

Gottfried Wilhelm Leibniz Universität Hannover
Fakultät für Elektrotechnik und Informatik
Institut für Praktische Informatik
Fachgebiet Software Engineering

**Quantitative und qualitative Messung von
Software-Anforderungen**

Bachelorarbeit

im Studiengang Informatik

von

Christian El Boustani

Prüfer: Prof. Dr. Kurt Schneider
Zweitprüfer: Prof. Dr.-Ing. Gabriele von Voigt
Betreuer: M. Sc. Eric Knauss

Hannover, 28. Januar 2008

Erklärung

Hiermit versichere ich, dass ich die vorliegende Bachelorarbeit selbstständig und ohne fremde Hilfe verfasst habe und keine anderen als die in der Arbeit angegebenen Quellen und Hilfsmittel verwendet habe.

Hannover, 28. Januar 2008

Christian El Boustani

Danksagung

Hiermit möchte ich mich bei allen Personen bedanken, die mich bei der Erstellung dieser Arbeit unterstützt haben.

Zusammenfassung

Bei der Vergleichbarkeit von Softwareprojekten spielt die Vergleichbarkeit der Anforderungen eine zentrale Rolle. Gängige Verfahren wie z.B. Function Points oder Use Case Points liefern zwar Kennzahlen, jedoch keine Aussagen über die Qualität der Anforderungen. Wie können also Projekte im Rahmen des Qualitätsmanagements qualitativ untereinander verglichen werden?

Im Rahmen dieser Arbeit werden mittels der GQM-Methode Metriken erstellt, um die Anforderungen von Softwareprojekten qualitativ miteinander zu vergleichen. Diese Metriken werden auf die Anforderungsdokumente von insgesamt 16 Softwareprojekten angewendet.

Auf Basis der Messergebnisse werden Einschätzungen bezüglich der Qualität der Anforderungen der jeweiligen Projekte getroffen. Die einzelnen Ergebnisse eines Projekts werden zusammengefasst, um die Softwareprojekte qualitativ miteinander vergleichen zu können.

Diese Arbeit zeigt auf, welche Qualitätsmerkmale von Anforderungen isoliert anhand des eines Anforderungsdokuments ermittelt werden können und wo die Grenzen der Messbarkeit liegen. Es wird auf die Qualität der Metriken eingegangen und dabei untersucht, welche Faktoren die Messergebnisse beeinflussen können. Es wird eine Einschätzung abgegeben, welche Verbindung zwischen der Qualität eines Anforderungsdokuments und dem Projekterfolg besteht. Ferner kann bestätigt werden, dass die Verwendung eines Formulierungsregelwerks sich positiv auf die Qualität von Anforderungen auswirken kann.

Inhalt

1 Einleitung	6
1.1 Motivation.....	6
1.2 Aufgabenstellung.....	6
1.3 Aufbau dieser Arbeit.....	7
1.4 Ziele dieser Arbeit.....	7
2 Grundlagen	8
2.1 Metriken in der Softwarequalität.....	8
2.2 Die GQM-Methode.....	9
2.3 Qualität von Anforderungen.....	9
2.4 Metriken zur Messung der Anforderungsqualität.....	10
2.5 Arten von Anforderungen.....	11
2.5.1 Funktionale Anforderungen.....	11
2.5.2 Nicht-funktionale Anforderungen.....	11
3 Erstellung eines Messplans	12
3.1 Vorgehensweise.....	12
3.2 Ziele (nach GQM)	13
3.3 Fragen (nach GQM).....	21
3.4 Metriken und Indikatoren (nach GQM).....	23
3.5 Aufstellen des Messplans.....	26
3.5.1 Auswahl der Metriken.....	26
3.5.2 Der Messplan.....	29
4 Analyse von Projekten	37
4.1 Vorgehensweise.....	37
4.2 Messung mittels Fragebogen-Programm.....	38
4.3 Auswertung	41
4.3.1 Messergebnisse.....	41
4.3.2 Abgleich der Ergebnisse mit den Hypothesen.....	44
4.3.3 Gewichtung und Vergleich der Ergebnisse.....	46
4.3.4 Zusammenhang zwischen Messergebnis und Projekterfolg.....	52
5 Bewertung der Metriken	53
5.1 Einschränkungen der Allgemeingültigkeit.....	54
5.1.1 Heuristiken / Metriken zur sprachlichen Qualität.....	55
5.1.2 Metriken zur inhaltlichen Qualität	58
5.1.3 Metriken für nicht-funktionale Anforderungen	59
5.2 Optimierungsmöglichkeiten.....	61
6 Fazit	63
7 Ausblick	64
8 Literaturverzeichnis	65

1 Einleitung

1.1 Motivation

Software stellt eine besondere Art von Werkstoff dar. Durch das Fehlen einer festen Struktur und der kaum vorhandenen Einschränkungen auf den Anwendungsbereich ist jede Software einzigartig.

Eine wesentliche Rolle in einem Softwareprojekt stellen die Anforderungen dar, welche die Funktionen und Leistungsmerkmale eines zu entwickelnden Programms [Bro08] darstellen. Will man zwei Softwareprojekte miteinander vergleichen, so stellt die Vergleichbarkeit von Anforderungen eine besondere Herausforderung dar.

Die Frage, wie man Anforderungen miteinander vergleichen kann, ist sowohl in der Aufwandsabschätzung für Softwareprojekte (wo man versucht, basierend auf vergleichbaren Projekten, die Dauer und die Kosten frühzeitig zu ermitteln), im Qualitätsmanagement (wo es um die Beurteilung der Qualität – auch im Vergleich zu andern Softwareprojekten geht) und beim Projektmanagement (wo die Messung eines Softwareprojekts z.B. mittels einer Kennzahl wünschenswert wäre) von zentraler Bedeutung.

Um Anforderungen bereits in frühen Projektphasen – wie während der Anforderungsanalyse – vergleichbar zu machen, gibt es bereits verschiedene Ansätze, die z.B. im Rahmen der Aufwandsabschätzung mittels Function Points oder Use Case Points Kennzahlen liefern. Jedoch stellt die Frage nach der Qualität der Anforderungen Grenzen dieser Methoden dar, auch die Interpretation dieser Daten ist oft schwierig.

Im Rahmen der Forschung wird versucht quantitative Daten zu diesen Fragen zu ermitteln. Auch im Rahmen von verschiedenen Arbeiten am Fachgebiet Software Engineering wurden bereits Anforderungen in einer Menge von Softwareprojekten quantitativ analysiert.

1.2 Aufgabenstellung

Ausgehend von einem Datenbestand von 16 Softwareprojekten sollen die Anforderungen von verschiedenen Projekten quantitativ und qualitativ analysiert und verglichen werden. Hierzu werden mittels der GQM-Methode Metriken zur Erfassung verschiedener Qualitätsaspekte von Anforderungen erstellt, welche in einen Messplan einfließen, der auf konkrete Projekte angewendet wird.

1.3 Aufbau dieser Arbeit

Diese Arbeit ist in sieben Kapitel gegliedert: Nach der Einleitung, folgt in Kapitel 2 die Einführung in die Grundlagen, wie z.B. Software-Metriken und das GQM-Modell. Außerdem werden die Aspekte eingeführt, die für die Qualität von Anforderungen maßgeblich sind, es werden Metriken zur Messung von Softwarequalität vorgestellt, wie sie bereits in der Literatur zu finden sind.

In Kapitel 3 wird mittels der GQM-Methode ein Messplan erstellt, der die Basis für die Analyse in Kapitel 4 bildet. Hierbei wird detailliert und schrittweise die Vorgehensweise erläutert.

Kapitel 4 beinhaltet die praktische Analyse von Anforderungsdokumenten aus den bekannten Softwareprojekten des FG Software Engineering. In der Auswertung wird auch eine differenzierte Gewichtung der einzelnen Teilergebnisse vorgenommen. Es wird eine Bilanz über die Gesamtqualität der Anforderungen eines Projekts gezogen.

Kapitel 5 nimmt eine Bewertung der Metriken vor und untersucht inwiefern die Messergebnisse zwischen verschiedenen Prüfern voneinander abweichen und welche Faktoren dazu führen.

In Kapitel 6 wird ein Fazit über die Ergebnisse dieser Arbeit gezogen.

Kapitel 7 schätzt ein, welche Aspekte dieser Arbeit in zukünftigen Arbeiten vertieft und weitergehend behandelt werden könnten.

1.4 Ziele dieser Arbeit

Ziel dieser Arbeit ist vor allem die Erstellung eines Messplans zur Analyse von Anforderungsdokumenten. Dieser Messplan soll auf konkrete Projekte angewendet werden; dadurch sollen Aussagen über die Qualität der Anforderungen in diesen Projekten getroffen und deren Anforderungen qualitativ vergleichbar gemacht werden.

2 Grundlagen

2.1 Metriken in der Softwarequalität

Eine Metrik ist eine „Funktion, die eine Softwareeinheit oder den Vorgang ihrer Erstellung in einen Zahlenwert oder in Zeichen (eventuell mit einer Einheit) abbildet. Dieser Wert zeigt an, wie gut ein zugehöriges Qualitätsziel erfüllt ist.“ [Sch07]

Bekannte Metriken sind z.B. die *Zyklomatische Komplexität* (oder auch „*McCabe-Metrik*“ genannt), die die Komplexität eines Programnteils anhand der Programmpfade misst, oder die *LoC-Metrik* (*Lines of Code*), welche die Anzahl der Quellcodezeilen misst.

Um die Messergebnisse einer Metrik zu interpretieren und sie vergleichen zu können, ist entscheidend, auf welcher Skala sie sich befinden. Je nach Skala sind unterschiedliche Operationen erlaubt.

Die wichtigsten Skalen sind [Sch07]:

- **Nominalskala:** Einzelne Werte auf dieser Skala sind unterscheidbar, es sind jedoch keinerlei mathematische Operationen oder Vergleiche erlaubt.
- **Ordinalskala:** Die Werte auf der Ordinalskala sind in einer Reihenfolge angeordnet, es lassen sich also Aussagen treffen, wie z.B. „Wert 1 ist besser als Wert 2“. Auf Grund der unterschiedlichen Abstände zwischen den Werten sind jedoch mathematische Operationen nicht erlaubt [Bro08].
- **Intervallskala:** Bei der Intervallskala sind die Abstände zwischen den Werten der Skala gleich groß und stehen für „gleiche Unterschiede im Merkmal eines Objekts“ [Bro08]. Mathematische Operationen sind bis auf die Multiplikation erlaubt, unzulässig sind jedoch numerische Vergleiche.
- **Rationalskala:** Auf einer Rationalskala sind Multiplikationen und Divisionen erlaubt. Auch dürfen Zahlenwerte zueinander in Verhältnis gesetzt werden.

Die Interpretation von Messergebnissen ist jedoch nicht immer einfach, und zum Teil kann es passieren, dass man am Ende der Messung zwar viele Zahlen ermittelt hat, jedoch nicht sicher weiß, was sie aussagen, bzw. wie sie mit den Qualitätszielen zusammenhängen [Sch07].

2.2 Die GQM-Methode

Problematisch bei der Anwendung von bekannten Metriken ist, dass man dazu neigt, die Eigenschaften zu messen, die einfach zu messen sind. Was jedoch sagen diese gemessenen Werte genau aus?

Die *Goal-Question-Metric*-Methode (im Folgenden *GQM* genannt) geht nun den umgekehrten Weg. Das Prinzip ist: „Man soll nicht das messen, was leicht zu messen ist, sondern das, was man braucht, um seine Verbesserungsziele zu erreichen.“ [Sch07].

GQM dient also dazu, eigene Metriken zu erstellen, die auf die eigenen Qualitätsziele abgestimmt sind.

Das prinzipielle Vorgehen ist:

1. Ziele definieren (**G**oal)
2. Ziele näher beschreiben durch Zielfacetten
3. Fragen aus den Zielen ableiten (**Q**uestion)
4. Metriken und Erhebungsmethoden zu den Fragen sammeln (**M**etric)
5. Formulare und Fragelisten erstellen, die in einen Messplan einfließen
6. Erhebung und Analyse mit Messplan

2.3 Qualität von Anforderungen

Um die Qualität von Anforderungen zu analysieren, existieren in der Literatur bereits einige Qualitätskriterien, die im Folgenden näher betrachtet werden sollen:

Folgende Qualitätskriterien zeichnen nach [Rup06] „exzellente Anforderungen“ aus:

- **Vollständig**
Jede Anforderung muss die geforderte und zu liefernden Funktionalität vollständig beschreiben.
- **Korrekt**
Eine Anforderung gilt als korrekt, wenn sie vollständig die Vorstellung des Stakeholders, der sie formuliert hat, wiedergibt.
- **Klassifizierbar**
Die rechtliche Relevanz einer Anforderung muss festgelegt sein.
- **Konsistent**
Eine Anforderung muss gegenüber allen anderen Anforderungen widerspruchsfrei sein; auch eine Anforderung in sich darf keine Widersprüche aufwerfen.
- **Prüfbar**
Eine Anforderung muss so beschrieben sein, dass sie testbar ist, d.h. die geforderte Funktionalität muss durch Messung nachweisbar sein.
- **Eindeutig**
Eine Anforderung darf nur auf eine Art und Weise verstanden werden; es darf nicht möglich sein, andere Sachverhalte hineinzuzinterpretieren.
- **Verstehbar**
Die Anforderungen müssen für alle Stakeholder verstehbar sein.
- **Gültig und aktuell**
Die Anforderung muss die Realität des Systems beschreiben – dies impliziert auch die Aktualität der Beschreibung.

