



Frank Witte

Metriken für das Testreporting

Analyse und Reporting
für wirkungsvolles Testmanagement

EBOOK INSIDE

 Springer Vieweg

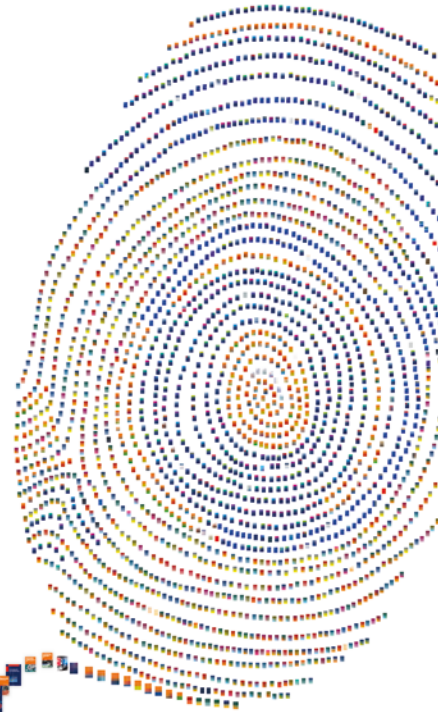


Metriken für das Testreporting

Lizenz zum Wissen.

Sichern Sie sich umfassendes Technikwissen mit Sofortzugriff auf tausende Fachbücher und Fachzeitschriften aus den Bereichen: Automobiltechnik, Maschinenbau, Energie + Umwelt, E-Technik, Informatik + IT und Bauwesen.




Exklusiv für Leser von Springer-Fachbüchern: Testen Sie Springer für Professionals 30 Tage unverbindlich. Nutzen Sie dazu im Bestellverlauf Ihren persönlichen Aktionscode **C0005406** auf www.springerprofessional.de/buchaktion/



Jetzt
30 Tage
testen!

Springer für Professionals.

Digitale Fachbibliothek. Themen-Scout. Knowledge-Manager.

-  Zugriff auf tausende von Fachbüchern und Fachzeitschriften
-  Selektion, Komprimierung und Verknüpfung relevanter Themen durch Fachredaktionen
-  Tools zur persönlichen Wissensorganisation und Vernetzung

www.entschieden-intelligenter.de

Springer für Professionals

 **Springer**

Frank Witte

Metriken für das Testreporting

Analyse und Reporting
für wirkungsvolles Testmanagement

Frank Witte
Landshut, Deutschland

ISBN 978-3-658-19844-2 ISBN 978-3-658-19845-9 (eBook)
<https://doi.org/10.1007/978-3-658-19845-9>

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

Springer Vieweg

© Springer Fachmedien Wiesbaden GmbH, ein Teil von Springer Nature 2018

Das Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung, die nicht ausdrücklich vom Urheberrechtsgesetz zugelassen ist, bedarf der vorherigen Zustimmung des Verlags. Das gilt insbesondere für Vervielfältigungen, Bearbeitungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen.

Die Wiedergabe von Gebrauchsnamen, Handelsnamen, Warenbezeichnungen usw. in diesem Werk berechtigt auch ohne besondere Kennzeichnung nicht zu der Annahme, dass solche Namen im Sinne der Warenzeichen- und Markenschutz-Gesetzgebung als frei zu betrachten wären und daher von jedermann benutzt werden dürften.

Der Verlag, die Autoren und die Herausgeber gehen davon aus, dass die Angaben und Informationen in diesem Werk zum Zeitpunkt der Veröffentlichung vollständig und korrekt sind. Weder der Verlag noch die Autoren oder die Herausgeber übernehmen, ausdrücklich oder implizit, Gewähr für den Inhalt des Werkes, etwaige Fehler oder Äußerungen. Der Verlag bleibt im Hinblick auf geografische Zuordnungen und Gebietsbezeichnungen in veröffentlichten Karten und Institutionsadressen neutral.

Gedruckt auf säurefreiem und chlorfrei gebleichtem Papier

Springer Vieweg ist ein Imprint der eingetragenen Gesellschaft Springer Fachmedien Wiesbaden GmbH und ist ein Teil von Springer Nature.

Die Anschrift der Gesellschaft ist: Abraham-Lincoln-Str. 46, 65189 Wiesbaden, Germany

Vorwort

In Software-Projekten ist eine der zentralen sich stellenden Fragen, wie man die eigenen Ergebnisse dem Management wirkungsvoll präsentiert. Das Projektmanagement, übergeordnete Stellen und verantwortliche Entscheidungsträger müssen erkennen, wie es um das Projekt steht und wo es Handlungsbedarf gibt. Das Testmanagement soll zeigen, ob es die vorgegebenen Ziele erreichen kann oder ob spezielle Maßnahmen getroffen werden müssen.

Leider ist heute das Reporting auf der einen Seite zu umfangreich, andererseits aber oft inhaltsleer, und leider ist das Management in vielen Fällen nicht bereit, weiter ins Detail zu gehen; man versucht sogar eher, kritischen Fragen aus dem Weg zu gehen. Das Reporting kann mit der technischen und inhaltlichen Komplexität, die von neuen Systemen und aktuellen Technologien gefordert ist, in vielen Fällen nicht mehr Schritt halten.

Es ist eine große Herausforderung, die wesentlichen Daten so aufzubereiten, dass man eine mehrdimensionale und schwierige Botschaft wirkungsvoll seinen Adressaten präsentieren kann. Das führt dazu, dass Metriken, die anhand ausgewählter Daten, Indizes und Kurven eine grafische Aufbereitung bestimmter Sachverhalte plakativ darstellen, in den vergangenen Jahren eine immer größere Bedeutung bekommen haben. Die zentrale Bedeutung des Softwaretests wird mehr und mehr als projektkritisch erkannt und führt dazu, dass man Metriken über den Projektstand vom Testmanager einfordert. Metriken sind wie ein Fieberthermometer: Sie sollen zeigen, wie gut oder wie schlimm es um das Testprojekt steht. Man kann durch die richtige Interpretation und Bewertung von Metriken weitaus besser dokumentieren, warum bestimmte Probleme im Projekt aufgetreten sind.

Messungen bestimmen unser Leben, ob im Bau, im Handwerk, in der Medizin oder in der industriellen Fertigung. Schon früh in der Geschichte der Menschheit war das Maß wichtig, um eine Orientierung zu schaffen und Erkenntnisse einordnen zu können, um Vergleichbarkeit herzustellen und um ein gemeinsames Verständnis, einen gemeinsamen Kompass für alle Betrachter zu schaffen. Messungen erlauben Vergleichbarkeit von Daten und eine Einordnung von Erkenntnissen in ein Koordinatensystem. Das Messen von Wegstrecken und Entfernungen ist eine Grundlage für die Schifffahrt, Logistik und Verkehr. In der Medizin werden zahlreiche Werte des Körpers gemessen, um den Gesundheitszustand des Patienten festzustellen. Um diese Werte sinnvoll zu messen, muss man

wissen, welche Parameter dazu herangezogen werden müssen: die Größe, das Gewicht, die Körpertemperatur, der Blutdruck, der Cholesterinspiegel oder der Körperfettanteil sind unterschiedliche Größen, die auf Probleme an der einen oder anderen Stelle hinweisen können.

