Ausarbeitung Softwarequalität

Software-Qualitätsbewertung mittels Software Metriken

Softwarequalität
Duale Hochschule Baden-Württemberg
Stuttgart

von
Paul Walker und Leon Kampwerth

Matrikelnummer: 3610783, 5722356
Abgabedatum: 22.03.2023

Inhaltsverzeichnis

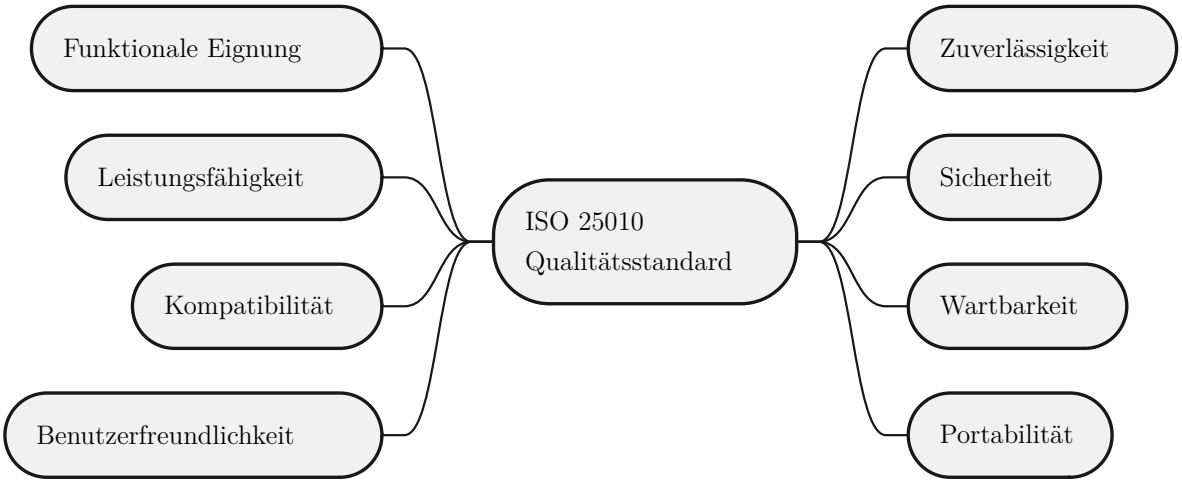

Abbildung 1: Softwarequalität nach ISO 25010 [4]

1 Einleitung

Die meisten Menschen arbeiten jeden Tag mit verschiedener Software. In der Arbeit mit Office-Software oder je nach Beruf spezialisierter Software für bestimmte Aufgaben und/oder Maschinen. Zu Hause nutzen viele Menschen Unterhaltungssoftware wie Spiele, oder Videoplattformen. Software ist heutzutage allgegenwärtig und die meisten Menschen, werden auch schon einmal Software schlechter Qualität genutzt haben. Das kann durch abruptes Schließen einer Applikation, unintuitives Design oder langer Ladezeiten für den Nutzer erkennbar sein. Es kann recht einfach sein, schlechte Qualitäten in Software zu erkennen, allerdings lässt sich Softwarequalität so nicht Objektiv vergleichen. Diese Arbeit beschäftigt sich damit Metriken zu finden, mit denen Software Qualität Objektiv gemessen werden kann.

2 Softwarequalität

Um Softwarequalität messen zu können muss zuerst festgelegt werden, was Softwarequalität überhaupt ist. Viele Aspekte von Softwarequalität sind für den Endnutzer gar nicht erkennbar oder sind nur dann erkennbar, wenn dieser Aspekt eine schlechte Qualität hat.

In dieser Arbeit wird die Definition von Softwarequalität nach International Organization for Standardization (ISO) 25010 [4] verwendet. In dem ISO 25010 Standard wird ein Produktqualitätsmodell für System und Softwarequalität. Nach diesem Modell wird die Qualität eines Softwaresystems oder Produkts durch acht Charakteristika definiert. Diese Charakteristika werden in Abbildung 1 dargestellt.

Jedes dieser Charakteristika wird weiter in Subcharakteristika aufgeteilt. In dieser Arbeit wird der Fokus auf zwei bestimmte Charakteristika gesetzt und für jedes eine Metrik vorgestellt, mit der die Qualität der Charakteristik gemessen werden kann.