- **Realisierbar**
Eine Anforderung muss bezüglich der technischen, zeitlichen und finanziellen Restriktionen umsetzbar sein.
- **Notwendig**
Eine Anforderung sollte nur jene Leistungen und Eigenschaften fordern, die (vom Kunden) tatsächlich gefordert werden.
- **Verfolgbar**
Jede Anforderung muss für sich eindeutig zu identifizieren sein.
- **Bewertet**
Eine Anforderung muss nach Wichtigkeit oder Priorität bewertet werden können.

2.4 Metriken zur Messung der Anforderungsqualität

Es existieren bereits Metriken, um die Qualität einer Anforderung bzw. des gesamten Anforderungsdokuments zu analysieren. [Rec+06] unterscheidet in:

- Textbasierte Metriken: Bewertung des Inhalts einer Anforderung

- $$\text{Eindeutigkeit} = \frac{\sum (\text{Anforderungen ohne Defekte})}{\sum (\text{Anforderungen})}$$

Die Eindeutigkeit einer Anforderung wird durch einen so genannten „sprachlichen Defekt“ herabgesetzt. Das Verhältnis der Summe von Anforderungen ohne sprachliche Defekte zu der Gesamtzahl der Anforderungen bestimmt die „Eindeutigkeit“.

- *Anzahl der Passivsätze*

Passive Formulierungen begünstigen das bewusste oder unbewusste Unterschlagen von wichtigen Informationen. Enthält eine Dokument also eine hohe Anzahl an Passivsätzen, so ist dies ein Indikator für unvollständige Anforderungen.

- $$\text{Klassifizierbarkeit} = \frac{\sum (\text{Klassifizierbare Anforderungen})}{\sum (\text{Anforderungen})}$$

Die Klassifizierbarkeit trifft eine Aussage über die Anzahl der Anforderungen, die eine Aussage über ihre rechtliche Verbindlichkeit treffen. Hierzu muss vorher festgelegt sein, wie eine rechtliche Verbindlichkeit gekennzeichnet ist.

- Strukturbasierte Metriken: Bewertung der Struktur von Anforderungen innerhalb eines Dokuments

- $$\text{Verfolgbarkeit} = \frac{\sum (\text{Anforderungen mit ID})}{\sum (\text{Anforderungen})}$$

Die Verfolgbarkeit untersucht, ob für jede Anforderung eine eindeutige Nummer (ID) vergeben worden ist, über die die Anforderung referenziert werden kann. Hierzu ist es notwendig, dass jede Anforderung klar gegen andere Textteile des Dokuments abgegrenzt ist.

$$\circ \quad \text{Traceability} = \frac{1}{\sum (\text{fehlende Verweise})}$$

Die Verfolgbarkeit bietet die Basis für die Traceability: Sie gibt an, inwieweit zu einer Anforderung zugehörige Informationen gefunden werden können (z.B. Begründungen, Testfälle, etc.). In der Praxis kommen bei der Messung der Traceability in der Regel Requirement-Tools zum Einsatz.

$$\circ \quad \text{Sortierbarkeit} = \frac{\sum (\text{IST-Verwaltungsattribute})}{\sum (\text{SOLL-Verwaltungsattribute})}$$

Die Sortierbarkeit misst, wie gut der Leser eines Anforderungsdokuments die für ihn relevanten Informationen erreichen kann. Hierzu wird eine Liste mit den geforderten Attributen erstellt (z.B. Anforderungen, Definition, Abnahmekriterien, etc.) und auf das Anforderungsdokument angewendet.

2.5 Arten von Anforderungen

Anforderungen lassen sich in zwei grundlegende Kategorien unterteilen: Die funktionalen und nicht-funktionalen Anforderungen.

2.5.1 Funktionale Anforderungen

Funktionale Anforderungen sind jene Anforderungen, die „vom System selbstständig ausgeführt werden sollen, Interaktionen des Systems (Eingaben, Ausgaben) mit menschlichen Nutzern und Anforderungen zu allgemeinen, funktionalen Vereinbarungen und Einschränkungen“ [Rup06]. Anders gesagt: Funktionale Anforderungen sind jene Anforderungen, die konkret das Systemverhalten betreffen.

2.5.2 Nicht-funktionale Anforderungen

Nicht-funktionale Anforderungen sind jene Anforderungen, die nicht funktionale Anforderungen sind. Konkret: Nach [Sch-AuE07] ist eine nicht-funktionale Anforderung „eine Qualitätseigenschaft, die nicht direkt die korrekte Funktion eines Systems bestimmt“.

Gerade nicht-funktionale Anforderungen werden häufig in der Anforderungsanalyse nicht beachtet, können jedoch kritisch sein: Wenn etwa ein Vorgang nicht in einer akzeptablen Zeit mit der Software realisierbar ist, so kann es sein, dass die gesamte Software für den praktischen Einsatz de facto unbrauchbar ist – selbst wenn sie alle funktionalen Anforderungen erfüllt.

Untersucht werden folgende nicht-funktionale Anforderungen [Rup06]:

- Technische Anforderungen
- Anforderungen an die Benutzerschnittstelle
- Qualitätsanforderungen

3 Erstellung eines Messplans

3.1 Vorgehensweise

Die in Kapitel 2.3 genannten Qualitätsaspekte werden zusammengefasst und in „Sprachliche Qualität“ und „Inhaltliche Qualität“ unterteilt.

Die sprachliche Qualität der Anforderungen bezieht sich vor allem auf die Einhaltung von Formulierungsregeln, wie sie in [Rup06] beschrieben sind. Es wurden gezielt jene Formulierungsregeln ausgewählt, deren Einhaltung nur anhand des Dokuments geprüft werden kann und keinen Kontakt zu Stakeholder wie z.B. den Kunden erfordern. Ferner werden die Anforderungen auf „Verstehbarkeit“, „Eindeutigkeit“ und „Konsistenz (innerhalb einer Anforderung)“ untersucht – entsprechende Metriken werden im Verlauf dieses Kapitels mittels der GQM-Methode erstellt.

- Sprachliche Qualität
 - Einhaltung von Formulierungsregeln [Rup06]:
 - Vermeidung von unvollständig spezifizierten Prozesswörtern:
 - Anforderung ist im Aktiv formuliert
(z.B. „Der **Kunde gibt** seine Login-Daten am Terminal **ein**“)
 - Prozesse sind durch Vollverben ausgedrückt
(falsch ist z.B. „Das System soll dem Gast die Möglichkeit zur **Anmeldung bieten**“)
 - Existieren unvollständig spezifizierte Prozesswörter?
(z.B. „Wichtige Informationen sollen **hervorgehoben** werden können“ , wie wird hervorgehoben?)
 - Vermeidung von unvollständigen Vergleichen und Steigerungen
 - Existieren unvollständige Vergleiche?
(z.B. „**Alle anderen** Kunden müssen sich anmelden“, welchen Kunden gibt es sonst noch?)
 - Fehlen bei einem Constraint die näheren Beschreibungen?
(z.B. „**Unter keinen Umständen** darf Zugriff auf den geschützten Bereich erfolgen“, wie wird der Zugriff verhindert?)
 - Einsatz von Universalquantoren?
(z.B. „**Alle** Benutzer dürfen auf die Hauptseite zugreifen“, wirklich „alle“?)
 - Substantive ohne Bezugsindex?
(z.B. „**Diese Funktion** ist wichtig, weil...“, welche „Funktion“ ist gemeint?)
 - Substantive im Plural geschrieben? (Dort wo Singular nicht ausdrücklich falsch wäre) (z.B. „**Die Benutzer** melden sich an“, „der Benutzer“ wäre auch korrekt)
 - Nominalstil?
(z.B. „Auf **Aufforderung** wird eine **Dateneingabe** durchgeführt“)
 - Verstehbar
 - Eindeutig
 - Konsistent, d.h. keine Widersprüche innerhalb **einer** Anforderung

Die inhaltliche Qualität bezieht sich auf inhaltliche Qualitätsaspekte einer Anforderung. Mittels GQM werden Metriken entwickelt, um Aussagen zu der inhaltlichen Qualität einer Anforderung treffen zu können.

- Inhaltliche Qualität:
 - Vollständig
 - Korrekt
 - Konsistent (d.h. keine Widersprüche zwischen verschiedenen Anforderungen)
 - Prüfbar
 - Notwendig
 - Gültig und aktuell
 - Realisierbar
 - Bewertet
 - (Verfolgbar)
 - Klassifizierbar

Der Qualitätsaspekt „Verfolgbarkeit“ wird nicht untersucht werden, da die Verfolgbarkeit von Anforderungen im Verlauf eines Projekts nur mit dem Einsatz von Requirement-Tools effektiv möglich ist [Sch-AuE07]. Diese Tools wurden jedoch für die zu analysierenden Projekte nicht verwendet.

3.2 Ziele (nach GQM)

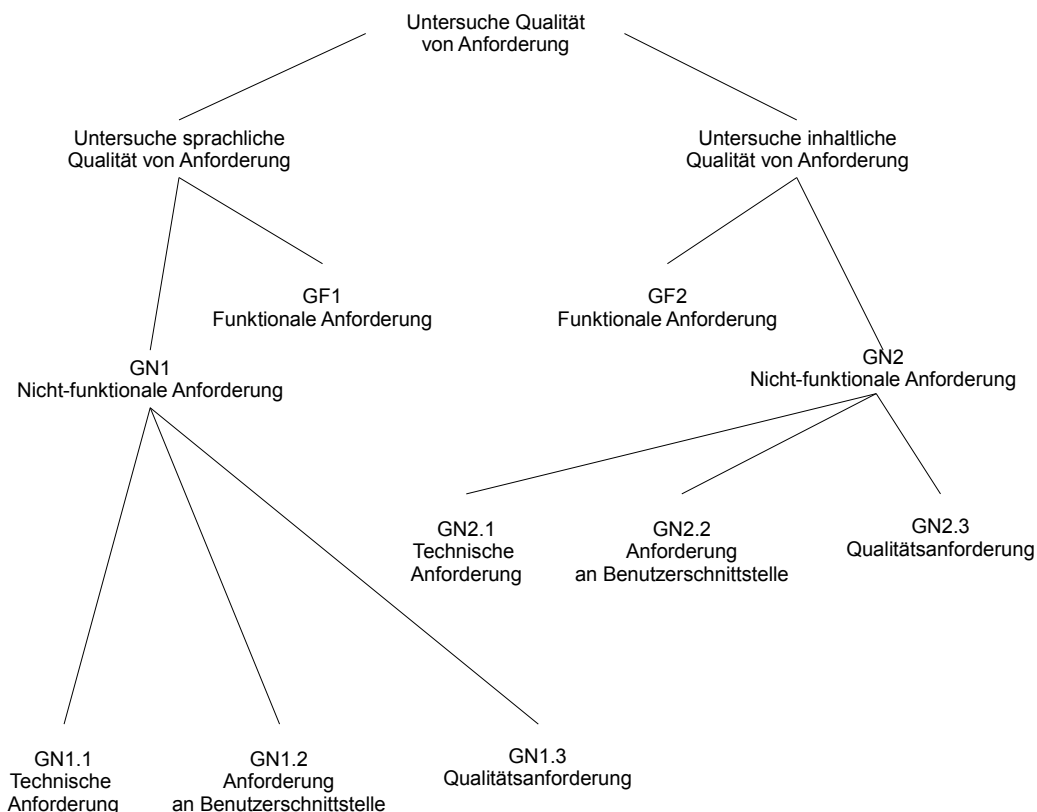


Abbildung 3.1: Zielbaum der Qualitätsziele

Das Hauptziel besteht darin, die Qualität von Anforderungen im Rahmen der Qualitätssicherung zu untersuchen. Die Teilziele, die sich daraus ergeben, bilden die unteren Ebenen des Zielbaums: Es werden zunächst die funktionalen Anforderungen auf ihre sprachliche (GF1) und inhaltliche Qualität (GF2) untersucht. Die Untersuchung der sprachlichen und inhaltlichen Qualität wird auch für nicht-funktionale Anforderungen durchgeführt, wobei sich die Ziele GN1 und GN2 jeweils in drei Unterziele aufteilen, die in der Untersuchung der technischen Anforderungen, der Anforderungen an die Benutzerschnittstelle und der Untersuchung der Qualitätsanforderung bestehen.

Die Ziele werden aus der Perspektive eines Qualitätsbeauftragten (im folgenden mit „QA“ abgekürzt) anhand der Anforderungsdokumente der Projekte untersucht.

Ein wesentlicher Schritt bei der GQM-Methode nach [Sol+99] ist auch, die aufgestellten Hypothesen mit den Messergebnissen abzugleichen.

Die Hypothesen sind als *erwartete* Antworten auf die formulierten Fragen aufzufassen [Sol+99]. Ohne den Abgleich der Hypothesen mit den Messergebnissen fehlt also ein zentraler Schritt bei der Durchführung der GQM-Methode, aber genau dieser Abgleich könnte bei den meisten zu untersuchenden Anforderungsdokumenten nicht mehr stattfinden, würde man z.B. auch aus Perspektive des Kunden die Ziele untersuchen: Die Kunden sind bei den meisten Projekten nicht mehr auffindbar oder aber die Projekte liegen soweit zurück, dass eine Einbindung der Kunden keine aussagekräftigen Ergebnisse mehr liefern würde. Aus diesem Grund wird als Perspektive die des QA verwendet.