Bei der Arbeit mit Metriken werden Indexwerte gebildet: Der BMI setzt Größe und Gewicht miteinander in Beziehung und erlaubt dadurch die Vergleichbarkeit von Sachverhalten, die aus zwei Größen bestehen. So ist es auch bei einem Softwareprojekt, und analog zum Body-Mass-Index gibt es auch hier Indexgrößen die unterschiedliche Kenngrößen sinnvoll in Relation zueinander setzen. Diese Indexwerte können erweitert werden, um mehrere Werte gleichzeitig miteinander in Beziehung zu setzen. Die Gewichtung mit Faktoren erlaubt eine individuelle Bewertung: Der Aktienindex sagt etwas über die gesamtwirtschaftliche Lage. Im DAX fließen die Aktienkurse der 30 größten deutschen Aktiengesellschaften ein, die wiederum unterschiedlich stark gewichtet werden: Am 07.12.16 wurde der Aktienkurs der Deutsche Lufthansa AG mit 0,58 %, der Wert der Siemens AG hingegen mit 9,44 % gewichtet. Diese Gewichtung ändert sich von Jahr zu Jahr. Indem man bestimmten Werten eine größere Bedeutung beimisst, kann man das Reporting von Ergebnissen besser an die Realität oder an die Botschaft anpassen, die man beim Reporting übermitteln will. Durch unterschiedliche Gewichtungen kann ein Ergebnis besser oder schlechter, eher beschwichtigend oder dramatischer dargestellt werden.

Projektzyklen werden tendenziell kürzer, neue Entwicklungen wegen der Fortschritte der Hardware-Entwicklung (Prozessoren, Speicher) und der dynamischen Prozesse (z. B. in agilen Projekten) immer schneller möglich. Diese Schnelligkeit erzeugt einen hohen Erwartungsdruck auf neue Softwareanwendungen und Erweiterungen bestehender Features. Dieser Termindruck wirkt sich auch auf Testaktivitäten aus. Projekte müssen zu bestimmten Terminen fertig werden, Meilensteine und Budgets sind (leider) dem Quartalsdenken und teilweise unrealistischen Renditeerwartungen unterworfen. Die Planung ist häufig nicht ausreichend filigran und berücksichtigt bestimmte Parameter und Rahmenbedingungen nicht hinreichend. So geht vieles an Qualität und Tiefe verloren, die Kostenvorgaben und Zeitschätzungen sind zu ambitioniert und stehen nachhaltigen, qualitativ hochwertigen Produkten entgegen. Die Globalisierung hat zwar in vielen Fällen das Tempo von Entwicklungszyklen erhöht, das geht aber meist auf Kosten der Produktreife („Bananensoftware reift beim Kunden“). Diese Problematik erkennt man häufig in Entwicklungsprojekten, hat aber schließlich nicht genügend handfeste und transparente Argumente, um entsprechend gegenüber dem Management oder dem Controlling zu argumentieren.

Ich bemerke immer wieder, dass der Projektaufwand unterschätzt wird, weil zu wenig auf Probleme in den Prozessen und der Organisation geachtet wird. Technisch sind die Probleme meist lösbar, aber die Technik macht eben tendenziell nur 20 % der gesamten Probleme aus, 80 % liegen in der Kommunikation, der Arbeitsorganisation und den Rahmenbedingungen verborgen. Diese Tatsache wird gerade im technologielastigen Umfeld nicht ausreichend gewürdigt. Dazu kommt immer wieder ein Phänomen, das menschlichen Schätzungen generell innewohnt: Wir unterschätzen die Ziele, die wir in 5 oder 10

Jahren erreichen können, aber überschätzen, was in 1 oder 2 Jahren erzielt werden kann. Intervalle von 1 oder 2 Jahren Dauer sind aber gerade die typische zeitliche Länge für Projekte in der Softwareerstellung.

Aussagen, mit denen man auf mehr Softwarequalität und Risikominimierung hinwirkt, müssen jedoch mit handfesten Daten untermauert werden, um auf Abhängigkeiten und Auswirkungen hinzuweisen. Dazu verhelfen Vergleichswerte und fundierte Zahlen. Je mehr unterschiedliche Indizes erhoben werden, desto genauer wird das Bild, was man über ein Projekt, ein Produkt oder eine Software erhält. Zahlen verhelfen dazu, von einem ungefähren, emotionalen Eindruck zu einer definierten, verlässlichen und belastbaren Messung zu gelangen. Nur mit Zahlen schafft man eine sichere Argumentationsbasis und erhält nachvollziehbare Aussagen. Viele Probleme sind von verschiedenen Einflussgrößen abhängig, die sich gegenseitig beeinflussen. Zahlen, Metriken und Analysen schaffen eine bessere Grundlage, die richtigen Entscheidungen zu treffen und frühzeitig die entscheidenden Weichenstellungen vorzunehmen.

Ein Reporting über 20 bis 50 Seiten, wie es bereits für ein mittleres Softwareprojekt erforderlich ist, erschlägt den Empfänger durch seinen Informationsumfang. In einem Project Review vor einem Project Director oder Manager, das in der Regel einmal pro Woche stattfinden sollte, hat man vielleicht eine Stunde Zeit, um den aktuellen Projektstatus darzustellen: Da sollte man mit wenigen Charts auskommen, die man mit wenigen Worten erklären kann. Es wird leider vermehrt dazu übergegangen, aus „Zeitgründen“ die Metriken nur online bereitzustellen. Auch das ist ein Trend, bei dem aus vermeintlicher Zeiteinsparung eher wesentliche Informationen verloren gehen. Bei einzelnen Aktivitäten, zum Beispiel im Reporting, bei Review-Maßnahmen oder bei der Erstellung der System Requirements, sieht man die Erfolge nicht sofort. Deswegen werden einige notwendige Aktivitäten aus falsch verstandener Prioritätensetzung weggelassen oder eingespart, was aber das Gesamtprojekt am Ende viel teurer und langwieriger werden lässt. Weiterhin muss man bedenken, dass das Reporting von Projekten nach oben hin immer mehr verdichtet wird. Wenn es in einem Konzern 100 unterschiedliche größere Entwicklungsprojekte gibt, dann will das Topmanagement am Ende nur über diejenigen im Detail informiert werden, bei denen Penalty-Zahlungen oder Abbrüche drohen, also die Projekte bei denen „die Hütte brennt“.

In jedem Fall besteht eine besondere Herausforderung darin, die wesentlichen Aussagen mit Hilfe von Metriken hervorzuheben und die Konzentration auf das Wesentliche zu lenken. Dazu verhelfen „Projektampeln“ und „Fieberkurven“, weil diese Darstellungen schon aus dem täglichen Leben bekannt sind. Viele Projekte verlaufen zwar sicherlich nicht gerade optimal, aber eben gerade noch so, dass man es mit ein paar Wochen Verzug und wenigen Prozent Kostenüberlauf und einer etwas geringeren Qualität als beabsichtigt oder einer abgespeckten Funktionalität immer noch vermeiden kann, im Fokus des Topmanagements zu stehen. Wenn man nämlich erst einmal zu den wenigen Katastrophenprojekten gehört, die im Rampenlicht stehen, gewinnt man dauernde Reportings und zusätzliche Panik und wird allein durch die ständige Aktualisierung der Fehlerlisten noch weniger fertig als ohnehin schon. In diesem Fall erlebt man häufig Aktionismus: Viele

zusätzliche Ressourcen werden in kurzer Zeit ins Projekt gepumpt, Maßnahmen wie Samstagsarbeit und angeordnete Überstunden getroffen, die aber in vielen Fällen nur akut das Schlimmste verhindern und letztlich Kosmetik sind, weil die eigentlichen Probleme in der Regel viel tiefer liegen und kurzfristig gar nicht behoben werden können. Wenn man aber nicht die tiefen Ursachen angeht, erschöpft man sich in der Besänftigung von Symptomen. Je eher man daher Projekte auf eine gesunde Grundlage stellt, desto besser sind die Projektziele zu erreichen.