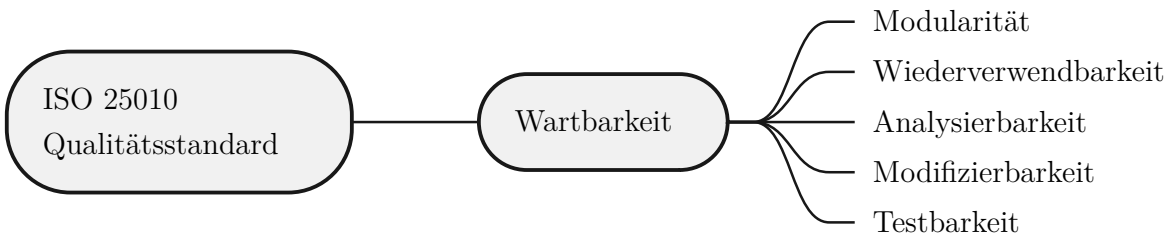

Abbildung 2: Wartbarkeit nach ISO 25010 [4]

2.1 Wartbarkeit

Der Grad der Wartbarkeit ist nach ISO, wie effektiv und effizient das Produkt oder System verändert werden kann [4]. Diese Veränderungen können Korrekturen, Verbesserungen oder Anpassungen der Software auf Änderungen des Einsatzgebiets oder auf Änderungen der Anforderungen sein [4]. Die Wartbarkeit bezieht sich auch auf die Installation von Updates [4]. Die Wartbarkeit ist nach ISO in weitere Subcharakteristika aufgeteilt. Diese Subcharakteristika sind in Abbildung 2 dargestellt.

2.1.1 Modularität

Die Modularität wird daran gemessen, inwieweit das System oder Programm aus einzelnen Teilen besteht und Änderungen an einem Teil der Software andere Teile nur minimal oder gar nicht beeinflussen [4]. Module sind Softwareattribute, die Software in unabhängige Einzelteile trennen [6].

2.1.2 Wiederverwendbarkeit

Die Wiederverwendbarkeit ist der Grad, zu dem ein Teil des Systems in mehr als einem System oder beim Erstellen von weiteren Teilen des Systems verwendet werden kann [4]. Die Wiederverwendung bezieht sich dabei nicht nur auf Software, sondern auch auf Dokumentation, Software-Design, Applikationstests und weitere Teile des Systems [3].

2.1.3 Analysierbarkeit

Wie effektiv und effizient es möglich ist, den Einfluss auf ein Produkt oder System abzuschätzen, wenn ein oder mehrere Teile verändert werden, ist durch die Analysierbarkeit definiert [4]. Außerdem wie gut der Ursprung von Fehlern im Produkt oder System diagnostiziert werden kann [4].

2.1.4 Modifizierbarkeit

Die Modifizierbarkeit ist ein Maßstab dafür wie effektiv und effizient sich ein Produkt oder System ändern lässt, ohne dabei Defekte einzuführen oder die Produktqualität zu verschlechtern [4].

Die Modifizierbarkeit kann durch die *Modularität* und die *Analysierbarkeit* beeinflusst werden [4].

2.1.5 Testbarkeit

Der Grad der Testbarkeit ist, wie effektiv und effizient Testkriterien für ein Produkt oder System definiert werden können und wie effektiv und effizient Tests für das Produkt oder System durchgeführt werden können [4]. Zusätzlich zur Erstellung der Testkriterien und der Durchführung der Tests, spielt es auch eine Rolle wie gut objektive und messbare Tests erstellt werden können, die feststellen, ob ein bestimmtes Kriterium erfüllt wurde [6].

2.2 TODO @Leon Charakteristika X

3 Software Metriken

Software Metriken sind quantitative Maße, die zur Bewertung von Softwareprodukten und Softwaresystemen genutzt werden können. Diese Metriken können verschiedene Aspekte der Software Bewerten. Hier werden Metriken vorgestellt, mit denen die Wartbarkeit und die analysiert werden könne.

Software Metriken geben sowohl den Entwicklern als auch den Stakeholdern der Software Informationen über Eigenschaften der Software. Diese Informationen können nützlich sein, um Risiken frühzeitig zu erkennen, Verbesserungen und Verschlechterungen zu quantifizieren und verschiedene Software vergleichen zu können. Software Metriken können in der Kommunikation zwischen Entwicklern und Stakeholdern helfen, da sie helfen Eigenschaften der Software zu objektifizierten und sich so beispielsweise bestimmte Entwicklungstätigkeiten, wie das Refactoring von Software besser gegenüber den Stakeholdern rechtfertigen lassen.

