

Monatliche Kurzberichte zur Softwarequalität, deren derzeitige Wahrnehmung im Entwicklungsprozess und Verbesserungspotenzial*

Dominik Klein[♣], Jakob Rott[♣]

[♣]dominik.klein@tum.de – TU München, School of Computation, Information and Technology, Garching b. München
[♣]rott@cqse.eu – CQSE GmbH, München

Zusammenfassung

Die Wahrnehmung von Berichten über die interne Softwarequalität, die im Rahmen der Qualitätssicherung von Systemen teils kostspielig erstellt werden, ist in der Praxis nicht hinreichend untersucht. Während sich die Beachtung durch Manager beobachten lässt, ist gerade die von Dev-Teams empfundene Nützlichkeit und, ob Handlungen abgeleitet werden, unbekannt. Werden festgestellte Qualitätsdefizite nicht behoben, blieben z. B. Wartungsprobleme im System bestehen und eingesetzte Aufwände für die Berichterstellung wären in diesem Aspekt folgenlos. In diesem Artikel^ℵ wird vorgestellt, wie viele von 1 800, in Berichten genannten, Findings in industriellen Systemen später behoben wurden. Außerdem wird eine softwareassistenten-gestützte Methode vorgestellt, die es erlaubt, Qualitätsberichte zu verteilen, und von den Lesern Feedback einzuholen. Die prototypische Implementierung wurde in einem Konzern mit gewachsenem Applikationsportfolio evaluiert.

Es zeigte sich, dass etwa ein Viertel der berichteten Findings später behoben wurden. Die Möglichkeiten einer assistentengestützten Führung durch Qualitätsberichte wurde von den Studienteilnehmern begrüßt.

1 Einleitung

Der bloße Einsatz von Analysewerkzeugen im Qualitätssicherungsprozess bewirkt nicht automatisch, dass Findings (mgl. Qualitätsdefizite) behoben werden [1]. Es bedarf z. B. der entsprechenden Awareness und Zeit, um die Qualität zu erhalten, bzw. eine Verbesserung zu erzielen. Berichte können als Sicherheitsnetz dienen und auf übrige Auffälligkeiten hinweisen. Im Artikel wird zunächst das Berichtsformat sog. *Monthly Assessments* (MAs) eingeführt, anschließend anhand von etwa 2 600 derselben ihr Aufbau analysiert und untersucht, wie viele der, in den Berichten genannten, Findings später behoben wurden. Im zweiten Teil wird eine neue softwareassistenten-gestützte Methode vorgestellt, um MAs mit besserem Nutzererlebnis zu verteilen und einen geeigneten Rückkanal für Feedback an die MA-Autoren zu schaffen.

2 Das Berichtsformat »Monthly Assessment«

Monthly Assessments (MAs) sind ein Format für regelmäßige Entwicklungsberichte zur Softwarequalität. Nach mehrjähriger Anwendung wurden sie von Jürgens und Proft im Jahr 2020 auf der OOP-Konferenz vorgestellt [2]. Mit ihnen soll auf Abweichungen vom Qualitätsziel[†] aufmerksam gemacht werden. MAs betrachten jeweils die Softwareentwicklung aus dem Vormonat und verarbeiten dabei Ergebnisse von toolgestützten Qualitätsanalysen. Verfasst werden diese Kurzberichte meist von entwicklungsteamexternen Personen, den sog. *Quality Engineers* (QEng.). In MAs werden gerade solche Findings genannt, die bisher vom Dev-Team weder behoben noch als toleriert gekennzeichnet wurden und vom QEng. als behebens- oder prüfenswert eingestuft werden. MAs sind im Wesentlichen dreiteilig und sprechen durch Inhalt auf unterschiedlichen Abstraktionsebenen unterschiedliche Zielgruppen (Managementstufen und Rollen im Dev-Team) an.

- *Assessmentfarbe*: Der Qualitätstrend, vornehmlich anhand des vergangenen Monats bemessen, wird durch eine Ampelfarbe dargestellt. Eine *grüne* Bewertung bedeutet die Einhaltung des Qualitätsziels, eine *gelbe* bzw. *rote*, dass kleinere bzw. größere Abweichungen vom Qualitätsziel beobachtet wurden.
- *Kopfzeile*: Fasst das Assessment zusammen, weist ggf. auf Abweichungen vom Qualitätsziel hin.
- *Details*: Einzelne Findings oder häufige Findingstypen werden hervorgehoben und ggf. durch einen Vorschlag zur Behebung des Findings ergänzt.