Bei der Präsentation, Analyse und Bewertung der Metriken gefällt es dem Management allgemein, wenn linear oder progressiv ansteigende Kurven zu sehen sind. In der Praxis gibt es aber in der Regel viel zu viele Einflussfaktoren, um diese optimalen Kurvenverläufe permanent präsentieren zu können. Um die richtigen Metriken zu nutzen, die einzelnen Fakten sinnvoll zu interpretieren und den Fortschritt korrekt zu bewerten, müssen geeignete Parameter in einer anschaulichen Darstellung angegeben werden.

In Metriken fließen unterschiedliche Werte ein. Wichtig ist zunächst einmal, möglichst viele Daten konsequent zu erheben. Metriken können auch unterschiedliche Projekte miteinander vergleichen. Es bringt wenig, die Schuld bei den Projektmitgliedern zu suchen, wenn ein Projekt nicht richtig läuft. Eine Metrik bringt den wirklichen Status klar und neutral hervor. Daher erkennt man auch, dass vielleicht bestimmte Voraussetzungen von Anfang an gar nicht gegeben waren und das Projekt gar keine höhergesteckten Ziele erreichen konnte.

Damit überhaupt Metriken unterschiedlicher Projekte miteinander verglichen werden können, ist es wichtig, innerhalb des Unternehmens gemeinsame Festlegungen zu treffen, sonst vergleicht man Äpfel mit Birnen. Das einzelne Projekt sollte sich daher aus einem unternehmensweit von einer Prozessgruppe definierten Bestand von Metriken diejenigen auswählen, die für das entsprechende Projekt sinnvoll und notwendig sind. Für diese Metriken werden die erforderlichen Daten für die Laufzeit des Projekts regelmäßig erhoben. Gerade diesen Aspekt sehe ich bisher in der betrieblichen Praxis in den wenigsten Fällen erfüllt, meist wird das Rad in jedem Projekt neu erfunden. Der geforderte unternehmensweit definierte Bestand von Metriken existiert nämlich in den seltensten Fällen bzw. ist so gut versteckt, dass der Projektmanager und der Testmanager ihn nicht finden und sich dann doch wieder eigene Metriken überlegen. Um in solchen Situationen nicht ganz bei null anfangen zu müssen, sind in diesem Buch einige geeignete Metriken näher beschrieben.

In den Projekten werden viele Daten nicht erhoben bzw. nicht archiviert. Und falls in einem Projekt dennoch brauchbare Ergebnisse erhoben wurden, fließen diese Daten selten in die Planung und Kalkulation von Folgeprojekten ein, die gewonnene Datenbasis wird nicht weitergenutzt. Dabei wird wertvolles Potenzial verschenkt. Meist wird beim Folgeprojekt ähnlich optimistisch kalkuliert wie beim vergangenen Projekt, das man schon kaum gestemmt hat und das nur mit Terminverzug und Kostenüberschreitung auf den letzten Drücker bei Abstrichen in der Qualität fertig wurde.

Die „Lessons Learned Meetings“ beschränken sich oft nur auf 30 Minuten oberflächliche Betrachtungen, an den Meetings nehmen nur wenige Stakeholder (oder diejenigen, die

mit der operativen Umsetzung am wenigsten betraut waren) teil, oder eine skeptische Rückschau fällt ganz aus. Wenn überhaupt notwendige Aktivitäten für die Zukunft protokolliert werden, sind sie beim nächsten Projekt längst vergessen.

Besonders zu beachten ist, dass immer nach dem Ziel der einzelnen Metrik gefragt werden sollte und ein gemeinsames Verständnis vorhanden sein sollte, was eine bestimmte Metrik aussagen kann, welche Folgerungen aus ihr gezogen werden können und welche nicht und wo die Grenze der einzelnen Metrik ist. Einzelne Metriken beantworten ganz bestimmte Fragen, und wenn einfach nur Hunderte von Zahlen erhoben werden, also ein Messen ohne Ziel am Ende steht, wenn der Nutzen den Entwicklern und Testern nicht transparent ist, entsteht eine Abwehrhaltung und es wird nur zusätzlicher Aufwand gesehen. Metriken können immer interpretiert werden und geben unter diesen Bedingungen ein Zerrbild der Realität ab.

Die Arbeit in der Produktion wird schon seit Jahrzehnten sehr genau getrackt und mit Hilfe von Metriken untersucht. Inzwischen wird Büroarbeit wie Fließbandarbeit in zahlreiche Teilschritte zerlegt. Das Tayloring der Prozesse, die verursachergerechte Zuordnung von Aktivitäten und die zunehmende Überwachung und Kontrolle von einzelnen Arbeitsschritten führt zu einem erhöhten Bedarf an Reporting und das betrifft auch Testaktivitäten. Demgegenüber entsteht eine Abwehrhaltung und es wird nur zusätzlicher Aufwand gesehen. Metriken sind dann sinnvoll, wenn alle Beteiligten im Projekt ihren Sinn verstehen und aus den Aussagen zielgerichtet Optimierungen eingeleitet und umgesetzt werden.

Im Zeitverlauf von Projekten kann es vorkommen, dass einzelne Fragestellungen sich ändern und unterschiedliche Aussagen im Fokus stehen. Die einzelnen Kapitel dieses Buches beziehen sich auf bestimmte Fragestellungen, die im Zusammenhang mit Testaktivitäten stehen. Die häufigsten und verbreitetsten Metriken werden dabei genannt. Es kann aber im individuellen Projekt durchaus erforderlich sein, Fragestellungen zu vertiefen oder Erweiterungen von Metriken vorzunehmen, spezielle Indexwerte zu erheben und Metriken zu erweitern. Daher ist es aus meiner Sicht erforderlich, dem Testmanagement Unterstützung in Analyse und Interpretation erhobener Daten zukommen zu lassen und sehr genau hinzusehen, wo die Probleme im Projekt sind und wie sie in eine sinnvolle Darstellung gebracht werden können, um den Ist-Zustand umfassend und dennoch prägnant darzustellen und die richtigen Fragen zu stellen, damit die richtigen Maßnahmen eingeleitet und die richtigen Entscheidungen getroffen werden. Richtig eingesetzt, können Metriken sehr effizient zum Projektcontrolling eingesetzt werden, falsch eingesetzt fristen sie schnell ein Schattendasein oder werden als überflüssiger Aufwand abgeschafft. Es ist daher von besonderer Bedeutung, dass das Management Metriken fordert und den Einsatz aktiv beeinflusst.

Metriken sollen Eigenschaften von Funktionen auf Zahlenwerte abbilden. Sie sollen möglichst objektiv Informationen verdichten, Komplexität reduzieren, adressatengerecht aufbereitet sein und schnell und frühzeitig Probleme erkennen lassen und diese Erkenntnisse transparent darstellen. Im Auto werden auf dem Armaturenbrett wesentliche Informationen (Geschwindigkeit, Drehzahl, Uhrzeit, Warnhinweise) dem Fahrer dargestellt.