Es ergeben sich folgende Ziele:

Ziel	Zweck	Qualitätsaspekt	Betrachtungsgegenstand	Perspektive
GF1	Untersuche	Sprachliche Qualität	Funktionale Anforderung	QA
GF2	Untersuche	Inhaltliche Qualität	Funktionale Anforderung	QA

Tabelle 3.2.1: Zielfacetten für funktionale Anforderungen

Ziel	Zweck	Qualitätsaspekt	Betrachtungsgegenstand	Perspektive
GN1.1	Untersuche	Sprachliche Qualität	Technische Anforderung	QA
GN1.2	Untersuche	Sprachliche Qualität	Anforderung an die Benutzerschnittstelle	QA
GN1.3	Untersuche	Sprachliche Qualität	Qualitätsanforderung	QA
GN2.1	Untersuche	Inhaltliche Qualität	Technische Anforderung	QA
GN2.2	Untersuche	Inhaltliche Qualität	Anforderung an die Benutzerschnittstelle	QA
GN2.3	Untersuche	Inhaltliche Qualität	Qualitätsanforderung	QA

Tabelle 3.2.2: Zielfacetten für nicht-funktionale Anforderungen

Für die einzelnen Ziele werden im Folgenden so genannte *Abstraction Sheets* erstellt. Abstraction Sheets sind als ein (tabellarisches) Hilfsmittel anzusehen, deren strukturiertes Ausfüllen hilft, sich über Probleme und Faktoren, die das Ziel betreffen, Gedanken zu machen [Sol+99].

Beim Ausfüllen der Abstraktion Sheet geht man – wie in der Literatur beschrieben [Sch07], [Sol+99] - in U-Form vor, d.h. man fängt links oben an, indem man zunächst Qualitätsfaktoren definiert, die die Qualität des Qualitätsaspekts ausmachen. Als nächstes stellt man Ausgangshypothesen auf, wie weit der jeweilige Qualitätsfaktor zum jetzigen Zeitpunkt erfüllt ist. Der nächste Schritt besteht im Aufstellen von Einfluss hypothesen: Wodurch werden die Qualitätsfaktoren auf welche Weise beeinflusst? Der letzte Schritt besteht in dem Aufstellen von Einflussfaktoren, die die Parameter darstellen, anhand derer die Qualitätsfaktoren beeinflusst werden können – oft können die Einflussfaktoren direkt aus den Einfluss hypothesen extrahiert werden.

Aus einem ausgefüllten Abstraktion Sheet lassen sich nun Fragen ableiten, deren Antworten angeben, inwieweit ein Ziel erfüllt ist [Sol+99]. Der letzte Schritt besteht nun in der Aufstellung von Metriken, deren Messergebnisse die gestellten Fragen (in quantifizierbarer Art und Weise) beantworten sollen [Sol+99].

Abstraction Sheet 1 (Ziel: GF1)			
Zweck: Untersuchung	Qualitätsaspekt: Sprachliche Qualität	Betrachtungsgegenstand: Funktionale Anforderung	Perspektive: QA
Qualitätsfaktoren 1 Grad der Einhaltung von Formulierungsregeln 2 Verstehbarkeit 3 Eindeutigkeit 4 Konsistenz 5 Korrekter Satzbau / Grammatik 6 Rechtschreibung		Einflussfaktoren: <ul style="list-style-type: none"> ● Anzahl der bekannten Formulierungsregeln (1) ● Anzahl der erklärten Fachbegriffen im Glossar (2,3) ● Granulat der Erklärung komplexer Zusammenhänge (2,3) ● Anzahl der Fehler in Grammatik, Rechtschreibung und Satzbau (2,3) ● Quantifizierbarkeit von Begriffsbedeutungen (3) ● Klarstellung der zeitlichen Zusammenhänge (3) ● Länge der geschriebenen Sätze (3,4) ● Kenntnisse der im Anforderungsdokument verwendeten Sprache (5,6) ● Zeitdruck (2,6) 	
Ausgangshypothesen 1 : 1.1 Beschreibung in natürlicher Sprache ohne Berücksichtigung von Formulierungsregeln 2 : 2.1 Fachliche Ausdrücke werden verwendet, aber nicht erläutert 2.2 Komplexe Vorgänge werden zu ungenau beschrieben 2.3 Zusammenhänge werden nur unzureichend erläutert 2.4 Zusammenhänge sind teilweise unverständlich 3 : 3.1 Ungenaue und vage Beschreibung lässt Raum für individuelle Interpretationen 3.2 Mehrdeutige Begriffe werden verwendet 4 : 4.1 Zusammenhänge innerhalb eines Satzes widersprechen sich 5 : 5.1 Teilweise grammatikalische Fehler 6 : 6.1 Rechtschreibfehler tauchen auf		Einflussshypothesen: Abhängigkeiten <ul style="list-style-type: none"> ● Kenntnis an Formulierungsregeln beeinflusst 1.1 ● Hohe fachliche Kompetenz der Autoren führt zu 2.1 ● Mangel an Zeit führt zu 2.2, 2.3 und zu 6.1 ● Fehler in Rechtschreibung, Satzbau und Grammatik führen zu 2.4 und zu 4.1 ● Unklare Zielvorstellungen führen zu 3.1 ● Verknüpfte Anforderungen mit „oder“ führen zu 3.1 ● Unvollständige Erklärung von Begriffen führt zu 3.2 ● Formulierung von langen Sätzen führt zu 4.1 ● Ausnahmen von der Regel im selben Satz führen zu 3.1 und 4.1 	

Abstraction Sheet 2 (Ziel: GF2)			
Zweck: Untersuchung	Qualitätsaspekt: Inhaltliche Qualität	Betrachtungsgegenstand: Funktionale Anforderung	Perspektive: QA
Qualitätsfaktoren <ol style="list-style-type: none"> 1 Vollständigkeit 2 Korrektheit 3 Konsistent 4 Prüfbar 5 Notwendig 6 Gültig und aktuell 7 Realisierbarkeit 8 Bewertbarkeit 9 Klassifizierbarkeit 		Einflussfaktoren: <ul style="list-style-type: none"> ● Art der Ermittlungstechnik (Workshop, Einzelinterviews,...) (1,2) ● Erwähnung von „Selbstverständlichkeiten“ (1,2) ● Komplexität der Geschäftsprozesse (1,2) ● Anzahl der Fachwörter (des Kunden) bei Kommunikation mit Entwicklern (2) ● Anzahl der redundanten Anforderungen (3) ● Anzahl der Veränderungen der Geschäftsprozesse (3,6) ● Klarheit von Zielen einer Anforderung (4) ● Anzahl der „Anforderungen“, die von Entwicklern aufgestellt wurden (5) ● Häufigkeit der Kommunikation (zwischen Kunden und Entwicklern) (6) ● Einschätzbarkeit der Komplexität der Umsetzung einer Anforderung (7) ● Kenntnisse über die technischen Grenzen (7) ● Kenntnis über die Wichtigkeit einer Anforderung (8) ● Klarheit der rechtlichen Verbindlichkeiten (9) 	
Ausgangshypothesen <ol style="list-style-type: none"> 1 : <ol style="list-style-type: none"> 1.1 Anforderung beschreibt nur teilweise die Bedürfnisse des Kunden 2 : <ol style="list-style-type: none"> 2.1 Anforderung beschreibt die vom Kunden gewünschte Funktionalität nicht korrekt 3 : <ol style="list-style-type: none"> 3.1 Inhalt der Anforderung steht im Widerspruch zu einer Anforderung oder zu mehreren anderen Anforderungen 4 : <ol style="list-style-type: none"> 4.1 Vage Formulierung der Anforderung ohne nachprüfbar Kriterien zur Erfüllung 4.2 Kriterien zur Erfüllung sind (z.B. durch schwer simulierbare äußere Bedingungen) nicht nachprüfbar 5 : <ol style="list-style-type: none"> 5.1 Entwickler erweitern Anforderung, ohne dies mit dem Kunden abzustimmen 5.2 Entwickler lassen ihre eigenen Wünsche in die Anforderung einfließen 6 : <ol style="list-style-type: none"> 6.1 Anforderung in der Realität verändert sich schnell, aber Dokument wird nicht angepasst 7 : <ol style="list-style-type: none"> 7.1 Anforderungen sind zeitlich oder technisch nicht realisierbar 8 : <ol style="list-style-type: none"> 8.1 Keine explizite Bewertung für die meisten Anforderungen 9 : <ol style="list-style-type: none"> 9.1 Keine Erwähnung der rechtlichen Verbindlichkeiten 		Einflusshypthesen: Abhängigkeiten <ul style="list-style-type: none"> ● Ungenaue Vorstellungen (des Kunden) führen zu 1.1 ● Unerwähnte „Selbstverständlichkeiten“ führen zu 1.1 und zu 2.1 ● Ungeeignete Ermittlungstechnik der Anforderung seitens der Entwickler führt zu 1.1 und zu 2.1 ● Komplexität der Geschäftsprozesse des Kunden führt zu 1.1 und 2.1 ● Missverständnisse (zwischen Kunde und Entwickler) führen zu 2.1 ● Redundanz innerhalb des Anforderungsdokuments begünstigt 3.1 ● Unterschiedliche Vorstellungen bzgl. der Anforderung führt zu 3.1 ● Unklare Zielvorstellungen führen zu 4.1 ● Falsch verstandene Kreativität der Entwickler führt zu 5.2 ● Mangel an Kommunikation zwischen Kunden und Entwicklern führt zu 5.1 und zu 6.1 ● Veränderte Geschäftsprozesse führen zu 3.1 und 6.1 ● Unkenntnis der technischen Grenzen führt zu 7.1 ● Falsche Einschätzung der Komplexität der Umsetzung führt zu 7.1 ● Unklare Vorstellung über die Wichtigkeit einer Anforderung begünstigt 8.1 ● Unklarheiten über Vertragsgestaltung auf Entwicklerseite führen zu 9.1 	

In den Abstraction-Sheets für die nicht-funktionalen Anforderungen wurden bewusst nur jene Aspekte aufgenommen, die sich von den funktionalen Anforderungen unterscheiden. Aspekte, die mit jenen der funktionalen Anforderungen identisch sind, sind mit „...“ gekennzeichnet und können in den Abstraction Sheet zu GF2 unter dem jeweiligen Aufzählungspunkt nachgeschlagen werden.

Die Abstraction Sheets zu den Qualitätszielen GN1.1, GN1.2 und GN1.3 sind mit Abstraction Sheet von GF1 identisch.

Abstraction Sheet 3 (Ziel: GN2.1)			
Zweck: Untersuchung	Qualitätsaspekt: Inhaltliche Qualität	Betrachtungsgegenstand: Technische Anforderung	Perspektive: QA
Qualitätsfaktoren 1 Vollständigkeit 2 Korrektheit 3 Konsistent 4 Prüfbar 5 Notwendig 6 Gültig und aktuell		Einflussfaktoren: <ul style="list-style-type: none"> ● Anzahl der erwähnten Kategorien von technischen Anforderungen ● Erwähnung der Kategorie einer technischen Anforderung 	
Ausgangshypothesen 1 : 1.1 ... 1.2 Beschreibung von Anforderungen ohne Nennung der Kategorie (z.B. „Es soll JavaScript verwendet werden“ ohne Benennung der Kategorie „Anforderung an Programmiersprache“) 1.3 Nur einzelne Kategorien überhaupt erwähnt (z.B. Programmiersprache) 2 : 2.1 ... 3 : 3.1 ... 4 : 4.1 ... 4.2 ... 5 : 5.1 ... 5.2 ... 6 : 6.1 ...		Einflussshypothesen: Abhängigkeiten <ul style="list-style-type: none"> ● Keine Vorgaben zu den Kategorien von nicht-funktionalen Anforderungen begünstigen 1.2 ● Kenntnisse der Kategorien beeinflusst 1.3 	

Abstraction Sheet 4 (Ziel: GN2.2)			
Zweck: Untersuchung	Qualitätsaspekt: Inhaltliche Qualität	Betrachtungsgegenstand: Anforderung an die Benutzerschnittstelle	Perspektive: QA
Qualitätsfaktoren 1 Vollständigkeit 2 Korrektheit 3 Konsistent 4 Prüfbar 5 Notwendig 6 Gültig und aktuell		Einflussfaktoren: <ul style="list-style-type: none"> ● Vorhandensein einer visuellen Darstellung der Benutzerschnittstelle (1) ● Vorhandensein von Bedienelementen in visueller Darstellung (1) ● Benennung von Benutzergruppen und ihren Fähigkeiten (1) ● Benennung von Eingabegeräten (1) 	
Ausgangshypothesen 1 : 1.1 ... 1.2 Es existiert keine visuelle Beschreibung der Benutzerschnittstelle 1.3 Die Bedürfnisse unterschiedlicher Benutzer-Gruppen wurden nur unzureichend berücksichtigt 1.4 Physikalische Aspekte (z.B. Bildschirmgröße) werden nicht genannt 1.5 Eingabegeräte (Maus, Tastatur, ...) werden nicht genannt 2 : 2.1 ... 3 : 3.1 ... 4 : 4.1 ... 4.2 ... 4.3 Keine Festlegung von quantifizierbaren Kriterien, die Grad der Bedienbarkeit messbar machen 5 : 5.1 ... 5.2 ... 6 : 6.1 ...		Einflusshypthesen: Abhängigkeiten <ul style="list-style-type: none"> ● Ungenaue Vorstellungen über Aussehen der Benutzerschnittstelle führen zu 1.2, 1.5 ● Unkenntnis der technischen Rahmenbedingungen führt zu 1.4 ● Ungenaue Vorstellungen über Benutzergruppen führen zu 1.3 	

Zu GN2.3 (Untersuchung der inhaltlichen Qualität einer Qualitätsanforderung) wird kein eigenes Abstraction Sheet erstellt. Der wesentliche Aspekt bei GN2.3 ist das Vorhandensein und der Inhalt eines so genannten *Qualitätsmodells*.