Mit Softwaremetriken die, die Softwarequalität quantifizieren können Regeln eingeführt werden, die bestimmte Mindestanforderungen an die Softwarequalität stellen. So kann in der Theorie eine konstante objektiv gute Qualität der Software gewährleistet werden. Um diese Regeln umzusetzen ist es hilfreich die nötigen Metriken automatisch zu generieren und Änderungen an der Software nur dann zu erlauben, wenn die Regeln eingehalten werden.

3.1 Wartbarkeit

Viele Charakteristika der Wartbarkeit lassen sich über die Dauer oder Effizienz von Wartungsaufgaben messen. Es kann beispielsweise die Zeit gemessen werden, die benötigt wird, um eine Änderung an der Software durchzuführen, um damit die Modifizierbarkeit zu charakterisieren [2].

Diese Metriken lassen sich allerdings nur bei konkreten Wartungsaufgaben messen. Das ist problematisch, da so die Qualität der Software nur durch großen Aufwand bestimmt werden kann. Ziel hier ist es daher eine Metrik zu finden, mit der sich die Wartbarkeit

ohne großen Aufwand bestimmen lässt. Das Paper *A Practical Model for Measuring Maintainability* von Ilja Heitlager, Tobias Kuipers und Joost Visser [2] stellt eine Möglichkeit dar, die Wartbarkeit ohne großen Aufwand zu messen. Allerdings basiert diese Paper auf dem älteren ISO 9126 Qualitätsmodell [2, 5]. In dieser Arbeit wird die Metrik, die in [2] daher erweitert und auf das neuere ISO 25010 Qualitätsmodell angepasst.

Für diese Metrik werden einige einfache Kennzahlen der Software kombiniert, um Kennwerte für die Jeweiligen Subcharakteristika der Wartbarkeit zu berechnen. Diese Kennwerte können dann verwendet werden, um einen Gesamtwert für die Wartbarkeit der Software zu bekommen. Die Kennzahlen die Genutzt werden sind.

Quellcodevolumen

Die Zahl der effektiven Quellcodezeilen im gesamten Bereich der analysiert werden soll. Je geringer das Quellcodevolumen ist, desto einfacher ist es eine Übersicht über den Bereich zu bekommen, der analysiert werden soll. Also ist niedriger besser.

Komplexität der Einheiten

Die zyklomatische Komplexität einer Einheit sagt aus wie Komplex eine Einheit ist, und damit auch wie komplex Wartungsaufgaben sein können. Eine Einheit ist dabei das kleinste ausführbare Stück Code [2]. Je nach Programmiersprache kann eine Einheit eine Funktion, eine Klasse oder aber das gesamte Programm sein [2]. Eine geringere Komplexität spricht für besser wartbare Software. Die zyklomatische Komplexität von Software kann mit Tools wie *scc* [1] automatisiert gemessen werden. Hier ist also niedriger besser.

Quellcode Duplikation

Quellcode Duplikation ist definiert durch die Anzahl Quellcode Zeilen, die eins zu eins an verschiedenen Stellen zu finden sind. Duplikation spricht für Probleme in der Software. Also ist für eine höhere Qualität der Software eine niedrige Duplikation besser.

Einheitengröße

Eine geringe Einheitengröße sorgt für eine höhere Softwarequalität, da kleinere Einheiten einfacher zu warten sind.

Unit-Testabdeckung

Die Testabdeckung durch Unit-Tests wird häufig verwendet, um Aussagen über die Softwarequalität zu machen. Hier wird diese Metrik auch verwendet. Allgemein gilt, je mehr Unit-Tests vorhanden sind, desto besser.

Modulanzahl

Die Modulanzahl gibt Aussage über die Modularisierung von Software. Je höher die Modulanzahl, desto modularisierter ist die Software. Hier ist also höher besser.

Wiederverwendung von Einheiten

Die Wiederverwendung von Einheiten sagt aus, wie oft eine Einheit an anderer Stelle im Quellcode wiederverwendet wird. Das kann gemessen werde, indem gezählt wird wie oft Klassen genutzt werden oder wie oft Funktionen aufgerufen werden. Hier gilt, höher ist besser.

Diese Kennzahlen werden wie folgt kombiniert um Metriken für die Subcharakteristika zu erhalten.