In einem Flugzeug sind schon erheblich mehr Informationen für den Piloten erforderlich, aber auch hier geht es um die sinnvolle Verdichtung von Daten. Ähnlich muss es einer Testmetrik gelingen, wesentliche Daten auf einen Blick zu zeigen. Die Verdichtung von Informationen ist immer ein Drahtseilakt: Tendenziell gehen dabei zu viele Nuancen verloren, der Wunsch nach dem „One Pager“ ist einerseits nachvollziehbar, andererseits benötigen komplexe Prozesse und schwierige Sachverhalte eben auch detaillierte Hintergründe. Oft nimmt man sich viel zu wenig Zeit, die Ursachen zu ergründen. Die Informationsfülle auf der einen Seite und der Wunsch nach einfachen, klaren Aussagen andererseits gehören zu den wesentlichen Ursachen für Kommunikationsprobleme in Projekten. Es ist wichtig, stringente Daten zu liefern, und manchmal muss man dazu Aussagen vereinfachen und reduzieren, weil sich Nebenbedingungen und Besonderheiten nur in einer Fußnote sinnvoll erklären lassen, aber die Metrik überfrachten und die Aussagen am Ende unverständlich machen würden. Nicht immer ist es sinnvoll, jedes Detail ausführlich zu beleuchten und umfassend zu erklären.

Eine wichtige Ursache dieser Informationsfülle ist das Mooresche Gesetz: Das „Moore'sche Gesetz“ besagt, dass sich die Komplexität integrierter Schaltkreise mit minimalen Komponentenkosten regelmäßig verdoppelt; je nach Quelle werden 12 bis 24 Monate als Zeitraum genannt. Unter Komplexität verstand Gordon Moore, der das Gesetz 1965 formulierte, die Anzahl der Schaltkreiskomponenten auf einem integrierten Schaltkreis. Gelegentlich ist auch von einer Verdoppelung der Integrationsdichte die Rede, also der Anzahl an Transistoren pro Flächeneinheit. Dieser Technologiefortschritt bildet eine wesentliche Grundlage der „digitalen Revolution“. Gordon Moore äußerte seine Beobachtung in einem 1965 erschienenen Artikel, nur wenige Jahre nach /Erfindung der integrierten Schaltung. Die Bezeichnung „Moore'sches Gesetz“ wurde um 1970 von Carver Mead geprägt. Ursprünglich sagte Moore eine jährliche Verdoppelung voraus, korrigierte diese Aussage jedoch 1975 auf eine Verdoppelung alle zwei Jahre. Die stürmische Entwicklung der Halbleitertechnik der ersten Jahre hatte sich etwas verlangsamt. Neben der Verkleinerung von Elementen und der Vergrößerung der Wafer spielte in den ersten Jahren noch die von Moore so genannte „cleverness“ eine Rolle, nämlich die Kunst, Bauelemente intelligent auf dem Chip zu integrieren. Die Grenzen dieser Cleverness waren in den 1970er-Jahren weitgehend ausgereizt. Moores damaliger Intel-Kollege David House brachte eine Abschätzung von 18 Monaten für die Verdoppelung der Rechenleistung von Chips ins Spiel, was heute die verbreitetste Variante des Mooreschen Gesetzes darstellt und den Rahmen bildet, an dem die Halbleiterindustrie ihre Entwicklungspläne auf mehrere Jahre hinaus festmacht. Real verdoppelt sich die Leistung neuer Computerchips im Mittel etwa alle 20 Monate, mit gewissen Schwankungen.

Moores Gesetz ist kein wissenschaftliches Naturgesetz, sondern eine Faustregel, die auf eine empirische Beobachtung zurückgeht. Gleichzeitig kann man von einer „sich selbsterfüllenden Prophezeiung“ sprechen, da verschiedene Industriezweige an der Entwicklung besserer Mikrochips beteiligt sind. Sie müssen sich auf gemeinsame Meilensteine einigen (z. B. optische Industrie mit verbesserten lithografischen Methoden), um wirtschaftlich arbeiten zu können. Die Formulierung des Mooreschen Gesetzes hat sich im

Laufe der Zeit stark verändert. Sprach Moore noch von der Komponentenanzahl auf einem integrierten Schaltkreis, so ist heute von der Transistoranzahl auf einem integrierten Schaltkreis die Rede, mitunter sogar von der Transistoranzahl pro Flächeneinheit. Zwar sieht man hin und wieder das Mooresche Gesetz an seine Grenzen kommen, die 18 Monate sind ein Mittelwert über mehrere Jahre, aber generell geht der Fortschritt in der Prozess- und Speichertechnologie immer noch weiter: Die aktuellen Prozessoren der großen Chiphersteller werden mit der 10-Nanometer-Technologie gefertigt. 14 Nanometer sind zurzeit die kleinsten Strukturen bei Mikroprozessoren. Gemäß der „International Technology Roadmap for Semiconductors“ sollen die Strukturen (Transistoren und Leiterbahnen) bis zum Jahr 2021 auf 5 Nanometer schrumpfen. Diese Größenordnungen übersteigen das menschliche Vorstellungsvermögen bei weitem. Ein Nanometer ist ein Milliardstel Meter, das bedeutet eine Zahl mit acht Nullen und einer eins hinter dem Komma. Auch der Vergleich zum menschlichen Haar bietet kaum mehr Hilfe. Ein durchschnittliches Haar ist weit mehr als 4000 Mal so dick.

Informationskompetenz stellt in der modernen, stark dynamischen Informationsgesellschaft eine Schlüsselqualifikation zur Bewältigung von Problemen dar. Sie gehört zum Bereich der sozialen Kompetenz (Soft Skills) und umfasst im Allgemeinen eine Reihe von Fähigkeiten, die dem Einzelnen den kompetenten, effizienten – unter Berücksichtigung von Rahmenbedingungen wie Zeit oder Programme – und verantwortungsbewussten Umgang mit Informationen ermöglicht. Diese Fähigkeiten beziehen sich auf alle Aspekte des problembezogenen Erkennens eines Bedarfs an Informationen, ihrer Lokalisation, ihrer Organisation, ihrer zielgerichteten Selektion durch Analyse und Evaluation und ihrer zweckoptimierten Gestaltung und Präsentation. Aus den Anforderungen an Informationskompetenz erwachsen Erfordernisse für Metriken.

Metriken sollen wenig Interpretationsaufwand abbilden. Metriken sollen stabil sein, unwichtige Änderungen sollen wenig Effekt auf das Ergebnis haben. Die einzelnen Werte müssen vergleichbar und analysierbar sein, und die Maße müssen rechtzeitig abgebildet werden, so dass man noch eingreifen kann. Das Kriterium der Reproduzierbarkeit besagt, dass das Ergebnis der Messung nicht vom Messenden abhängt und nicht manipulierbar sei.

Dieses Buch soll dabei helfen, einen Überblick über die verschiedenen Parameter zu gewinnen, den Einsatzzweck zu beschreiben und dadurch für das eigene, individuelle Projekt die besten Bewertungsparameter zu ermitteln, denn diese können nach Art der Fragestellung, Projektumfang oder Projektfortschritt durchaus unterschiedlich sein. In meinem Buch „Softwaretest und Testmanagement“ habe ich das gesamte Thema Softwaretest umfassend betrachtet. Es liegt in der Natur der Sache, dass ich manche Aspekte nur anreißen und nicht in seiner gesamten Bandbreite in der Tiefe behandeln konnte. Mit dem vorliegenden Buch möchte ich auf einen wesentlichen Teilaspekt des Testmanagements besonders eingehen, da seine Wichtigkeit enorm gestiegen ist.