MAs werden an Projektstakeholder per E-Mail verschickt und im Intranet des jeweiligen Auftraggebers als Webseite veröffentlicht, die eine Übersicht (inkl. MAs) über das gesamte Applikationsportfolio bietet.

	Compare	Delta	Finding	Metrics	Task
Kopfzeile	-	-	2	2	-
Details	376	666	2115	225	2

Tabelle 1: Genutzte Verweise innerhalb der untersuchten MAs. Meist wurde aus dem *Details*-Bereich verlinkt. *Compare* zeigt eine Vergleichsansicht von Quelltextduplikaten (Klonen); *Delta* bietet einen (meist Findings-)Vergleich zw. zwei Codeständen; *Finding* zeigt ein einzelnes Finding, sowie Metadaten und eine Erklärung; *Metrics* zeigt aggregierte Qualitätsmetriken auf Basis der Datei- und Ordnerstruktur; *Task* verweist auf eine, im Tool festgehaltene, Aufgabe zur Behebung von Findings.

[†]Ein Qualitätsziel legt z. B. den Fokus auf die Behebung von lediglich neu hinzugekommenen Findings bzw. auf die proaktive Behebung bestehender Findings (vgl. auch [2], S. 12).

*Das dem Artikel zugrundeliegende Vorhaben wurde mit Mitteln des Bundesministeriums für Bildung und Forschung unter dem Förderkennzeichen **Q-Soft, 01IS22001A** gefördert. Die Verantwortung für den Inhalt der Veröffentlichung liegt bei den Autoren.
^ℵAlle vorgestellten Ergebnisse und Techniken wurden im Rahmen der Bachelorarbeit von Dominik Klein an der Technischen Universität München erhoben und entwickelt.

Finding behoben innerhalb 45 Tage	182 (10,1%)
F. behoben später als 45 Tage danach	234 (13%)
Beständiges Finding	957 (53,2%)
Info technisch nicht nachvollziehbar	427 (23,7%)

Tabelle 2: Werden Findings, die in Berichten von Quality Engineers explizit referenziert werden, behoben? Hier gezeigt ist die rohe Behebungsverteilung über die letzten Jahre. Hinweis zur Interpretation: Nicht-Behebungen können begründet sein und wurden möglicherweise mit den QEngs. diskutiert.

3 Ist-Aufnahme: Wahrnehmung von MAs

In einer vorbereitenden Analyse von 2 604 MAs (exkl. MAs über entwicklungsfreie Zeit; inkl. »grüne« MAs meist ohne verlinkte Findings) zeigte sich, dass häufig auf Findings verwiesen wird, die ein Qualitätsziel verletzen (Tab. 1). Dies geschieht, um den Bericht zu kräftigen und Verbesserungsvorschläge anzubringen. Diese Vorstudie diente als Grundlage, um Verbesserungspotenzial im Prozess um die MAs aufzudecken. So gibt es nach dem Versand und der Veröffentlichung der MAs derzeit keinen geregelten Rückkanal, um festzustellen, ob und von wie vielen sie gelesen werden und wie hilfreich die Vorschläge für die Empfänger sind. Um die Wirkung annäherungsweise festzustellen, wurde untersucht, ob die Findings, auf die in den MAs verwiesen wurde, im Anschluss (45 Tage nach Berichtverteilung) von den Dev-Teams behoben wurden. Dies spräche dafür, dass der Bericht sowohl gelesen wurde, als auch hinsichtlich seiner Relevanz Zustimmung gefunden hat. Zudem wurden die Empfänger der MAs gebeten, die folgenden Fragen zu beantworten.

- Lesen Sie üblicherweise die MA-E-Mails?
- Wussten Sie, dass MAs von Quality Engineers verfasst und nicht automatisiert erstellt werden?
- Haben Sie bewusst Findings aus MAs behoben?

Findingsbehebungen Die etwa 2 600 MAs enthielten 1 800 Verweise auf konkrete Findings. Tabelle 2 zeigt, inwieweit diese nach der Verteilung der MAs behoben wurden. Etwa ein Viertel der Findings wurde behoben: 10% innerhalb und weitere 13% später als 45 Tage nach der Berichtverteilung. Grob die Hälfte der Findings verblieb in den Systemen. Wiederum gut ein Viertel der untersuchten Findings konnte aus technischen Gründen nicht mehr nachvollzogen werden.

Stakeholderbefragung Die Rücklaufquote des verbreiteten Fragebogens war gering. Von den 257 kontaktierten Projektbeteiligten antworteten nur 19. Die Rückmeldungen sind in Abbildung 1 dargestellt.