Ein Qualitätsmodell gliedert sich in drei Ebenen [Sch07]:

1. **Abstrakte Qualitätsziele:** In der ersten Ebene werden die abstrakten Qualitätsziele festgelegt, indem man Begriffe wie z.B. „Sicherheit“, „Effizienz“, etc. aus einer genormten Liste (z.B. IEEE 610.12-1990) wählt.
2. **Konkrete Qualitätsaspekte:** Hier werden die abstrakten Ziele der ersten Ebene auf das aktuelle Projekt angewendet. Es wird konkretisiert, was z.B. mit „Sicherheit“ gemeint ist – Sicherheit gegen Datenverlust? Sicherheit gegen Manipulation der Daten? usw.
3. **Indikatoren und Metriken:** Zu guter Letzt werden Metriken und Indikatoren erstellt, um zu messen inwieweit ein Qualitätsaspekt erfüllt ist. Dies könnte beim Qualitätsaspekt „Sicherheit“ z.B. bedeuten, dass ein simulierter Ausfall des Systems zu maximal 5% Datenverlusten führt.

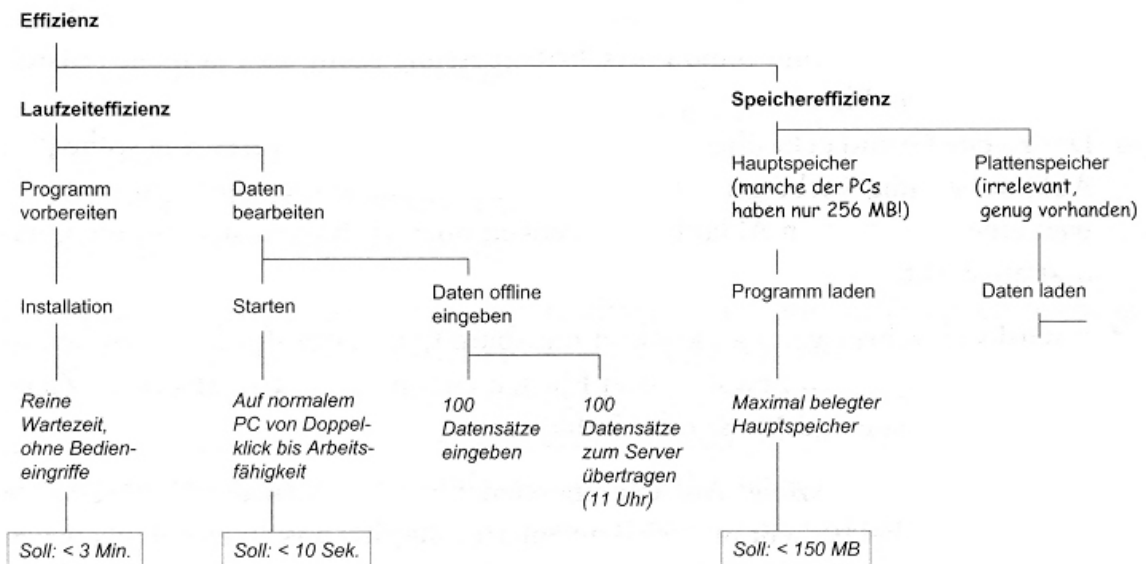


Abbildung 3.2.: Ausschnitt aus einem Qualitätsmodell nach [Sch07]

3.3 Fragen (nach GQM)

Aus den in Kapitel 3.2 erstellten Abstraction Sheets werden nun zu jedem Qualitätsaspekt eine oder mehrere Fragen abgeleitet:

GF1-Q1.1	Wie viele Formulierungsregeln pro Anforderung wurden eingehalten?
GF1-Q2.1	Wie viele Fachbegriffe werden im Glossar erklärt?
GF1-Q2.2	Wie viele Fehler, die die Verstehbarkeit der Anforderung beeinflussen, enthält die Anforderung?
GF1-Q3.1	Existieren pro Satz mehrere verbundene Anforderungen?
GF1-Q3.2	Wie viele mehrdeutige Begriffe enthält die Anforderung?
GF1-Q3.3	Wie viele Fehler, die zu Mehrdeutigkeit führen, enthält die Anforderung?
GF1-Q4.1	Wie viele Aspekte einer Anforderung stehen im Widerspruch zueinander?
GF1-Q5.1	Wie viele Anforderungen sind grammatikalisch korrekt formuliert?
GF1-Q6.1	Wie viele Anforderungen sind frei von Rechtschreibfehlern?

Tabelle 3.3.1: Fragen zu GF1 (Sprachliche Qualität einer funktionalen Anforderung)

GF2-Q1.1	Wie viele Aspekte der Anforderung sind dokumentiert worden?
GF2-Q2.1	Wie viele inhaltliche Defekte hat die Anforderung?
GF2-Q3.1	Wie viele Anforderungen tauchen mehrfach im Dokument auf?
GF2-Q3.2	Wie viele Anforderungen widersprechen sich gegenseitig?
GF2-Q4.1	Wie viele Aspekte der Anforderung sind messbar?
GF2-Q5.1	Wie viele Anforderungen im Dokument sind vom Kunden nicht explizit genannt worden?
GF2-Q.6.1	Wie viele Anforderungen aus Geschäftsprozessen sind nicht mehr aktuell?
GF2-Q7.1	Welche Aspekte einer Anforderung können (aus technischen, finanziellen, zeitlichen oder rechtlichen Aspekten) nicht umgesetzt werden?
GF2-Q.8.1	Wurde die Wichtigkeit einer Anforderung eingeschätzt?
GF2-Q.9.1	Wurde die rechtliche Verbindlichkeit einer Anforderung eingeschätzt?

Tabelle 3.3.2: Fragen zu GF2 (Inhaltliche Qualität einer funktionalen Anforderung)

GN2.1-Q1.1	Wurde zu einer technischen Anforderung die Kategorie genannt?
GN2.1-Q1.2	Wie viele Kategorien von technischen Anforderungen tauchen im Dokument auf?

Tabelle 3.3.3: Fragen zu GN2.1 (Inhaltliche Qualität einer technischen Anforderung)

GN2.2-Q1.1	Existiert eine visuelle Beschreibung der Benutzerschnittstelle?
GN2.2-Q1.2	Wird präzisiert, welche Benutzer das Programm bedienen sollen?
GN2.2-Q1.3	Werden bei Anforderung mit Benutzerinteraktion die Eingabegeräte benannt?

***Tabelle 3.3.4: Fragen zu GN2.2
(Inhaltliche Qualität einer Anforderung an die Benutzerschnittstelle)***

GN2.3-Q1.1	Existiert ein Qualitätsmodell?
GN2.3-Q1.2	Werden konkrete Qualitätsziele genannt?
GN2.3-Q1.3	Wird angegeben, wann die Qualitätsziele erfüllt sind?

***Tabelle 3.3.5: Fragen zu GN2.3
(Inhaltliche Qualität einer Qualitätsanforderung)***

3.4 Metriken und Indikatoren (nach GQM)

Nachdem zu jedem Qualitätsaspekt (mindestens) eine Frage formuliert worden ist, werden in diesem Kapitel nun konkrete Metriken zur Messung erstellt, deren Messergebnisse die Antworten auf die jeweiligen Fragen bezüglich der Qualitätsaspekte darstellen [Sol+99].

Genauere Erläuterungen zu den einzelnen Metriken finden sich in Kapitel 3.5.2.

Tabelle 3.4.1 zeigt die Metriken zum Ziel GF1 (sprachliche Qualität einer funktionalen Anforderung). Diese Metriken gelten gleichermaßen für alle nicht-funktionalen Anforderungen.

Metrik	Formel
GF1-M1.1: Formulierungsregelkonformität einer Anforderung	$\frac{\sum \text{Eingehaltenen Formulierungsregeln}}{\sum \text{Formulierungsregeln}}$
GF1-M2.1: Erläuterung von Fachbegriffen	$\frac{\sum \text{Erklärte Fachbegriffe}}{\sum \text{Fachbegriffe im Dokument}}$
GF1-M2.2: Anzahl der sprachlichen Fehler, die die Verstehbarkeit beeinflussen	Zähle die sprachlichen Fehler, die die Verstehbarkeit der Anforderung beeinflussen.
GF1-M3.1: Identifizierbarkeit von Anforderungen	Suche pro Satz nach mehr als einer Anforderung.
GF1-M3.2: Anzahl mehrdeutiger Begriffe innerhalb einer Anforderung	Zähle die mehrdeutigen Begriffe innerhalb der Anforderung.
GF1-M3.3: Mehrdeutigkeit durch sprachliche Fehler	Zähle die sprachlichen Fehler, die zu Mehrdeutigkeit führen.
GF1-M4.1: Widersprüchlichkeit einer Anforderung	$\frac{\sum \text{Widersprüchliche Aspekte der Anforderung}}{\sum \text{Aspekte einer Anforderung}}$
GF1-M5.1: Grammatikalische Korrektheit	$\frac{\sum \text{Grammatikalisch korrekte Anforderungen}}{\sum \text{Anforderungen}}$
GF1-M6.1: Korrektheit in Bezug auf die Rechtschreibung	$\frac{\sum \text{Anforderungen ohne Rechtschreibfehler}}{\sum \text{Anforderungen}}$

Tabelle 3.4.1: Metriken zu GF1 (sprachliche Qualität einer funktionalen Anforderung)

Tabelle 3.4.2 zeigt die Metriken zum Ziel GF2 (inhaltliche Qualität einer funktionalen Anforderung). Diese Metriken gelten gleichermaßen für alle nicht-funktionalen Anforderungen, wobei für einzelne nicht-funktionale Anforderungen noch spezielle Metriken hinzukommen.

Metrik	Formel
GF2-M1.1: Vollständigkeit einer Anforderung	$\frac{\sum \text{Dokumentierte Aspekte einer Anforderung}}{\sum \text{Aspekte einer Anforderung}}$
GF2-M2.1: Korrektheit einer Anforderung	$1 - \frac{\sum \text{Fehlerhafte Aspekte einer Anforderung}}{\sum \text{Aspekte einer Anforderung}}$
GF2-M3.1: Redundanz	$\frac{\sum \text{Redundante Anforderungen}}{\sum \text{Anforderungen}}$
GF2-M3.2: (Externe) Widersprüchlichkeit einer Anforderung	$\frac{\sum \text{Widersprüchliche Anforderungen}}{\sum \text{Anforderungen}}$
GF2-M4.1: Prüfbarkeit einer Anforderung	$\frac{\sum \text{Prüfbare Aspekte der Anforderung}}{\sum \text{Aspekte der Anforderung}}$
GF2-M5.1: Notwendigkeit	$\frac{\sum \text{Anforderungen, die nicht vom Kunden genannt wurden}}{\sum \text{Dokumentierte Anforderungen}}$
GF2-M6.1: Obsoleszenz	$\frac{\sum \text{Geänderte Geschäftsprozesse seit letzter Kommunikation mit Entwicklern}}{\sum \text{Dokumentierte Anforderungen aus Geschäftsprozessen}}$
GF2-M7.1: Realisierbarkeit	Keine Metrik vorhanden
GF2-M8.1: Bewertbarkeit	Suche nach Einschätzung der Wichtigkeit einer Anforderung.
GF2-M9.1: Klassifizierbarkeit	Suche nach Benennung der rechtlichen Verbindlichkeit einer Anforderung.

Tabelle 3.4.2: Metriken zu GF2 (inhaltliche Qualität einer funktionalen Anforderung)

Tabelle 3.4.3 zeigt die Metriken, die sich speziell auf technische Anforderungen beziehen; es werden für technische Anforderungen sowohl die Metriken aus Tabelle 3.4.2 wie auch aus Tabelle 3.4.3 angewendet.

Metrik	Formel
GN2.1-M1.2: Zuordnung zu einer Kategorie	Suche nach Nennung der Kategorie einer technischen Anforderung.
GN2.1-M1.3: Anzahl Kategorien von technischen Anforderungen im Dokument	Zähle die Kategorien von technischen Anforderungen im Dokument.

Tabelle 3.4.3: Metriken zu GN2.1 (inhaltliche Qualität einer technischen Anforderung)

Tabelle 3.4.4 zeigt die Metriken, die sich speziell auf Anforderungen an die Benutzerschnittstelle beziehen; es werden für Anforderungen an die Benutzerschnittstelle sowohl die Metriken aus Tabelle 3.4.2 wie auch aus Tabelle 3.4.4 angewendet.

Metrik	Formel
GN2.2-M1.1: Vorhandensein einer visuellen Beschreibung der Benutzerschnittstelle	Suche nach visueller Beschreibung der Benutzerschnittstelle.
GN2.2-M1.2: Vorhandensein eines Benutzerprofils	Suche nach Beschreibung der zukünftigen Benutzer und deren Fähigkeiten.
GN2.2-M1.3: Benennung von Eingabegeräten	Suche in Anforderungen mit Benutzerinteraktion nach Nennung der Eingabegeräte.