Inhaltsverzeichnis

1	Definition, Historie und Nutzen von Metriken	1
2	Klassifizierung von Metriken	7
3	Komplexitätsmaße	17
4	Normen und Standards für Messungen	35
5	Verwendung unterschiedlicher Darstellungen	47
6	Einsatz von Metriken	55
7	Softwaremetriken	61
8	Metriken für die Softwarequalität	63
9	Anforderungsbasierte Testmetriken	71
10	Metriken zur Aufwandsschätzung	75
11	Metriken zur Testvorbereitung und -nachbereitung	81
12	Metriken zur Testplanung	83
13	Metriken für unterschiedliche Teststufen	87
14	Metriken zur Messung der Testüberdeckung	91
15	Testfallbasierte Metriken	107
16	Testobjektbasierte Metriken	115
17	Metriken zur Testdurchführung	125
18	Metriken für Performance- und Lasttests	129
19	Metriken zur Testautomatisierung	133
20	Kostenbasierte Testmetriken	137
21	Testproduktivitätsmetriken	141

22	Goal Question Metric	145
23	Fehlerbasierte Testmetriken	151
24	Metriken für Testdokumente	169
25	Metriken für Usability-Tests	173
26	TPI Next	179
27	Metriken für Testendekriterien	193
28	Metriken und Testreporting	197
29	Vergleich von Testmetriken	207
30	Problemlösung für den Einsatz von Testmetriken im Testbetrieb	213
31	Wirkung beim Reporting mit Hilfe von Testmetriken	221
	Nachwort	233
	Weiterführende Literatur	239
	Stichwortverzeichnis	241



Definition, Historie und Nutzen von Metriken

1

Zusammenfassung

Metriken werden dafür eingesetzt, um das Reporting des Projektfortschritts zu erleichtern. Sie werden zur Kontrolle und Prognose in Entwicklungsprojekten eingesetzt und erleichtern das Verständnis für komplexe Zusammenhänge.

Das Maß gehört zu den frühesten Werkzeugen, das die Menschheit erfunden hat. Das alltägliche Leben von Menschen ist unvorstellbar ohne solche Einheiten, die ein gemeinsames System für Verständnis und eine Grundlage für alle Wissenschaften bilden. Beim Anfang des Handels, der Produktion von Alltagswaren oder bei der Entstehung von Bautätigkeiten entwickelten sich die ersten Maße. Je mehr in der Naturwissenschaft erforscht wurde, desto mehr entstand die Notwendigkeit für übersichtliche und nützliche Darstellung von neuer Information. Dabei musste die Information nicht nur von anderen korrekt interpretiert werden können, sondern auch vergleichbar sein. Die Anforderung an die Vergleichbarkeit hat sich als sehr wichtig in der Entwicklung der Wissenschaft herausgestellt. Diese Anforderung führte im Verlaufe des 18. Jahrhunderts in Europa zur Entstehung des auf dem Meter basierenden Einheitssystems. Das höhere Ziel für die Gültigkeit des metrischen Systems lautet: „Für alle Welt, für alle Völker.“ Jedes Land, das den internationalen Handel sowie wissenschaftlichen und technischen Austausch unterstützen wollte, musste das System akzeptieren [Alyo2017].

Das Maß ist eine der wesentlichen Voraussetzungen für Standardisierung und Globalisierung. Metriken stellen eine Vergleichbarkeit her.

Eine Metrik (Zählung, Messung) bezeichnet im Allgemeinen ein Verfahren zur Messung einer quantifizierbaren Größe. Die „quantifizierbare Größe“ im Softwaretest kann z. B. die Testbarkeit der Software oder die Komplexität oder Qualität der Testfälle darstellen. Metriken wandeln den Softwaretest zu einer messbaren Technologie.

Ziele der Erstellung von Metriken sind dabei u. a.:

- Bewertung der Software
- Aufdecken von Fehlern und fehlerhaften Prozessen
- Schaffen einer Vertrauensbasis in die Software
- Gewichtung und Rechtfertigung der Testkosten im Vergleich zum gesamten Softwareherstellungsprozess

Metriken ermöglichen Objektivität durch die Quantifizierung von Beobachtungen [Soge2011]. Eine Metrik ist eine Messung für quantitative Bewertungen oder Verbesserungen eines Prozesses oder Produkts. **Key Performance Indicators** (KPI) bewerten quantifizierbare Metriken, welche die Performance der Organisation bezüglich eines Ziels bewerten

Metriken werden dafür eingesetzt, um den Testprozess zu steuern. Sie untermauern die Testberatung und finden beim Vergleich von Systemen und Prozessen Anwendung. Eine Testmetrik ist die messbare Eigenschaft eines Testfalls oder Testlaufs mit Angabe der zugehörigen Messvorschrift, wie zum Beispiel dem Testabdeckungsgrad.

Metriken sollen gewisse Voraussetzungen erfüllen, damit deren Erhebung regelmäßig fortgeschrieben und nicht vernachlässigt wird:

1.1 Kriterien für Metriken

Folgende Kriterien soll eine Metrik erfüllen:

- Einfachheit der Erhebung
- kompakte und objektive Beschreibung
- Vorhersagen zu geordneter Planung und Prognose von Verbesserungen
- Repräsentanz (die Darstellung als Zahl kann sinnvoll umgesetzt werden)
- Eindeutigkeit (viele Abbildungen sind möglich)
- Nachvollziehbarkeit
- Skalierbarkeit

1.2 Gütekriterien für Metriken

Folgende Gütekriterien sollen Metriken erfüllen:

- Objektivität: kein subjektiver Einfluss durch Messenden möglich
- Validität und Konsistenz: die Metrik misst wirklich das, was sie vorgibt zu messen
- Zuverlässigkeit (Reproduzierbarkeit): die Wiederholung liefert gleiche Ergebnisse
- Nützlichkeit: die Metrik hat eine praktische Bedeutung
- Normiertheit: eine Skala für die Messergebnisse existiert

- Vergleichbarkeit: die Metrik ist mit anderen Maßen vergleichbar
- Ökonomie: die Metrik ist mit vertretbaren Kosten messbar [Kosc2011]

1.3 Erfolgskriterien für Metriken

Es gibt einige wichtige Erfolgsfaktoren für die Arbeit mit Metriken:

- Metriken müssen proaktiv sein und in der täglichen Arbeit eingesetzt werden können. Metriken, die am Monatsende vergangenheitsbezogen erfasst werden müssen und deren Bezug zur täglichen Arbeit sich nicht erschließt oder Fehlerzahlen ohne Analyse der Fehlerursachen werden ungern erhoben und erscheinen nur als Befriedigung einer Vorgabe, werden aber von den betroffenen Mitarbeitern nicht ernst genommen.
- Metriken müssen einfach und in Ziel und Auswirkungen nachvollziehbar sein. Komplexe und aufwändige Rechenmodelle eignen sich nicht für die Einführung von Metriken, sondern höchstens im weiteren Verlauf und von den Modellen abgeleiteter Fragestellungen.
- Metriken müssen im Projekt erhoben und gesammelt werden und nicht durch externe Planer oder eigene „Metrikgruppen“.
- Der Messgegenstand muss klar definiert sein; es muss eindeutig darstellbar sein, welcher Sachverhalt untersucht wird. Gerade an diesem Punkt, der sich zunächst relativ trivial anhört, scheitert oft das Reporting mit Metriken. Dazu gehört, dass die in der Organisationseinheit verwendeten Begriffe klar definiert sind, am besten in einem Glossar, und dass alle Stakeholder von einem bestimmten Terminus dieselben Vorstellungen haben. Eine erfolgreiche Kommunikation ist dazu das wesentliche Element und gerade die richtige Kommunikation wird häufig vernachlässigt.
- Programme für Metriken verlangen Training und exakte Anleitungen, auf welche Weise sie ermittelt werden sollen. Ein hoher manueller Aufwand bei der Erhebung der Metriken ist zu vermeiden, die Ermittlung, Erfassung und Verdichtung der Daten soll so automatisiert wie möglich erfolgen.
- Die Arbeit mit Metriken verlangt eine detaillierte Analyse und individuelle Interpretation der Daten ohne personenbezogene Schuldzuweisungen.
- Metriken müssen vom Management gefordert und aktiv unterstützt werden. Erfolgreiche Projekte und neue Erkenntnisse aufgrund von Metriken führen zu stärkerer Akzeptanz ihres Einsatzes [Eber1996].

