-umfrage.de/2016/GTB_Survey_Forschung_Summary_V1.pdf.
- Spillner, Andreas; Vosseberg, Karin; Winter, Mario (2017a): „Technischer Report Umfrage 2016 Softwaretest in Praxis und Forschung“. Abgerufen am 11.06.2020 von http://www.softwaretest-umfrage.de/2016/Technischer_Report_Umfrage_2016.pdf.
- Spillner, Andreas; Vosseberg, Karin; Winter, Mario (2017b): „Umfrage 2016: Softwaretest In Praxis und Forschung“. Abgerufen am 11.06.2020 von http://www.softwaretest-umfrage.de/2016/Softwaretest_Broschuere_2016.pdf.
- Statistisches Bundesamt (2018): „Anteil der Computernutzer* in Deutschland von 2004 bis 2017“.
- Trittmann, Ralph (2004): *Wirtschaftlichkeit von Wissenstransfers in der Softwareentwicklung: ein kostenorientiertes Gestaltungskonzept*. Aachen: Shaker (Berichte aus der Wirtschaftsinformatik).
- van Veenendaal, Erik (2020): „ERIK VAN VEENENDAAL internationally recognized testing expert“. *ERIK VAN VEENENDAAL internationally recognized testing expert*. Abgerufen am 11.06.2020 von <http://www.erikvanveenendaal.nl/en/>.

van Veenendaal, Erik (2018): „Test Maturity Model integration (TMMi®) Guidelines for Test Process Improvement Release 1.2“. Abgerufen am 11.06.2020 von <https://tmmi.org/tm6/wp-content/uploads/2018/11/TMMi-Framework-R1-2.pdf>.

Vosseberg, Karin; Spillner, Andreas; Winter, Mario; u. a. (2013): „Technischer Report Umfrage 2011 Softwaretest in Praxis“. Abgerufen am 11.06.2020 von <http://www.softwaretest-umfrage.de/2011/TechnischerReport130317.pdf>.

Wittsiepe, Mike; Kahlert, Detlef (2020): *Interview Referatsleiter 04.06.20*.

Eigenständigkeitserklärung

Hiermit erkläre ich an Eides statt, dass ich die vorliegende Arbeit selbstständig und ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe.

Die aus anderen Quellen direkt oder indirekt übernommenen Daten und Konzepte sind unter Angabe der Quelle gekennzeichnet. Dies gilt auch für Quellen aus eigenen Arbeiten.

Ich versichere, dass ich diese Arbeit oder nicht zitierte Teile daraus vorher nicht in einem anderen Prüfungsverfahren eingereicht habe.

Mir ist bekannt, dass meine Arbeit zum Zwecke eines Plagiatsabgleichs mittels einer Plagiatserkennungssoftware auf ungekennzeichnete Übernahme von fremdem geistigem Eigentum überprüft werden kann.

Ich versichere, dass die elektronische Form meiner Arbeit mit der gedruckten Version identisch ist.

Köln, 17.06.2020



Ort, Datum, Unterschrift