*Tabelle 3.4.4: Metriken zu GN2.2
(inhaltliche Qualität einer Anforderung an die Benutzerschnittstelle)*

Tabelle 3.4.5 zeigt die Metriken, die sich auf Qualitätsanforderungen beziehen; es werden für Qualitätsanforderungen sowohl die Metriken aus Tabelle 3.4.2 wie auch aus Tabelle 3.4.5 verwendet.

Metrik	Formel
GN2.3-M1.1: Vorhandensein eines Qualitätsmodells	Suche nach einem Qualitätsmodell.
GN2.3-M1.2: Nennung von Qualitätszielen	Suche zu einem allgemeinen Qualitätsziel nach konkreten (!) Qualitätszielen.
GN2.3-M1.3: Metriken für Qualitätsziele	Suche zu einem konkreten Qualitätsziel nach prüfbareren Kriterien für die Erfüllung.

Tabelle 3.4.5: Metriken zu GN2.3 (inhaltliche Qualität einer Qualitätsanforderung)

3.5 Aufstellen des Messplans

3.5.1 Auswahl der Metriken

Für die Aufstellung des Messplans wird nun eine Auswahl an Metriken getroffen, die sich für die (isolierte) Analyse eines Anforderungsdokuments eignen. Einige der aufgestellten Metriken erfordern den Kontakt zu Stakeholdern (z.B. Kunden) und eignen sich daher nicht zur isolierten Analyse eines Anforderungsdokuments.

Folgende Metriken werden nicht in den Messplan einfließen:

GF2-M1.1: Vollständigkeit einer Anforderung	
Zur Frage:	GF2-Q1.1 Wie viele Aspekte der Anforderung sind dokumentiert worden?
Metrik:	$\frac{\sum \text{Dokumentierte Aspekte einer Anforderung}}{\sum \text{Aspekte einer Anforderung}}$
Einheit:	Prozentual [%]
Skala:	0% – 100%

Bei der Metrik GF2-M1.1 taucht das Problem auf, dass eine Referenz bezüglich der Vollständigkeit fehlt. Die dokumentierten Aspekte einer Anforderung sind jene, die aus dem Anforderungsdokument zu entnehmen sind. Was jedoch schwierig zu ermitteln ist, sind die gesamten Aspekte einer Anforderung: Hierzu müssten entweder Stakeholder befragt werden oder Protokolle aus Kundeninterviews mit dem Anforderungsdokument verglichen werden. Und selbst der Kontakt zum Kunden würde eventuell nicht alle Aspekte einer Anforderung zu Tage fördern, da der Kunde vermutlich selber nicht alle tatsächlich benötigten Anforderungsaspekte überblicken und nennen kann. Um die Vollständigkeit einer Anforderung zu prüfen, sind also Verfahren wie z.B. Inspektion, Walkthrough, etc. wie sie in [Rup06] erwähnt sind, anzuwenden.

GF2-M2.1: Korrektheit einer Anforderung	
Zur Frage:	GF2-Q2.1 Wie viele inhaltliche Defekte hat die Anforderung?
Metrik:	$1 - \frac{\sum \text{Fehlerhafte Aspekte einer Anforderung}}{\sum \text{Aspekte einer Anforderung}}$
Einheit:	Prozentual [%]
Skala:	0% – 100%

Bei GF2-M2.1 fehlt ebenfalls eine Referenz bezüglich der Korrektheit: Auch hier müsste ggf. das Anforderungsdokument gegen andere Dokumente oder Ton- und Bildaufzeichnungen aus der Anforderungsphase geprüft werden, bzw. auch der Kunde direkt befragt werden. Auch bezüglich der Korrektheit einer Anforderung sind also Verfahren notwendig, die auch Stakeholder von Kundenseite einschließen.

GF2-M5.1: Notwendigkeit	
Zur Frage:	GF2-Q5.1 Wie viele Anforderungen im Dokument sind vom Kunden nicht explizit genannt worden?
Metrik:	$\frac{\sum \text{Anforderungen, die nicht vom Kunden genannt wurden}}{\sum \text{Dokumentierte Anforderungen}}$
Einheit:	Prozentual [%]
Skala:	0% – 100%

Die Anwendung von GF2-M5.1 setzt voraus, dass jede im Anforderungsdokument formulierte Anforderung mit einer konkreten Anforderungsquelle (z.B. Protokoll eines Kundengesprächs) verknüpft ist – diese Verknüpfungen existieren für die zu analysierenden Projekte nicht. Eine weitere Schwäche dieser Metrik ist, dass es sehr wohl Anforderungen geben kann, die faktisch notwendig sind, vom Kunden jedoch nicht explizit genannt wurden, da sie in seinen Augen „selbstverständlich“ seien (Stichwort: „*Symmetrie of Ignorance*“). Um die Notwendigkeit von Anforderungen zu bestimmen, sind also abermals kommunikative Verfahren notwendig.

GF2-M6.1: Obsoleszenz	
Zur Frage:	GF2-Q.6.1 Wie viele Anforderungen aus Geschäftsprozessen sind nicht mehr aktuell?
Metrik:	$\frac{\sum \text{Geänderte Geschäftsprozesse seit letzter Kommunikation mit Entwicklern}}{\sum \text{Dokumentierte Anforderungen aus Geschäftsprozessen}}$
Einheit:	Prozentual [%]
Skala:	0% – 100%

Bei GF2-M6.1 ist eine Kommunikation mit dem Kunden zwingend notwendig, da aus dem Anforderungsdokument allein nicht zu erfahren ist, inwieweit sich die Geschäftsprozesse des Kunden geändert haben.

Ferner werden die Metriken GF1-M5.1 und GF1-M6.1 nicht in den Messplan einfließen, da von einem professionellen Anforderungsdokument zwar erwartet werden darf, dass es frei von Rechtschreib- und Grammatikfehlern ist, aber diese beiden Aspekte nicht direkt die Qualität von Anforderungen beeinflussen – spezielle Rechtschreib- und Grammatikfehler, die sich auf die Qualität der Anforderungen auswirken, sind schon durch andere Metriken berücksichtigt.

Obwohl die Messung von bestimmten Qualitätsaspekten wie z.B. der „Vollständigkeit einer Anforderung“ anhand des Anforderungsdokuments allein nicht möglich zu sein scheint, können jedoch hier die Formulierungsregeln helfen, Indikatoren für *mögliche* Defekte zu finden, d.h. es handelt sich hierbei lediglich um Heuristiken: Nicht jeder Verstoß einer Formulierungsregel muss zwangsläufig (aber kann!) einen inhaltlichen Defekt zur Folge haben.

Die folgende Tabelle gibt eine Übersicht über die Metriken, die nicht in den Messplan mit eingeflossen sind, für die es jedoch Formulierungsregeln gibt, die heuristische Aussagen über mögliche Defekte ergeben. In Kapitel 4.3.3 wird dieser Zusammenhang noch näher erläutert werden.

Metrik-ID	Metrik	Formulierungsregeln
GF2-M1.1	Vollständigkeit einer Anforderung	Anforderung enthält kein Passiv, Prozesse sind durch Vollverben ausgedrückt, Anforderung enthält keine unvollständig spezifizierte Prozesswörter, Anforderung enthält keine unvollständigen Vergleiche, Constraints sind ausreichend beschrieben, Anforderung enthält keinen Nominalstil
GF2-M2.1	Korrektheit einer Anforderung	Anforderung enthält keine Universalquantoren
GF2-M5.1	Notwendigkeit	-
GF2-M6.1	Obsoleszenz	-

Tabelle 3.5.1: Metriken, die nicht im Messplan sind mit dazugehörigen Heuristiken

Durch das Wegfallen der Metriken aus Tabelle 3.5.1 aus dem Messplan, handelt es sich bei der „inhaltlichen Qualität“ faktisch nur noch um eine „formale Qualität“. Der Ausdruck „inhaltliche Qualität“ wird im folgenden synonym zu „formale Qualität“ verwendet.

3.5.2 Der Messplan

Im Folgenden wird nun der konkrete Messplan fertiggestellt, wobei für jede Metrik auch detailliert die Vorgehensweise der Messung beschrieben wird.

Metriken zur „sprachlichen Qualität“:

Die Metrik GF1-M1.1 misst wie viele Formulieringsregeln pro Anforderung eingehalten werden. Hierzu wird zunächst eine einzelne Anforderung identifiziert und dann überprüft welche Formulieringsregeln eingehalten bzw. nicht eingehalten worden sind.

GF1-M1.1: Formulieringsregelkonformität einer Anforderung	
Zur Frage:	GF1-Q1.1 Wie viele Formulieringsregeln pro Anforderung wurden eingehalten?
Metrik:	$\frac{\sum \text{Eingehaltenen Formulieringsregeln}}{\sum \text{Formulieringsregeln}}$
Einheit:	Prozentual [%]
Skala:	0% – 100%
Datenquelle:	Anforderungsdokument
Beschreibung:	<p>Identifiziere eine einzelne Anforderung im Anforderungsdokument. Prüfe, welche der folgenden Formulieringsregeln sie einhält:</p> <ol style="list-style-type: none"> 1. Anforderung ist im Aktiv formuliert 2. Prozesse innerhalb der Anforderung sind durch Verben ausgedrückt 3. Es gibt keine unvollständig spezifizierten Prozesswörter 4. Es gibt keine unvollständigen Vergleiche 5. Constraints sind ausreichend genau beschrieben 6. Es werden keine Universalquantoren eingesetzt 7. Keine Substantive ohne Bezugsindex? 8. Substantive sind im Singular geschrieben (wo Plural nicht explizit notwendig wäre) 9. Die Anforderung ist nicht im Nominalstil formuliert <p>Zähle die Formulieringsregeln, die von dieser Anforderung eingehalten werden und setze sie zu der Gesamtzahl der Formulieringsregeln ins Verhältnis.</p>

Die Metrik GF1-M2.1 misst, wie viele der Fachbegriffe im Glossar des Dokuments erklärt worden sind.

GF1-M2.1: Erläuterung von Fachbegriffen	
Zur Frage:	GF1-Q2.1 Wie viele Fachbegriffe werden im Glossar erklärt?
Metrik:	$\frac{\sum \text{Erklärte Fachbegriffe}}{\sum \text{Fachbegriffe im Dokument}}$
Einheit:	Prozentual [%]
Skala:	0% – 100%
Datenquelle:	Anforderungsdokument
Beschreibung:	Zähle die Fachbegriffe im Dokument und überprüfe wie viele davon im Glossar erklärt sind. Setze die Anzahl der Fachbegriffe im Glossar zu der Gesamtzahl der Fachbegriffe im Dokument ins Verhältnis.

GF1-M2.2 zählt die sprachlichen Fehler, die sich auf die Verstehbarkeit einer Anforderung auswirken.

GF1-M2.2: Anzahl der sprachlichen Fehler, die die Verstehbarkeit beeinflussen	
Zur Frage:	GF1-Q2.2 Wie viele Fehler, die die Verstehbarkeit der Anforderung beeinflussen, enthält die Anforderung?
Metrik:	Zähle die sprachlichen Fehler, die die Verstehbarkeit der Anforderung beeinflussen.
Einheit:	-
Skala:	Ordinalskala
Datenquelle:	Anforderungsdokument
Beschreibung:	Identifiziere eine einzelne Anforderung im Anforderungsdokument. Zähle die sprachlichen Fehler, die die Verständlichkeit der Anforderung beeinflussen.

Im Idealfall enthält ein Satz nur eine einzelne Anforderung. GF1-M3.1 überprüft, ob ein Satz mehr als eine Anforderung enthält.

GF1-M3.1: Identifizierbarkeit von Anforderungen	
Zur Frage:	GF1-Q3.1 Existieren pro Satz mehrere verbundene Anforderungen?
Metrik:	Suche pro Satz nach mehr als einer Anforderung.
Einheit:	-
Skala:	Ja / Nein
Datenquelle:	Anforderungsdokument
Beschreibung:	Prüfe für jeden Satz, ob er mehr als eine Anforderung enthält.

GF1-Q3.2 dient zur Ermittlung der Anzahl der mehrdeutigen Begriffe innerhalb einer Anforderung.

GF1-M3.2: Anzahl mehrdeutiger Begriffe innerhalb einer Anforderung	
Zur Frage:	GF1-Q3.2 Wie viele mehrdeutige Begriffe enthält die Anforderung?
Metrik:	Zähle die mehrdeutigen Begriffe innerhalb der Anforderung.
Einheit:	-
Skala:	Ordinalskala
Datenquelle:	Anforderungsdokument
Beschreibung:	Identifiziere eine einzelne Anforderung im Anforderungsdokument. Suche in der Anforderung nach Begriffen (Substantiven, die nicht als Fachwörter im Glossar auftauchen), die mehr als eine Bedeutung haben können (z.B. „Browser“ - welcher Browser genau? MS Internet Explorer, Mozilla Firefox,...?) und zähle diese.

GF1-M3.3 zählt jene sprachlichen Fehler, die dazu führen, dass die Anforderung auf mehr als eine Art und Weise verstanden werden kann.

GF1-M3.3: Mehrdeutigkeit durch sprachliche Fehler	
Zur Frage:	GF1-Q3.3 Wie viele Fehler, die zu Mehrdeutigkeit führen, enthält die Anforderung?
Metrik:	Zähle die sprachlichen Fehler, die zu Mehrdeutigkeit führen.
Einheit:	-
Skala:	Ordinalskala
Datenquelle:	Anforderungsdokument
Beschreibung:	Identifiziere eine einzelne Anforderung im Anforderungsdokument. Zähle die sprachlichen Fehler, die dazu führen, dass die Anforderung auf mehr als eine Art und Weise verstanden werden kann.