1.4 Validierung von Maßen

Man unterscheidet bei der **Validierung** von Maßen zwischen interner und externer Validierung. Die interne Validierung ist der Nachweis, dass ein Maß eine gültige numerische Charakterisierung eines Attributs ist, durch

- Nachweis der Erfüllung der Repräsentanzbedingung und
- Prüfung des Skalentyps.

Eine externe Validierung dient als Vorhersagemodell:

- Hypothese über den Zusammenhang zwischen zwei Maßen
- Erfassung der Messwerte beider Maße auf gleicher Testmenge
- statistische Analyse von Ergebnissen durch Bestimmung von Parametern und Prüfung der Allgemeingültigkeit
- Untersuchung von Prozessen, Ressourcen und Produkten
- isoliert (intern) oder extern (mit Umgebung)
- zu unterschiedlichen Phasen des Prozesses
- objektiv oder subjektiv, direkt oder abgeleitet

Bei Metriken unterscheidet man zwischen internen und externen Metriken. Eine interne Metrik ist darüber definiert, dass sie nur Eigenschaften innerhalb des untersuchten Objekts misst, wohingegen externe Metriken Interaktionen und Wechselwirkungen des Objektes mit seiner Umgebung berücksichtigen.

Metriken werden sowohl für den Rückblick als auch für Prognosen eingesetzt. Testmetriken dienen dem Testmanagement dazu, im Softwareprojekt den Aspekt einer nachhaltigen Qualitätssicherung zu dokumentieren.

Eine Metrik liefert dabei immer nur Aussagen bezüglich eines bestimmten untersuchten Aspekts. Die ermittelten Maßzahlen sind immer erst im Vergleich zu den Zahlen aus anderen untersuchten Programmteilen oder Werten zu anderen Kriterien aussagekräftig.

1.5 Probleme im Softwareentwicklungsprozess

In Zusammenhang mit der Erstellung neuer Anwendungen treten typischerweise Probleme in folgenden Bereichen auf. Metriken sollen dazu geeignet sein, diese Problembereiche zu untersuchen und mit Hilfe von Indexwerten auszudrücken:

- Überschreitung der geplanten Projektdauer
- unvollständige Realisierung der gewünschten Qualität
- unvollständige Realisierung von Funktionen
- Überschreitung der geplanten Projektkosten

1.6 Historie der Testmetriken

Aktivitäten im Softwaretest entwickelten sich um 1970, und schon bald wurden Testüberdeckungen gemessen. Da die Programmzuverlässigkeit damals zunehmend wichtig wurde, war es auch notwendig, gefundene Fehler zu messen. Bill Hetzel beschrieb 1993

mögliche Ansätze zur Testmessung und Stephan Kan behandelte in einer Studie Ende der 1990er-Jahre ebenfalls unterschiedliche Messmodelle. Im Jahre 2006 wurde von Sneed und Jungmayr der erste deutschsprachige Beitrag zum Thema Testmetrik publiziert [Snee2010]. Inzwischen ist der Einsatz von Testmetriken in Software-Projekten allgemein verbreitet.

1.7 Schritte zum Aufsetzen von Testmetriken

Es sollte möglichst zu Anfang des Projekts definiert werden, welche Metriken sinnvoll sind und im Projektverlauf erhoben werden sollen. Leider wird zu Projektbeginn (wie für viele andere Aktivitäten) in der Praxis in den weit überwiegenden Fällen viel zu wenig Zeit in diese notwendige Vorarbeit investiert und notwendige Definitionen werden unterschätzt. Gerade Dinge, die auf den ersten Blick trivial erscheinen, werden oft zu wenig beachtet und führen zu dauerhaften Problemen im Projekt.

Die Metriken sollen bereits im Testkonzept definiert werden (Tab. 1.1). Eine nachträgliche Einführung von Testmetriken ist in der Regel mit Mehraufwand verbunden, vor allem dann, wenn man Werte aus der Historie rekonstruieren muss. Wenn man Metriken während des Projektablaufs erst definiert und ab einem bestimmten Punkt einsetzt, hat man nicht die Entwicklung im Projektverlauf vollständig aufgezeigt. Es ist aber immer noch besser, eine Metrik nachträglich einzuführen als Werte gar nicht zu messen.

Tab. 1.1 Schritte zur Definition von Testmetriken

Lfd. Nr.-	Schritte zur Testmetrik	Beispiel
1	Schlüsselprozesse und Abläufe, die zu messen sind, im Test bestimmen	Trackingprozess des Testfortschritts
2	Daten verwenden zur Definition der Grundlagen der Metrik	Anzahl Testfälle, die pro Tag zur Testdurchführung geplant sind
3	Bestimmung der zu verfolgenden Daten, Häufigkeit der Erhebung der und für die Datenerhebung verantwortlichen Mitarbeiter	Die Anzahl durchgeführter Testfälle wird am Ende des Tages durch den Testmanager erhoben
4	Kalkulation, Management und Interpretation der definierten Metriken	Anzahl Testfälle, die pro Tag durchgeführt wurden
5	Identifikation der Verbesserungspotenziale aufgrund der Interpretation der Ergebnisse	Die Anzahl der durchgeführten Testfälle ist niedriger als erwartet, die Ursachen müssen analysiert und Gegenmaßnahmen vorgeschlagen werden

1.8 Lifecycle von Testmetriken

Im Allgemeinen haben alle Testmetriken folgenden Ablauf:

1. Analyse:
Identifikation und Definition der zu verwendenden Testmetriken
2. Kommunikation:
 - Erläuterung der Notwendigkeiten gegenüber den Stakeholdern und dem Testteam
 - Erläuterung der Verfahren zur Datenerhebung
3. Evaluierung:
 - Sammeln und Verifizieren der Daten
 - Kalkulation der Indexwerte
4. Report:
 - Erstellen des Testberichts mit Hilfe der Metriken
 - Unterrichtung von Projektmanagement und Einsammeln des Feedbacks

Literatur

- [Eber1996] Ebert C, Dumke R (1996) Software-Metriken in der Praxis. Springer, Heidelberg
- [Soge2011] TPI Next (2011) Geschäftsbasierte Verbesserung der Testprozesse. dpunkt-Verlag, Sogeti
- [Snee2010] Sneed, Baumgartner, Seidl (2010) Software in Zahlen. Hanser-Verlag, München
- [Kosc2011] Koschke. Softwaretechnik, Fachbereich Mathematik und Informatik, Universität Bremen, WS 2011/12. <https://www.informatik.uni-bremen.de/st/lehre/swt11/metriken.pdf>. Zugegriffen am 05.10.2017
- [Botr2006] Projektübergreifendes Testmanagement am Beispiel des Produkts „Quality Center“ in einem Finanzunternehmen, Bachelorarbeit Sinan Botros im Studiengang Informatik. Hannover (2006) <http://www.se.uni-hannover.de/pub/File/pdfpapers/Botros2006.pdf>. Zugegriffen am 07.10.2017
- [Alyo2017] Alyokhin V. Management von Softwaresystemen, Systembewertung: Metriken und Prozesse. http://www4.in.tum.de/lehre/seminare/hs/WS0506/mvs/files/Ausarbeitung_Alyokhin.pdf. Zugegriffen am 12.10.2017



Zusammenfassung

Metriken lassen sich nach mehreren Klassifizierungen in unterschiedliche Kategorien unterteilen.