Schlussfolgerung Der Anteil der durch die Dev-Teams behobenen Findings spricht dafür, dass in den MAs relevante Findings genannt werden. Die Analyse zeigt Tendenzen, lässt allerdings keine klaren Ableitungen zu. Die Befragung deutet auf ein grundlegendes Interesse an den MAs hin, so werden entsprechende E-Mails gelesen und Findings bewusst behoben.



Abbildung 1: Ergebnisse der Stakeholderbefragung. Werden Verbreitungsmails gelesen? – Ist klar, dass MAs nicht automatisiert erstellt werden? – Wurden vom Befragten schon Findings aus MAs behoben?

Gleichzeitig wird Bedarf deutlich, besser aufzuklären, dass MAs von Experten erstellt werden, um so deren Wertigkeit klarer zu machen. Ein entwicklungsbegleitendes Tooling, das, sowohl fachliches als auch prozedurales, Feedback zu MAs zulässt, böte Gelegenheit der Frage nach der Wirksamkeit von MAs gezielter zu begegnen und Optimierungspotential zu finden.

4 Softwareassistent ⇒ Wirksamkeitserhöhung?

Auf Basis von *Teamscale*, einem Analysetool für Softwarequalität, wurde ein prototypischer Softwareassistent entwickelt, der es Nutzern ermöglicht, MAs interaktiv durchzuarbeiten. Absatzweise wird den Lesern ein MA präsentiert und jeder Teil kann separat bewertet werden: Einerseits durch einen Smiley (☺, ☹ und ☹), andererseits durch Freitexteingabe.

Verprobung Der MA-Assistent wurde für zweieinhalb Monate zur Evaluation beim o. g. Unternehmen eingesetzt. Stakeholder, insb. Entwickler, wurden aufgefordert, den Prototypen zu nutzen und ihre Erfahrung in einer Onlinebefragung rückzumelden. In einer fünfstufigen Likert-Skala (5: *stimme voll zu*, 1: *stimme überhaupt nicht zu*) wurde abgefragt, ob MAs aufmerksam durchgearbeitet würden – mit oder ohne Assistentenführung. Die Antworten finden sich in Abb. 2. Bei einem Viertel der Teilnehmern trüge der Assistent dazu bei, MAs mehr zu beachten. Obgleich keineswegs eine Allgemeingültigkeit abgeleitet werden kann, ist das Ergebnis vielversprechend und lädt ein, die Verteilung von MAs zu modernisieren. So könnten kleinteilige Rückmeldungen aus Dev-Teams stetig dazu beitragen, dass Folgeassessments an Relevanz gewinnen.

5 Diskussion und Ausblick

Diese Arbeit stellte heraus, dass einige der Findings, die in MAs von QEngs. genannt wurden, später aus der Codebasis entfernt werden. Zudem zeigt sich, dass die Auswertungsmöglichkeiten im bisherigen Prozess um die MAs beschränkt sind und diese durch einen Softwareassistenten deutlich erweitert werden könnten. Dieser würde zudem einen Feedbackkanal von Dev-Teams zu QEngs. aufbauen.

Eine zukünftige Arbeit kann der längere Einsatz des Prototypen sein, um die Nutzbarkeit weiter zu evaluieren und einen dauerhaften Einsatz anzustreben.

Literatur

- Elmar Jürgens. Vom Wiegen allein wird die Sau nicht fett-Erfahrungen aus einem Jahrzehnt Qualitätsanalyse in Forschung und Praxis. *Software Engineering*, 2020.
- Elmar Jürgens und Uwe Proft. Kosten-Nutzen-Berechnung von Qualitätsanalysen – Erfahrungen bei der Munich Re. cqse.eu/sl/nutzenberechnung-qa-munichre, 2020. OOP Konferenz.

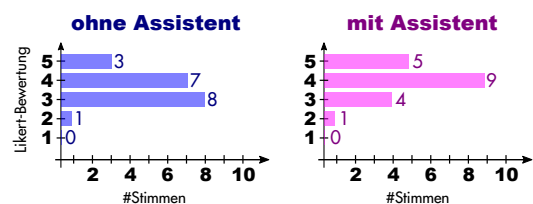


Abbildung 2: In der Umfrage, ob Beteiligte ein MA mit oder ohne Assistentenunterstützung aufmerksam durcharbeiten würden, zeigen sich unter Nutzung des Assistenten höhere Werte.