GF1-M4.1 ermittelt die Anzahl der sich widersprechenden Aspekte einer Anforderung und setzt sie zu der Gesamtanzahl der Aspekte ins Verhältnis.

GF1-M4.1: Widersprüchlichkeit einer Anforderung	
Zur Frage:	GF1-Q4.1 Wie viele Aspekte einer Anforderung stehen im Widerspruch zueinander?
Metrik:	$\frac{\sum \text{Widersprüchliche Aspekte der Anforderung}}{\sum \text{Aspekte einer Anforderung}}$
Einheit:	Prozentual [%]
Skala:	0% – 100%
Datenquelle:	Anforderungsdokument
Beschreibung:	Identifiziere eine einzelne Anforderung im Anforderungsdokument. Zähle, wie viele Aspekte dieser Anforderung zueinander im Widerspruch stehen und setze diese Anzahl zu der Gesamtanzahl der Aspekte der Anforderung ins Verhältnis.

Metriken zur „inhaltlichen Qualität“¹:

GF2-M3.1 dient zur Ermittlung der Redundanz von Anforderungen.

GF2-M3.1: Redundanz	
Zur Frage:	GF2-Q3.1 Wie viele Anforderungen tauchen mehrfach im Dokument auf?
Metrik:	$\frac{\sum \text{Redundante Anforderungen}}{\sum \text{Anforderungen}}$
Einheit:	Prozentual [%]
Skala:	0% – 100%
Datenquelle:	Anforderungsdokument
Beschreibung:	Identifiziere eine einzelne Anforderung im Anforderungsdokument. Prüfe, ob sie an einer weiteren Stelle im Dokument vorkommt bzw. ob es eine Anforderung mit identischem Inhalt im Dokument gibt. Zähle die redundanten Anforderungen und setze sie zu der Gesamtanzahl der Anforderungen ins Verhältnis.

GF2-M3.2 misst, ob eine Anforderung im Widerspruch zu einer oder mehreren anderen Anforderungen im Dokument steht.

GF2-M3.2: (Externe) Widersprüchlichkeit einer Anforderung	
Zur Frage:	GF2-Q3.2 Wie viele Anforderungen widersprechen sich gegenseitig?
Metrik:	$\frac{\sum \text{Widersprüchliche Anforderungen}}{\sum \text{Anforderungen}}$
Einheit:	Prozentual [%]
Skala:	0% – 100%
Datenquelle:	Anforderungsdokument
Beschreibung:	Identifiziere eine einzelne Anforderung im Anforderungsdokument. Prüfe, ob sie inhaltlich im Widerspruch zu einer oder mehrerer anderer Anforderungen im Anforderungsdokument steht.

¹ Es handelt sich faktisch eher um eine „formale Qualität“ - da einige Metriken zu wesentlichen inhaltlichen Aspekten nicht in den Messplan eingeflossen sind

GF2-M4.1 überprüft, wie viele Aspekte einer Anforderung sich messen bzw. objektiv auf Erfüllung prüfen lassen.

GF2-M4.1: Prüfbarkeit einer Anforderung	
Zur Frage:	GF2-Q4.1 Wie viele Aspekte der Anforderung sind messbar?
Metrik:	$\frac{\sum \text{Prüfbare Aspekte der Anforderung}}{\sum \text{Aspekte der Anforderung}}$
Einheit:	Prozentual [%]
Skala:	0% – 100%
Datenquelle:	Anforderungsdokument
Beschreibung:	Identifiziere eine einzelne Anforderung im Anforderungsdokument. Zähle die Aspekte der Anforderung, die messbar (d.h. objektiv auf Erfüllung überprüfbar) sind und setze sie zu allen Aspekten der Anforderung ins Verhältnis.

GF2-M8.1 überprüft, ob für eine Anforderung die Wichtigkeit angegeben worden ist.

GF2-M8.1: Bewertbarkeit	
Zur Frage:	GF2-Q.8.1 Wurde die Wichtigkeit einer Anforderung eingeschätzt?
Metrik:	Suche nach Einschätzung der Wichtigkeit einer Anforderung.
Einheit:	-
Skala:	Ja / Nein
Datenquelle:	Anforderungsdokument
Beschreibung:	Identifiziere eine einzelne Anforderung im Anforderungsdokument. Suche nach der Einschätzung der Wichtigkeit dieser Anforderung beispielsweise durch Verwendung von Ausdrücken wie „Hohe Priorität hat...“, „Am wichtigsten ist...“ etc. Hinweis: Die Anforderungen, die im Anforderungsdokument im Kapitel „1.2 Wünsche und Prioritäten des Kunden“ in einer nummerierten Liste auftauchen, gelten als bewertet.

GF2-M9.1 überprüft, ob die rechtlichen Verbindlichkeiten einer Anforderung geklärt sind.

GF2-M9.1: Klassifizierbarkeit	
Zur Frage:	GF2-Q9.1 Wurde die rechtliche Verbindlichkeit einer Anforderung eingeschätzt?
Metrik:	Suche nach Benennung der rechtlichen Verbindlichkeit einer Anforderung.
Einheit:	-
Skala:	Ja / Nein
Datenquelle:	Anforderungsdokument
Beschreibung:	Identifiziere eine einzelne Anforderung im Anforderungsdokument. Suche nach der rechtlichen Verbindlichkeit dieser Anforderung.

GN2.1-M1.2 überprüft, ob eine technische Anforderung in eine entsprechende Kategorie eingeordnet worden ist.

GN2.1-M1.2: Zuordnung zu einer Kategorie	
Zur Frage:	GN2.1-Q1.2: Wurde zu einer technischen Anforderung die Kategorie genannt?
Metrik:	Suche nach Nennung der Kategorie einer technischen Anforderung.
Einheit:	-
Skala:	Ja / Nein
Datenquelle:	Anforderungsdokument
Beschreibung:	Identifiziere eine einzelne technische Anforderung im Anforderungsdokument. Überprüfe, ob die Anforderung einer technischen Kategorie zugeordnet worden ist (z.B. „Anforderungen an die Programmiersprache“, „Anforderungen an Hard-/Software auf dem Zielsystem“, etc.).

GN2.1-M1.3 ermittelt die Anzahl der Kategorien von technischen Anforderungen im Anforderungsdokument.

GN2.1-M1.3: Anzahl Kategorien von technischen Anforderungen im Dokument	
Zur Frage:	GN2.1-Q1.3: Wie viele Kategorien von technischen Anforderungen tauchen im Dokument auf?
Metrik:	Zähle die Kategorien von technischen Anforderungen im Dokument.
Einheit:	-
Skala:	Ordinalskala
Datenquelle:	Anforderungsdokument
Beschreibung:	Zähle die Anzahl der genannten technischen Kategorien im Anforderungsdokument.

GN2.2-M1.1 überprüft, ob es eine visuelle Beschreibung der Benutzerschnittstelle gibt – etwa in der Form einer (handschriftlichen) Skizze o.Ä.

GN2.2-M1.1: Vorhandensein einer visuellen Beschreibung der Benutzerschnittstelle	
Zur Frage:	GN2.2-Q1.1 Existiert eine visuelle Beschreibung der Benutzerschnittstelle?
Metrik:	Suche nach visueller Beschreibung der Benutzerschnittstelle.
Einheit:	-
Skala:	Ja / Nein
Datenquelle:	Anforderungsdokument
Beschreibung:	Suche im Anforderungsdokument nach einer visuellen Beschreibung der Benutzerschnittstelle (z.B. Screenshots, GUI-Entwürfe, handschriftliche Skizzen, etc.).

GN2.2-M1.2 prüft, ob im Anforderungsdokument eine Beschreibung der zukünftigen Benutzer der Software und deren Fähigkeiten enthalten ist.

GN2.2-M1.2: Vorhandensein eines Benutzerprofils	
Zur Frage:	GN2.2-Q1.2 Wird präzisiert welche Benutzer das Programm bedienen sollen?
Metrik:	Suche nach Beschreibung der zukünftigen Benutzer und deren Fähigkeiten.
Einheit:	-
Skala:	Ja / Nein
Datenquelle:	Anforderungsdokument
Beschreibung:	Suche im Anforderungsdokument nach einer Beschreibung der zukünftigen Benutzer des Programms.

GN2.2-M1.3 überprüft, ob in den Anforderungen an die Benutzerschnittstelle, die eine Benutzerinteraktion beinhalten, die Eingabegeräte genannt werden, mit denen die Interaktion erfolgen soll.

GN2.2-M1.3: Benennung von Eingabegeräten	
Zur Frage:	GN2.2-Q1.3 Werden bei Anforderung mit Benutzerinteraktion die Eingabegeräte benannt?
Metrik:	Suche in Anforderungen an die Benutzerschnittstelle mit Benutzerinteraktionen nach Nennung der Eingabegeräte.
Einheit:	-
Skala:	Ja / Nein
Datenquelle:	Anforderungsdokument
Beschreibung:	Identifiziere eine Anforderung an die Benutzerschnittstelle, die eine Benutzerinteraktion beinhaltet. Überprüfe, ob genannt wird, mit welchen Eingabegeräten die Interaktion erfolgen soll (z.B. Maus, Tastatur, biometrische Systeme [Fingerabdrucksensor, Iris-Scan, etc.], etc.).

GN2.3-M1.1 prüft, ob ein Qualitätsmodell im Anforderungsdokument existiert.

GN2.3-M1.1: Vorhandensein eines Qualitätsmodells	
Zur Frage:	GN2.3-Q1.1 Existiert ein Qualitätsmodell?
Metrik:	Suche nach einem Qualitätsmodell
Einheit:	-
Skala:	Ja / Nein
Datenquelle:	Anforderungsdokument
Beschreibung:	Suche im Anforderungsdokument nach einem Qualitätsmodell.

GN2.3-M1.2 überprüft, ob allgemeine Qualitätsziele konkretisiert werden.

GN2.3-M1.2: Nennung von Qualitätszielen	
Zur Frage:	GN2.3-Q1.2 Werden konkrete Qualitätsziele genannt?
Metrik:	Suche zu einem allgemeinen Qualitätsziel nach konkreten (!) Qualitätszielen
Einheit:	-
Skala:	Ja / Nein
Datenquelle:	Anforderungsdokument
Beschreibung:	Identifiziere ein allgemeines Qualitätsziel im Anforderungsdokument. Prüfe, ob es konkretisiert worden ist bzw. ob konkrete Qualitätsziele für das allgemeine Qualitätsziel angegeben worden sind.

GN2.3-M1.3 überprüft, ob zu den konkreten Qualitätszielen Metriken bzw. objektiv prüfbare Erfüllungskriterien genannt werden.

GN2.3-M1.3: Metriken für Qualitätsziele	
Zur Frage:	GN2.3-Q1.3 Wird angegeben, wann die Qualitätsziele erfüllt sind?
Metrik:	Suche zu einem konkreten Qualitätsziel nach prüfbaren Kriterien für die Erfüllung
Einheit:	-
Skala:	Ja / Nein
Datenquelle:	Anforderungsdokument
Beschreibung:	Identifiziere ein konkretes Qualitätsziel im Anforderungsdokument. Prüfe, ob angegeben worden ist, wann es erfüllt ist (z.B. durch eine Zeitbeschränkung).

4 Analyse von Projekten

4.1 Vorgehensweise

Zur Analyse der Projekte werden ausschließlich Anforderungsspezifikationen untersucht, die zum ersten Quality Gate eines Softwareprojekts im FG Software Engineering eingereicht worden sind und dieses Quality Gate erfolgreich passiert haben. Die grundsätzliche Struktur dieser Dokumente ergibt sich aus einem Word-Template des FG Software Engineering und wird von daher als ideal angenommen und im Folgenden nicht weiter untersucht werden.

Im Gegensatz zu [Hor07], wo jede Anforderung zunächst identifiziert und mit Hilfe der *Anforderungsschablone* [Rup06] (vgl. Abbildung 4.1) umformuliert worden ist, werden im Folgenden die Anforderungen in sprachlich unverfälschter Form analysiert. Die Anforderungsschablone hat die (sonst positive) Eigenschaft, sprachliche Fehler (also Verstöße gegen die Formulierungsregeln) zu eliminieren; Ziel bei der folgenden Analyse ist es jedoch auch, gerade diese Verstöße auch zu finden.

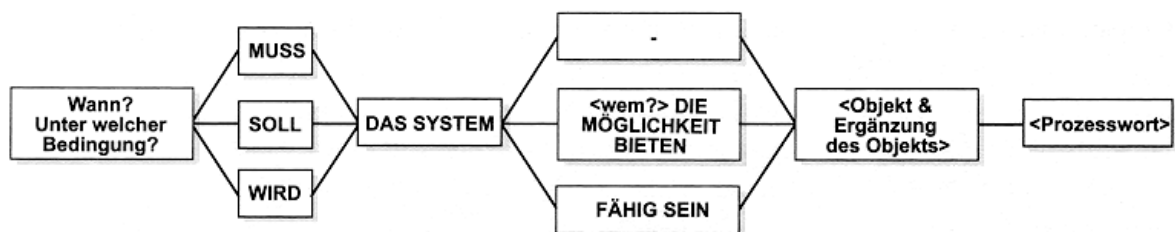


Abbildung 4.1.: Anforderungsschablone nach [Rup06]

Anforderungen aus Use-Cases werden wie folgt behandelt: Die Aktionen des Benutzers werden bis zur ersten Systemreaktion zu einem Satz zusammengefasst, die Aktionen bis zur zweiten Systemreaktion zu einem zweiten Satz etc.