2.1 Kennzahlen

Als Kennzahlen bezeichnet man Zahlen, welche messbare, betriebswirtschaftlich relevante Daten zusammenfassen. Sie bündeln damit mehrere Daten zu einer aussagekräftigen Größe. Mit Hilfe von Kennzahlen können Unternehmen oder Organisationen bewertet und Ergebnisse, zum Beispiel aus den Vorjahren, verglichen werden. Gleichzeitig können Unternehmen und Verwaltungen sich mit diesen Kennzahlen auch an anderen Organisationen messen. Kennzahlen basieren zumeist auf aggregierten und kompakten Informationen. Grob lassen sich Kennzahlen gliedern in

- absolute Kennzahlen: z. B. Fahrzeit, Gesamtkosten, Personalkapazität, gemittelter Wochenumsatz
- relative Kennzahlen (Verhältniskennzahlen, Indexwerte), dabei
 - dimensionsbehaftete relative Kennzahlen: z. B. Stückkosten, Spesen pro Tag, Umsatz pro Kunde
 - dimensionslose relative Kennzahlen: z. B. Preisindex, Aktienindex, Fertigstellungsgrad, prozentualer Anteil, Umsatzrendite
- Bestandskennzahlen: z. B. Krankenstand, Marktpreis, Temperatur (Gültigkeit zu einem festgelegten Zeitpunkt)
- Verlaufskennzahlen: z. B. Trends und Durchschnittswerte (Gültigkeit für einen festgelegten Zeitraum)

Kennzahlen fassen messbare, relevante Daten zusammen und stellen sie in einen größeren Zusammenhang. Kennzahlen sind quantitative Informationen, die für die Bedürfnisse der Analyse und Steuerung aufbereitet werden. Kennzahlen leisten

- die Messung betrieblicher Vorgänge,
- die Beurteilung unternehmensrelevanter Sachverhalte,
- die kurze und prägnante Darstellung komplexer Tatsachen,
- die Festlegung zukünftiger Maßstäbe sowie
- die Festlegung kritischer Erfolgsfaktoren.

Ein Kennzahlenwert ist der Wert der Kennzahl zu einem bestimmten Zeitpunkt (zum Beispiel Mitarbeiterzahl am 31. Dezember 2016) oder über einen festgelegten Zeitraum (zum Beispiel Gewinn in einem Geschäftsjahr). Für Kennzahlen finden sich typischerweise Namen wie -anteil, -faktor, -grad, -index, -koeffizient, -quote, -verhältnis, -zahl und ähnliche Zusätze, die teilweise nach den messtechnischen Normen speziellen Typen von Kennzahlen vorbehalten sind. Streng genommen ist eine Kennzahl für sich genommen nicht aussagekräftig; gemeint ist immer eine Kenngröße, also das Produkt aus Zeichen (z. B. Buchstaben, auch in Kombination mit mathematischen Zeichen oder Sonderzeichen bei Ratings [z. B. AA+] oder eine Zahl bei numerischen Angaben und ihre Einheit (z. B. Celsius, Euro, Meter).

In der mathematischen Statistik existieren verschiedene Kennzahlen. Mit diesen Kennzahlen gelingt es zum Beispiel in der deskriptiven Statistik, sich mit wenigen quantitativen Daten bereits eine gute Übersicht über Verteilungen, Mittelwerte etc. zu verschaffen. Im Folgenden sind einige Beispiele für statistische Kennzahlen beschrieben.

Arithmetisches Mittel: Mittelwert, der als Quotient aus der Summe der betrachteten Zahlen und ihrer Anzahl berechnet ist. Die zwei Zahlen 1 und 2 haben zum Beispiel den arithmetischen Mittelwert $1,5 = (1 + 2) / 2$. In der Statistik wird das arithmetische Mittel einer Stichprobe auch empirischer Mittelwert genannt.

Kurtosis ist eine Maßzahl für die Steilheit bzw. „Spitzigkeit“ einer Wahrscheinlichkeitsfunktion mit einem Gipfel, statistischer Dichtefunktion oder Häufigkeitsverteilung. Die Wölbung ist das zentrale Moment 4. Ordnung. Verteilungen mit geringer Wölbung streuen relativ gleichmäßig; bei Verteilungen mit hoher Wölbung resultiert die Streuung mehr aus extremen, aber seltenen Ereignissen.

Median: Der Median oder Zentralwert ist ein Mittelwert in der Statistik und ein Lageparameter. Der Median einer Auflistung von Zahlenwerten ist der Wert, der an der mittleren (zentralen) Stelle steht, wenn man die Werte der Größe nach sortiert. Beispielsweise ist für die Werte 4, 1, 37, 2, 1 die Zahl 2 der Median, nämlich die mittlere Zahl in 1, 1, 2, 4, 37.

Allgemein teilt ein Median einen Datensatz, eine Stichprobe oder eine Verteilung so in zwei (gleich große) Hälften, dass die Werte in der einen Hälfte nicht größer als der Medianwert sind, und in der anderen nicht kleiner. Der Median wird z. B. häufig für die Verteilung der Haushaltsvermögen in einem Land verwendet.

Modalwert: Als Modus oder Modalwert bezeichnet man in der Stochastik, einem Teilgebiet der Mathematik, eine Kennzahl der Verteilung einer Zufallsvariable oder eines Wahrscheinlichkeitsmaßes. Der Modus gehört zu den Lagemaßen und hat somit

wie der Erwartungswert und der Median die Aufgabe, die Position einer Verteilung zu charakterisieren.

Varianz: In der Stochastik ist die Varianz ein wichtiges Streuungsmaß der Wahrscheinlichkeitsverteilung einer reellen Zufallsvariablen. Sie beschreibt die erwartete quadratische Abweichung der Zufallsvariablen von ihrem Erwartungswert. Damit stellt die Varianz das zweite zentrale Moment der Zufallsvariablen dar. Die Quadratwurzel der Varianz wird Standardabweichung der Zufallsvariablen genannt.

2.2 Indikatoren

Indikatoren sind quantitative Informationen, die als Ersatzgröße Schlüsse auf die Ausprägung oder Veränderung anderer wichtiger Größen zulassen. Indikatoren sind häufig nicht direkt messbar und werden auch als „weiche Faktoren“ bezeichnet.

Kennzahlen und Indikatoren werden häufig zur Entscheidungsfindung, Strategieplanung und Absicherung verwendet. Dabei wird grundsätzlich zwischen absoluten und relativen Kennzahlen unterschieden, die jedoch meist aus Bilanzen oder anderen Unterlagen entnommen werden. Durch Kennzahlen und Indikatoren werden Werte und Prozesse dargestellt und abgebildet, die als Grundlage für Entscheidungen dienen. Neben der Anwendung für Management oder Controlling kann diese Methode auf Entwicklungsprojekte übertragen werden.