Modularität

Eine einfache Metrik für die Modularität ist die Modulanzahl in dem betrachteten Teil der Software. Nachteil dieser Metrik ist, dass sie nicht aussagt, inwieweit Änderungen an einem Modul die anderen Module beeinflussen. Nach ISO 25010 sollen Änderungen an einem Modul, die anderen Module möglichst wenig beeinflussen [4].

Eine andere Metrik ist die Einheitengröße. Diese gibt auch eine gewisse objektive Aussage über die Modularität. Für eine hohe Modularität ist eine geringe Kopplung und hohe Kohäsion, also hohe Zusammengehörigkeit, innerhalb der Einheiten nötig um zu gewährleisten, dass Änderungen an einem Modul, die anderen Module minimal oder gar nicht beeinflussen. Viele kleine Einheiten sprechen für eine hohe Kohäsion, da in einer kleinen Einheit nur wenig Funktionalität, die dann aber stark zusammenhängt, implementiert werden kann.

Sowohl die Modulanzahl als auch die Einheitengröße werden für die Modularität-Metrik verwendet.

Wiederverwendbarkeit

Die Wiederverwendbarkeit lässt sich nicht direkt messen, allerdings kann gemessen werden wie oft Softwarekomponenten tatsächlich wiederverwendet werden, also die Wiederverwendung. Bei hoher Wiederverwendung muss auch die Wiederverwendbarkeit hoch sein. Eine niedrige Wiederverwendung ist ein starker Indikator für niedrige Wiederverwendbarkeit.

Eine weitere Metrik für die Wiederverwendung ist die Code-Duplikation und die Einheitengröße. Wenn hohe Code-Duplikation vorhanden ist, ist die Wiederverwendung in der Regel gering, da duplizierter Code oft wiederverwendbarer implementiert werden kann. Bei geringer Code-Duplikation kann davon ausgegangen werden, dass eine hohe Wiederverwendung des Codes gegeben ist. Für eine hohe Wiederverwendbarkeit ist es wichtig, dass Einheiten nur wenige Aufgaben haben. Eine geringe Einheitengröße deutet daher auf eine höhere Wiederverwendbarkeit.

Analysierbarkeit

Nach [2] werden Code-Volumen, Code-Duplikation, Einheitengröße und Unit-Testabdeckung verwendet, um die Analysierbarkeit zu quantifizieren. Geringes Code-Volumen und geringe Einheitengröße hilft schnell einen Überblick über den Quellcode zu bekommen. Unit-Tests helfen die Funktionsweise des Quellcodes zu verstehen. Eine hohe Quellcode-Duplikation erschwert hingegen das Verständnis.

Modifizierbarkeit

In [2] wird die Komplexität der Einheiten und die Code-Duplikation als Metrik für die Modifizierbarkeit genannt. Code-Duplikation verringert die Modifizierbarkeit, da an Stellen mit dupliziertem Code, alle Teile des duplizierten Codes gleichermaßen geändert werden müssen. Das erhöht den Modifikations-Aufwand stark und erhöht die Chance, dass bei Modifikationen vergessen wird Teile des duplizierten Codes zu ändern und so Fehler im Programm entstehen. Eine hohe Komplexität der Einheiten sorgt dafür, dass es schwieriger ist diese Einheiten zu modifizieren, ohne Fehler in der Software zu erzeugen.

Testbarkeit

Die Testbarkeit von Software hängt nach [2] unter anderem von der Komplexität der Einheiten, der Einheitengröße und der Zahl der Unit-Testabdeckung ab. Eine hohe Unit-Testabdeckung ist ein Indikator für gute Testbarkeit, denn das bedeutet, dass die Software im allgemeinen testbar ist. Kleine Einheitengröße sorgt auch für eine bessere Testbarkeit, da es einfacher ist für kleine Einheiten, die weniger Aufgaben haben, Testfälle zu spezifizieren. Die Komplexität der Einheiten spielt auch eine Rolle, da für komplexe Einheiten voraussichtlich mehr und komplexere Testfälle spezifiziert werden müssen.