Beispiel:

1.	Benutzer	stellt Eigenschaften der Komponente ein und klickt „übernehmen“
2.	Programm	übernimmt Änderungen

Anforderung in einem Satz: *Benutzer stellt Eigenschaften der Komponente ein und klickt „übernehmen“, Programm übernimmt Änderungen.*

Bei Anforderungen aus Use-Cases werden nur die Schritte aus dem Hauptszenario behandelt.

4.2 Messung mittels Fragebogen-Programm

Um die manuelle Analyse zu beschleunigen und die Daten der Auswertungen elektronisch zu speichern, wurde ein Fragebogen-Programm in PHP / MySQL erstellt, welches es dem Benutzer erlaubt, einzelne Metriken auf ein konkretes Anforderungsdokument anzuwenden.

Um ein Anforderungsdokument eines Projekts zu analysieren, muss der Benutzer des Programms folgende Schritte durchführen:

1. Anlegen eines neuen Projekts im Fragebogen-Programm
2. Projektdaten (Projektname, optionale Beschreibung und Anforderungen in Textform) eingeben
3. Eine Metrik auswählen und die Anforderungen satzweise abarbeiten

Die Anforderungen eines Projekts werden beim Anlegen eines neuen Projekts im Fragebogen-Programm direkt aus dem Anforderungsdokument mit Copy-Paste übernommen. Im nächsten Schritt zerlegt das Programm den Anforderungstext in einzelne Sätze, welche nun vom Benutzer klassifiziert werden können. So kann der Benutzer für jeden einzelnen Satz angeben, um welchen Anforderungstyp es sich handelt oder ob es sich eventuell gar nicht um eine Anforderung, sondern nur um einen Kommentar handelt. Die Abbildungen 4.2 und 4.3 zeigen die Eingabe der Projektdaten und die Klassifikation der Anforderungen.

Neues Projekt anlegen

Projektdaten

Projektname:* FVE-XML Editor

Semester:* 2007 WS

Beschreibung:

Anforderungen:* Höchste Priorität besitzt für den Kunden die Entwicklung eines Editors, welcher ein schnelles Überblicken und Editieren von XML-Dokumenten ermöglichen soll. Die Usability ist dabei die zentrale Anforderung. Es soll mit einem Blick möglich sein zu erkennen, wie das Dokument strukturiert ist. Ebenso sollen das Erstellen neuer Dokumente sowie das Editieren bereits bestehender XML-Bäume möglichst schnell und einfach sein. Der Kunde wünscht sich deshalb einen Editor, welcher Baumstruktur und Texteditor vereint und dabei die für den Nutzer redundanten Informationen der Extensible Markup Language (XML) vermeidet. Der Kunde favorisiert die Realisierung des Produkts als Plug-In für die integrierte Entwicklungsumgebung „Eclipse“ und den damit bestehenden XML-Editor des Web Tools Platform Projects

Projekt anlegen

Abbildung 4.2: Angabe der Projektdaten beim Anlegen eines neues Projekts

Anforderung

Höchste Priorität besitzt für den Kunden die Entwicklung eines Editors, welcher ein schnelles Überblicken und Editieren von XML-Dokumenten ermöglichen soll

Kein Requirement

F.A.: Funktionale Anforderung

Die Usability ist dabei die zentrale Anforderung

Kein Requirement

NF.A.: Qualitätsanforderung

Abbildung 4.3: Klassifikation von Anforderungen

Nach der Auswahl einer Metrik, zeigt das Programm dem Benutzer je einen Satz an, welchen er gegen eine angezeigte Checkliste prüft. Abbildung 4.4 veranschaulicht diesen Vorgang.

Checklisten-Tool

Projektauswahl
Ausgewähltes Projekt: FVE-Edit XML-Editor

Bewerten Sie hier, welche der aufgeführten Formulieringsregeln die angezeigte Anforderung erfüllt.

Hinweis: Benutzen Sie die TAB-Taste, um zur nächsten Checkbox zu navigieren. Die Checkboxes können Sie mit der SPACE-Taste steuern!

Anforderung	Checkliste
Um den Überblick über das Dokument zu erleichtern, sollen bekannte Features wie Code-Folding und Syntax-Highlighting implementiert werden.	<input type="checkbox"/> Kein Passiv <input checked="" type="checkbox"/> Prozesse durch Vollverben <input checked="" type="checkbox"/> Keine unvollständig spez. Prozesswörter <input checked="" type="checkbox"/> Keine unvollständigen Vergleiche <input checked="" type="checkbox"/> Constraints ausreichend genau beschrieben <input checked="" type="checkbox"/> Keine Universalquantoren <input checked="" type="checkbox"/> Keine Substantive ohne Bezugsindex <input checked="" type="checkbox"/> Substantive im Singular (wo Plural nicht explizit gefordert ist) <input checked="" type="checkbox"/> Kein Nominalstil (Checkboxen mit "Space" steuern)
	<input type="button" value="weiter"/>

[Zurück zur Metrikauswahl](#)

Abbildung 4.4.: Anwendung einer Metrik auf eine Anforderung

Für jede Metrik kann sich der Benutzer nach Bewertung der Anforderungen des Projekts eine Auswertung anzeigen lassen.

Checklisten-Tool

Auswertung	
Projekt:	FVE-Edit XML-Editor
Anteil der Anforderungen, die Regeln verletzen, in Prozent:	
Enthält Passiv:	29,23 %
Prozesse nicht durch Vollverben ausgedrückt:	10,77 %
Unvollst. spez. Prozesswörter:	7,69 %
Unvollständige Vergleiche:	1,54 %
Constraints nicht genau genug beschrieben:	1,54 %
Enthält Universalquantoren:	1,54 %
Enthält Substantive ohne Bezugsindex:	4,62 %
Enthält Plural (wo auch Singular richtig wäre):	0,00 %
Nominalstil:	21,54 %
<hr/>	
GF1-M1.1: Wie viele Regeln wurden im Durchschnitt eingehalten?	91,28 %

[Zurück zur Metrikauswahl](#)

Abbildung 4.5.: Messergebnisse einer Metrik

4.3 Auswertung

Mit Hilfe des Fragebogen-Programms wurden die Software-Projekte aus dem Wintersemester 2007 / 2008 und aus dem Wintersemester 2006 / 2007 analysiert (insgesamt 16 Projekte). Die Messergebnisse und die Interpretation dieser Daten ist Gegenstand dieses Kapitels.

4.3.1 Messergebnisse

In diesem Kapitel werden die jeweils besten und schlechtesten Messergebnisse zu jeder Metrik vorgestellt.

Metrik	Bestes Messergebnis	Schlechtestes Messergebnis
GF1-M1.1: Formulierungsregelkonformität einer Anforderung	Sinologie in Würzburg (92,92% der Formulierungsregeln wurden im Durchschnitt eingehalten)	NEH-Sortieranlage – Cybertec (84,68%)
GF1-M2.1: Erläuterung von Fachbegriffen	RiskQG (97,47% der Fachbegriffe sind im Glossar erklärt)	Sinologie in Würzburg (27,91%)
GF1-M2.2: Anzahl der sprachlichen Fehler, die die Verstehbarkeit beeinflussen	(Hypothese, dass sprachliche Fehler die Verstehbarkeit beeinflussen, kann nicht bestätigt werden, 0 bei allen Projekten)	
GF1-M3.1: Identifizierbarkeit von Anforderungen	Key-Sign-Org 1 (98,39% der Sätze enthalten genau eine Anforderung)	SOA-Me-Client (64,47%)
GF1-M3.2: Anzahl mehrdeutiger Begriffe innerhalb einer Anforderung	<ul style="list-style-type: none"> • Key-Sign-Org 1 • RiskQG Beide 0,05 mehrdeutige Begriffe pro Anforderung (Mittelwert)	Process Pattern Matching (0,51 - Mittelwert)
GF1-M3.3: Mehrdeutigkeit durch sprachliche Fehler	(Hypothese, dass sprachliche Fehler zu Mehrdeutigkeit führen, kann nicht bestätigt werden, 0 bei allen Projekten)	
GF1-M4.1: Widersprüchlichkeit einer Anforderung	<ul style="list-style-type: none"> • UML Paint 2 • Fast View & Edit XML Editor • Software-Quanten-Simulation • NEH-Sortieranlage (Drehkreuz-Ritter) • SOA-Me Integrationsserver 0,00 % Widersprüche pro Anforderung (Mittelwert)	UML Paint (3,75 % - Mittelwert)
GF2-M3.1: Redundanz	KeysignOrg2 (3,64% redundante Anforderungen)	Key-Sign-Org 1 (32,26%)

Metrik	Bestes Messergebnis	Schlechtestes Messergebnis
GF2-M3.2: (Externe) Widersprüchlichkeit einer Anforderung	<ul style="list-style-type: none"> • UML Paint 2 • KeysignOrg2 • SOA-Me-Client <p>0,00% widersprüchliche Anforderungen</p>	Fast View & Edit XML Editor (18,46%)
GF2-M4.1: Prüfbarkeit einer Anforderung	Goal-Question-Metric Unterstützung (89,40% aller Anforderungsaspekte sind prüfbar)	NEH-Sortieranlage (Gruppe A) (63,70%)
GF2-M8.1: Bewertbarkeit	UML Paint 2 (42,06% der Anforderungen sind nach ihrer Wichtigkeit bewertet worden)	NEH-Sortieranlage (Drehkreuz-Ritter) (9,23%)
GF2-M9.1: Klassifizierbarkeit	NEH-Sortieranlage (Cybertec) (Bei 5,41% der Anforderungen sind die rechtlichen Verbindlichkeiten angeben)	<ul style="list-style-type: none"> • Key-Sign-Org 1 • Fast View & Edit XML Editor • UML Paint • KeysignOrg2 • Goal-Question-Metric Unterstützung • Software-Quanten-Simulation • Sinologie in Würzburg • Process Pattern Matching • SOA-A3EI • SOA-Me Integrationserver • SOA-Me-Client <p>(0,00%)</p>
GN2.1-M1.2: Zuordnung zu einer Kategorie	RiskQG (83,33% der technischen Anforderungen sind einer Kategorie zugeordnet)	<ul style="list-style-type: none"> • NEH-Sortieranlage (Gruppe A) • UML Paint <p>(0,00 %)</p>
GN2.1-M1.3: Anzahl Kategorien von technischen Anforderungen im Dokument	Eine Kategorie von technischen Anforderungen in jedem Dokument (ggf. auch auf Grund der Vorgabe im Dokumententemplate), Einzige Ausnahme: „Fast View & Edit XML Editor“: 2 Kategorien	

Metrik	Bestes Messergebnis	Schlechtestes Messergebnis
GN2.2-M1.1: Vorhandensein einer visuellen Beschreibung der Benutzerschnittstelle	<ul style="list-style-type: none"> • SOA-A3EI • SOA-Me-Client • Key-Sign-Org 1 • RiskQG • UML Paint 2 • Fast View & Edit XML Editor • UML Paint • KeysignOrg2 • Goal-Question-Metric Unterstützung • Software-Quanten-Simulation <p>Es existiert eine visuelle Beschreibung der Benutzerschnittstelle</p>	<ul style="list-style-type: none"> • Sinologie in Würzburg • NEH-Sortieranlage (Cybertec) • NEH-Sortieranlage (Drehkreuz-Ritter) • NEH-Sortieranlage (Gruppe A) • Process Pattern Matching • SOA-Me Integrationsserver <p>Es existiert keine visuelle Beschreibung der Benutzerschnittstelle</p>
GN2.2-M1.2: Vorhandensein eines Benutzerprofils	Benutzerprofil in jedem Projekt vorhanden, Ausnahme: „NEH-Sortieranlage (Cybertec)“: Keine Benutzerprofil vorhanden	
GN2.2-M1.3: Benennung von Eingabegeräten	<ul style="list-style-type: none"> • Software-Quanten-Simulation • Fast View & Edit XML Editor <p>(33,33% der Anforderungen an die Benutzerschnittstelle enthalten die Angabe von Eingabegeräten)</p>	<ul style="list-style-type: none"> • Key-Sign-Org 1 • RiskQG • UML Paint 2 • KeysignOrg2 • Goal-Question-Metric Unterstützung • Software-Quanten-Simulation • Sinologie in Würzburg • NEH-Sortieranlage (Gruppe A) • Process Pattern Matching • SOA-A3EI • SOA-Me Integrationsserver <p>(0,00%) (In den restlichen Projekten, war die Metrik nicht anwendbar, da es keine entsprechenden Anforderungen gab)</p>
GN2.3-M1.1: Vorhandensein eines Qualitätsmodells	In keinem Projekt wurde ein Qualitätsmodell angegeben	

Metrik	Bestes Messergebnis	Schlechtestes Messergebnis
GN2.3-M1.2: Nennung von Qualitätszielen	<ul style="list-style-type: none"> • Sinologie in Würzburg • NEH-Sortieranlage (Drehkreuz-Ritter) • SOA-A3EI • SOA-Me Integrationsserver • UML Paint • Goal-Question-Metric Unterstützung (100,00% der Qualitätsziele ist konkretisiert)	NEH-Sortieranlage (Gruppe A) (22,22%)
GN2.3-M1.3: Metriken für Qualitätsziele	SOA-Me Integrationsserver (100,00% der konkretisierten Qualitätsziele besitzen Metriken)	<ul style="list-style-type: none"> • Key-Sign-Org 1 • UML Paint 2 • Fast View & Edit XML Editor • KeysignOrg2 • Software-Quanten-Simulation (0,00%)

Tabelle 4.3.1: Bestes - / schlechtestes Messergebnis pro Metrik

Die detaillierten Auswertungen zu allen Projekten finden sich in tabellarischer Form auf der beigelegten CD-ROM.