Eine wichtige Frage, die sich Projektleiter und Testmanager vor dem Test stellen ist, welche Eigenschaften ein Projekt hat und welche Parameter das Projekt zu welchen Anteilen beeinflussen. Diese Komplexität wird aber häufig durch Erfahrungswerte und Gefühl ausgedrückt. Die Komplexität eines Projekts bestimmt aber maßgeblich den Testumfang und die Auswahl der Testmethoden. Um dies auch objektiv und nachvollziehbar zu dokumentieren, kann man sich Kennzahlen und sogenannter Expertenaudits bedienen, um die Erfahrung der Projektmitarbeiter sichtbar und messbar darzustellen. Weiterführend wird hier klar werden, dass man nicht ausschließlich von dem Begriff „Kennzahlen“ sprechen kann, da nicht alle Faktoren für die Komplexität eines Projekts auf messbaren Größen beruhen, sondern häufig zusätzlich Schätzwerte, Einschätzungen und Meinungen einbezogen werden. Es werden also auch Indikatoren verwendet. Es zeigt sich jedoch, dass es sinnvoll ist, Projekte nicht nur technisch mit Kennzahlen messbar, sondern auch auf sozio-technischer Ebene durch Indikatoren zu bewerten, um die realen Einflüsse abzubilden.

2.3 Klassifizierung von Testmetriken

Eine grobe Klassifizierung teilt Metriken in Prozessmetriken und Produktmetriken ein:

- **Prozessmetriken** bewerten den Test als Dienstleistung und drücken dazu Indexwerte wie Testüberdeckung oder die Fehlerfindungsrate aus.
- **Produktmetriken** bewerten alle Elemente des Tests (Konzept- und Entwurfsdokumente, Source-Dateien, Testfälle usw.) und deren Beziehung untereinander.

Einige Metriken sind dabei sowohl Prozess- als auch Produktmetriken:

- **Projektmetriken** werden dazu verwendet, um die Effizienz eines Projektteams oder eines im Projekt eingesetzten Tools zu messen.

Eine andere Form der Unterteilung von Metriken führt zu folgenden Gruppen:

- Anforderungsbasierte Metriken (Beurteilung von Testumfang, Testbarkeit)
- testfallbasierte Metriken (Testspezifikation, Testdurchführung, Entwicklung der Testabdeckung, Qualität der Testfälle)
- Metriken zur Testautomatisierung
- Testproduktivitätsmetriken (kostenbasierte Metriken, Testeffektivität)
- fehlerbasierte Testmetriken (Fehleranzahl, Fehlerdichte, Entwicklung im Zeitverlauf)

Inhaltsbasierte Metriken prüfen den Informationsgehalt der einzelnen Anforderungen. Dabei handelt es sich um Metriken wie Vollständigkeit, Redundanzfreiheit und Notwendigkeit von Anforderungen. Kombinierte Metriken sind aus mehreren Metriken zusammengesetzt. Beispiele für kombinierte Metriken sind Metriken für Modifizierbarkeit oder Testbarkeit.

Je nach **Testklassen** bilden sich folgende Metriken heraus:

a) **Modultest:**

- Größe und Kopplung zu anderen Modulen
- Steuerflusskomplexität (Pfadkomplexität)
- Schnittstellenkomplexität (Anzahl und Art der Parameter je Funktionsaufruf)
- Datennutzungskomplexität (künstliche Parameter pro Kombination)

b) **Integrationstest**

- **Schnittstellendichte** (Generierung der Schnittstellen zwischen Systemkomponenten)
Die hier verwendeten Testfälle lassen sich aus den Schnittstellen zwischen den einzelnen Komponenten ableiten. Diese müssen während der Testdurchführung generiert werden und erzeugen mit zunehmender Anzahl höhere Komplexität. Daher sollte man bereits beim Systementwurf darauf achten, nur so viele Schnittstellen wie unbedingt erforderlich zu verwenden.

$$\text{Schnittstellendichte} = \frac{\text{Anzahl Komponenten}}{\text{Anzahl Komponenten} + \text{Anzahl Schnittstellen}}$$

- **Schnittstellenbreite** (Komponentenschnittstelle versus Messageschnittstelle)
Hier unterscheidet man zwei Arten von Schnittstellen. Bei einer Komponentenschnittstelle (z. B. CORBA, RMI, RPC) werden Parameterlisten übergeben. Im Falle einer message-orientierten Schnittstelle wird eine geschlossene Nachricht übergeben (z. B. XML, SOAP). Man verwendet in diesem Fall eine Metrik aus der

Nachrichtentechnik (Bandbreite) und misst die Anzahl der Zeichen, die auf einmal übertragen werden. Außerdem ermittelt man die Anzahl der (logischen) Attribute. Die Schnittstellenbreite kann somit wie folgt ermittelt werden:

$$\text{Schnittstellenbreite} = \text{Anzahl Schnittstellen} / (\text{Anzahl Schnittstellen} + \text{Anzahl Attribute})$$

- **Schnittstellentransparenz** (ASCII-Zeichenformat versus Bytestream)

Je klarer die Schnittstellen beschrieben sind, desto einfach ist die Erzeugung geeigneter Testdaten und Testfälle. Das ist zum Beispiel bei einer XML-Datei im ASCII-Zeichenformat der Fall, während ein Bitstring, dessen Struktur nur über eine zusätzliche Beschreibung erkennbar wird, eine geringe Transparenz aufweist.

- **Zugriffshäufigkeit** (Anzahl der Systemkomponenten, die z. B. auf Datenbanken zugreifen)

Je häufiger eine Komponente z. B. auf eine Datenbank zugreift, desto schwieriger wird die Erzeugung geeigneter Testfälle. Die Systemarchitektur hat auf diesen Parameter einen hohen Einfluss: Sind diese Zugriffe über mehrere Komponenten verteilt, so entsteht eine große Abhängigkeit. Eine zentrale Zugriffsschicht, also wenige Komponenten mit Anbindung zu Datenbanken, verringert die Zugriffshäufigkeit und erleichtert die Prüfung der restlichen Komponenten.

c) Systemtest

- Anwendungstestfall-Testbarkeit (Anwendungsfälle ergeben sich aus dem Nutzungsprofil des Softwaresystems und können mehrere Ausgänge haben)
- Benutzeroberflächen-Testbarkeit (Anzahl der Testfälle steigt mit Anzahl der Oberflächenobjekte)
- Systemschnittstellen-Testbarkeit (Generierung der Import- und Exportdaten)
- Datenbank-Testbarkeit (Generierung geeigneter Datensätze)
- Gesamt-Testbarkeit als Produkt aller Testbarkeitswerte [Snee2012]

Testfallbasierte Metriken lassen sich wie folgt weiter unterteilen:

a) **Testfallquantität:** absolute Anzahl aller Testfälle, Testprozeduren und Testszenarien, ergibt sich durch Analyse der

- Konzepte: Anwendungsfall, Ausprägung der Systemschnittstellen, spezifische Fehlermeldungen usw.
- Entwurfsdokumente: Abhängigkeiten, Vererbung, Ablaufverzweigung, Objektzustandswechsel usw.
- Source-Dateien: Entscheidungsweig, Schleifen, Ausnahmebedingungen usw.
- minimale Anzahl der Testfälle für jede Systemebene

Die Testfälle sollten in einer Datenbank bzw. Anwendung gespeichert werden. Eine professionelle, toolgestützte Testfallverwaltung ist dafür unbedingt zu empfehlen.