In Tabelle 1 wird die Zugehörigkeit der zuvor vorgestellten Quellcode Kennzahlen zu den Wartbarkeit-Subcharakteristika dargestellt. Diese Tabelle erweitert die in [2] vorgestellte Tabelle um die Änderungen die der neuere ISO 25010 Qualitätsstandard gegenüber dem alten ISO 9126 Standard bringt. Damit kann nun Software verglichen werden und Qualitätsänderungen durch Änderungen an der Software gemessen werden. Im Gegensatz zu [2] wird hier keine konkrete Kennzahl für die Wartbarkeit vorgestellt. Vielmehr können mit dieser Metrik Softwarestände und Softwaresysteme relativ zueinander verglichen werden. Dafür werden die relativen Unterschiede in den einzelnen Metriken eins zu eins nach der Tabelle verrechnet, um einen Wert für die Jeweilige Subcharakterstik der Wartbarkeit zu erhalten. Um daraus einen Wert für die Wartbarkeit zu erhalten, werden die Werte der Subcharakteristika eins zu eins verrechnet.

Um diese Metrik zu verbessern, kann die Verrechnung der einzelnen Quellcode-Kennzahlen gewichtet werden. Diese Entscheidung kann je nach Anwendung der Metrik getroffen werden.

	Quellcodevolumen	Komplexität der Einheiten	Quellcode Duplikation	Einheitengröße	Unit-Testabdeckung	Modulanzahl	Wiederverwendung von Einheiten
Modularität				x		x	
Wieder-verwendbarkeit			x	x			x
Analysierbarkeit	x		x	x	x		
Modifizierbarkeit		x	x				
Testbarkeit		x		x	x		

Tabelle 1: Zugehörigkeit der Kennzahlen zu den Wartbarkeit-Subcharakteristika

Es ist denkbar diese Metrik automatisch durch Softwaretools zu generieren. Das ermöglicht es diese Metrik, als Auswertung von Software in ein Continuous Integration / Continuous Delivery (CI/CD) System zu integrieren. Damit können Änderungen an der Software ausgewertet werden und gewarnt werden, wenn die Wartbarkeit durch Änderungen an der Software verschlechtert wird. Im extremsten Fall, kann eine Änderung auch abgelehnt werden, wenn die Wartbarkeit verschlechtert wird.

3.2 TODO @Leon Metrik X

4 Ausblick

Die hier vorgestellte Metrik zur Wartbarkeit macht es zwar einfach die Wartbarkeit von Softwareprodukten und Softwaresystemen zu messen, allerdings hat diese Metrik auch einige Nachteile. So kann diese Metrik keine direkte Aussage über die Dauer von Wartungsaufgaben an der Software machen. Außerdem wurde nicht verglichen wie sich diese Metrik im Vergleich mit anderen Wartbarkeit-Metriken verhält. Es ist denkbar, dass die hier vorgestellte Metrik völlig andere Werte als andere Wartbarkeit-Metriken liefert. In einer weiteren Arbeit könnte darauf eingegangen werden und die hier vorgestellte Wartbarkeit-Metrik mit andere Wartbarkeit-Metriken verglichen werden, um festzustellen, inwieweit sich diese unterscheiden.

Außerdem bietet die hier vorgestellte Wartbarkeit-Metrik die Möglichkeit Gewichte zu setzen, um die Metrik weiter zu verbessern und anzupassen. In einer weiteren Arbeit könnte diese Metrik durch die Gewichtungen an andere Wartbarkeit-Metriken angepasst

werden.

In einer weiteren Arbeit ist es außerdem Denkbar weitere Metriken zu finden, mit denen die übrigen Qualitätscharakteristika einfach quantifiziert werden können.

Literatur

[1] Ben Boyter. *Sloc Cloc and Code (scc)*. `https://github.com/boyter/scc`. 2022.

[2] Ilja Heitlager, Tobias Kuipers und Joost Visser. „A Practical Model for Measuring Maintainability“. In: *6th International Conference on the Quality of Information and Communications Technology (QUATIC 2007)*. IEEE. 2007, S. 30–39. DOI: `10.1109/QUATIC.2007.8`.

[3] „IEEE Standard for Information Technology–System and Software Life Cycle Processes–Reuse Processes“. In: *IEEE Std 1517-2010 (Revision of IEEE Std 1517-1999)* (2010), S. 1–51. DOI: `10.1109/IEEESTD.2010.5551093`.

[4] ISO 25010. *Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) - System and software quality models*. Standard. Geneva, CH: International Organization for Standardization, 2011.

[5] ISO 9126-1. *Information technology — Software product quality. Part 1: Quality model*. Standard. Geneva, CH: International Organization for Standardization, 2000.

[6] „ISO/IEC/IEEE International Standard - Systems and software engineering–Vocabulary“. In: *ISO/IEC/IEEE 24765:2017(E)* (2017), S. 1–541. DOI: `10.1109/IEEESTD.2017.8016712`.