4.3.2 Abgleich der Ergebnisse mit den Hypothesen

Im Folgenden werden nun die Messergebnisse mit den Hypothesen (s. Kapitel 3.2, Abstraction Sheets) abgeglichen und damit überprüft, inwiefern die Einschätzungen richtig waren.

Betrachtet man die Messergebnisse zum Ziel „GF1: *Untersuche die sprachliche Qualität einer funktionalen Anforderung*“ (Ergebnisse gelten analog auch für nicht-funktionale Anforderungen), so kann die Hypothese, dass die Anforderungen in natürlicher Sprache ohne Berücksichtigung von Formulierungsregeln formuliert worden sind, bestätigt werden (Metrik: GF1-M1.1). Obwohl die Formulierungsregeln im Durchschnitt zwar zum Großteil eingehalten worden sind, deutet die Anzahl der Verstöße gegen die Regeln im Detail darauf hin, dass die Einhaltung der Regeln eher einen Zufall darstellt und nicht durch gezielte Anwendung eines Formulierungsregelwerkes (vgl. Abbildung 4.1) zu Stande gekommen ist. Hervorzuheben ist hier das Projekt „Sinologie in Würzburg“, wo offensichtlich die Anforderungsschablone (s. Abbildung 4.1) verwendet worden ist.

Die Hypothese, dass Fachbegriffe zwar verwendet, jedoch nicht erläutert werden, kann teilweise bestätigt werden (Metrik: GF1-M2.1). Oft wurde nur ca. die Hälfte der Fachbegriffe im Glossar erläutert. Die Qualität der Erläuterung wurde nicht weiter untersucht.

Die Hypothesen, dass sprachliche Fehler dazu führen können, dass eine Anforderung unverständlich bzw. mehrdeutig wird, können nicht bestätigt werden (Metriken: GF1-M2.2 und GF1-M3.3).

Die Hypothese, dass sich Zusammenhänge innerhalb eines Satzes widersprechen, kann mit den Messergebnissen widerlegt werden (Metrik: GF1-M4.1). Es ließen sich i.A. nur wenige Sätze finden, die Widersprüche enthielten.

Betrachtet man die Messergebnisse zum Ziel „GF2: *Untersuche die inhaltliche Qualität einer funktionalen Anforderung*“ (Ergebnisse gelten auch für nicht-funktionale Anforderungen), so kann die Hypothese, dass der Inhalt einer Anforderung in Widerspruch zu einer oder mehreren anderen Anforderungen steht (Metrik: GF2-M3.2), teilweise bestätigt werden. So war der Anteil der widersprüchlichen Anforderungen bei den meisten Projekten eher gering (0% - 10%), wobei auffällt, dass Projekte aus dem WS 2006 / 2007 öfter Werte über 10% zeigten.

Die Hypothese, dass Anforderungen vage und ohne nachprüfbare Kriterien zur Erfüllung formuliert wurden (Metrik: GF2-M4.1), kann nicht bestätigt werden. So existierten für die meisten Anforderungsaspekte nachprüfbare Kriterien zur Erfüllung (> 70%). Auffallend ist, dass Projekte aus dem WS 2007 / 2008 oft einen noch höheren Anteil an prüfbaren Anforderungsaspekten enthielten (öfter Werte > 80%), als Projekte aus WS 2006 / 2007.

Die Hypothese, dass Anforderungen nicht bewertet (also nach ihrer Priorität eingeschätzt) worden sind (Metrik: GF2-M8.1), kann bestätigt werden. Oft wurden nur ca. 20% der Anforderungen nach ihrer Priorität bewertet. Abermals fallen Projekte aus dem WS 2007 / 2008 hierbei positiv auf: Sie zeigen teilweise Werte von ca. 30% - 40%.

Die Hypothese, dass rechtliche Aspekte nicht erwähnt worden sind (Metrik: GF2-M9.1), kann bestätigt werden. So waren in fast keinem Projekt rechtliche Verbindlichkeiten in den Anforderungen erwähnt.

Betrachtet man die Messergebnisse zum Ziel „GN2.1: *Untersuche die inhaltliche Qualität einer technischen Anforderung*“, so kann die Hypothese, dass technische Anforderungen häufig keiner Kategorie zugeordnet sind, nicht klar bestätigt oder widerlegt werden (Metrik: GN2.1-M1.2). Die Ergebnisse schwanken hier sehr stark: Während in manchen Dokumenten keinerlei technische Anforderungen einer technischen Kategorie zugeordnet sind (0,00%), sind in anderen Projekten so gut wie alle technischen Anforderungen zugeordnet (70% - 80%). Die besten Ergebnisse finden sich abermals im WS 2007 / 2008.

Bezüglich der Messergebnisse zum Ziel „GN2.2: *Untersuche die inhaltliche Qualität einer Anforderung an die Benutzerschnittstelle*“, kann die Hypothese, dass keine visuelle Beschreibung der Benutzerschnittstelle vorliegt, teilweise bestätigt werden (Metrik: GN2.2-M1.1). Im WS 2006 / 2007 fanden sich nur in zwei von acht Anforderungsdokumenten eine visuelle Beschreibung der Benutzerschnittstelle. Positiv fällt auf, dass im WS 2007 / 2008 in jedem Dokument eine visuelle Beschreibung der Benutzerschnittstelle vorhanden war.

Die Hypothese, dass die Bedürfnisse der unterschiedlichen Benutzer nur unzureichend berücksichtigt werden (es ist kein Benutzerprofil vorhanden – Metrik: GN2.2-M1.2), kann nicht bestätigt werden. Bis auf eine Ausnahme existierte in jedem Anforderungsdokument eine Beschreibung der zukünftigen Nutzer der Software; dies ist vor allem durch die Vorgabe im Dokumenten-Template zu erklären, welches eine Nennung der zukünftigen Nutzer der Software explizit fordert.

Die Hypothese, dass bei Anforderungen an die Benutzerschnittstelle keine Eingabegeräte genannt werden, mit denen die Benutzerinteraktion erfolgen soll (Metrik: GN2.2-M1.3), kann nicht klar bestätigt werden. Zu erwähnen ist, dass in manchen Projekten es keine Anforderungen gab, die als Anforderungen an die Benutzerschnittstelle gewertet werden konnten, da oftmals in den Anforderungen nicht nur Details zur graphischen Benutzerschnittstelle vorhanden waren, sondern auch das Systemverhalten beschrieben worden ist. Somit wurden jene Anforderungen häufig als funktionale Anforderungen gewertet.

Beim Ziel „GN2.3: *Untersuche die inhaltliche Qualität einer Qualitätsanforderung*“, fällt insbesondere auf, dass es in keinem Anforderungsdokument ein Qualitätsmodell gibt (Metrik: GN2.3-M1.1). Bezüglich der Frage, wie viele Qualitätsziele konkretisiert worden sind (Metrik: GN2.3-M1.2), ergibt sich ein leichter Vorteil für die Projekte aus dem WS 2006 / 2007, wo gleich in mehreren Projekten alle genannten Qualitätsziele konkretisiert worden sind (100% konkretisierte Qualitätsziele). Auch bei der Frage, für wie viele Qualitätsziele messbar Kriterien zur Erfüllung angegeben worden sind (Metrik: GN2.3-M1.3), schneiden die Projekte aus dem WS 2006 / 2007 besser ab: Es finden sich häufiger Werte um die 50% oder höher, während bei den Projekten aus dem WS 2007 / 2008 bereits 25% den Spitzenwert darstellt, wobei in der Hälfte der Projekte für keinerlei Qualitätsziele messbare Kriterien angegeben worden sind (0,00% der konkretisierten Qualitätsziele sind messbar).

4.3.3 Gewichtung und Vergleich der Ergebnisse

Wie kommt man nun von den einzelnen Messergebnissen zu der Vergleichbarkeit von Anforderungen bzw. der Vergleichbarkeit der Anforderungsdokumente der einzelnen Projekte? Um die Vergleichbarkeit zu erreichen, werden die Messergebnisse den ursprünglichen Qualitätsanforderungen von Anforderungen zuordnet.

Tabelle 4.3.2 stellt die gewichteten Qualitätsfaktoren dar. Die **hervorgehoben** gedruckten Einträge sind an die Prioritäten aus [Dav05] angelehnt.

Qualitätsaspekt	Gewicht
Vollständigkeit	2
Korrektheit	10
Klassifizierbarkeit	1
Konsistenz	10
Prüfbarkeit	8
Eindeutigkeit	2
Verstehbarkeit	5
Gültig und aktuell	9
Realisierbarkeit	10
Notwendigkeit	3
Bewertbarkeit	1

Tabelle 4.3.2: Gewichtete Qualitätsfaktoren

Tabelle 4.3.3 gibt einen Überblick über die Qualitätsfaktoren und die Metriken und Heuristiken, die ihnen zugeordnet sind. Heuristiken werden dabei weniger stark gewichtet als Metriken. Dies erklärt sich dadurch, dass die Befunde, die durch Heuristiken gefunden

werden, weniger zuverlässig sind als die Befunde, die durch Metriken aufgedeckt werden – insbesondere die Überprüfung der Formulierungsregeln ergibt lediglich Hinweise auf mögliche Defekte [Rup06].

Qualitätsaspekt	Gewicht	Metriken	Typ	Gewicht
Vollständigkeit	2	Anforderung enthält kein Passiv?	Heuristik	1
		Prozesse durch Verben?	Heuristik	1
		Keine unvollständig spez. Prozesswörter?	Heuristik	1
		Keine unvollständigen Vergleiche?	Heuristik	1
		Constraints ausreichend beschrieben?	Heuristik	1
		Kein Nominalstil?	Heuristik	1
Korrektheit	10	Kein Einsatz von Universalquantoren?	Heuristik	1
Klassifizierbarkeit	1	GF2-M9.1: Klassifizierbarkeit	Metrik	2
Konsistenz	10	GF1-M4.1: Widersprüchlichkeit einer Anforderung	Metrik	2
		GF2-M3.1: Redundanz	Metrik	2
		GF2-M3.2: (Externe) Widersprüchlichkeit einer Anforderung	Metrik	2
Prüfbarkeit	8	GF2-M4.1: Prüfbarkeit einer Anforderung	Metrik	2
Eindeutigkeit	2	GF1-M3.1: Identifizierbarkeit von Anforderungen	Metrik	2
		Keine Substantive ohne Bezugsindex?	Heuristik	1
		Kein Plural, wo auch Singular richtig?	Heuristik	1
Verstehbarkeit	5	GF1-M2.1: Erläuterung von Fachbegriffen	Metrik	2
		GF1-M3.2: Anzahl mehrdeutiger Begriffe innerhalb einer Anforderung	Metrik	2
		GF1-M3.3: Mehrdeutigkeit durch sprachliche Fehler	Metrik	2
		GF1-M2.2: Anzahl der sprachlichen Fehler, die die Verstehbarkeit beeinflussen	Metrik	2
Gültig und aktuell	9	(Metrik vorhanden, jedoch nicht im Messplan)	Metrik	2
Realisierbarkeit	10	-	-	
Notwendigkeit	3	(Metrik vorhanden, jedoch nicht im Messplan)	Metrik	2
Bewertbarkeit	1	GF2-M8.1: Bewertbarkeit	Metrik	2

Tabelle 4.3.3: Übersicht der Qualitätsaspekte mit dazugehörigen Metriken / Heuristiken und den dazugehörigen Gewichten

Metrik	Gewicht
GN2.1-M1.2: Zuordnung zu einer Kategorie	2
(GN2.1-M1.3: Anzahl Kategorien von technischen Anforderungen im Dokument)	2
GN2.2-M1.1: Vorhandensein einer visuellen Beschreibung der Benutzerschnittstelle	2
GN2.2-M1.2: Vorhandensein eines Benutzerprofils	2
GN2.2-M1.3: Benennung von Eingabegeräten	2
GN2.3-M1.1: Vorhandensein eines Qualitätsmodells	2
GN2.3-M1.2: Nennung von Qualitätszielen	2
GN2.3-M1.3: Metriken für Qualitätsziele	2

Tabelle 4.3.4: Übersicht der Metriken für nicht-funktionale Anforderungen mit Gewichten

Ziel ist es, eine Kennzahl „Gesamtqualität“ für jedes Projekt zu erhalten, die die Messergebnisse aller Metriken einschließt. Bei „*positiven Metriken*“ (d.h. je höher das Messergebnis, um so höher die Qualität) wird das Messergebnis mit dem Gewicht der Metrik und danach mit dem Gewicht des Qualitätsfaktors multipliziert, dem die Metrik zugeordnet ist. Bei „*negativen Metriken*“ (d.h. je höher das Messergebnis, umso schlechter die Qualität) wird analog verfahren, das Ergebnis jedoch noch mit „-1“ multipliziert. Wie die Messergebnisse der einzelnen Metriken / Heuristiken berechnet bzw. gewichtet werden, geht aus Tabelle 4.3.5 hervor.

Metrik	Formel
Formulierungsregel	$(100 \% - [\text{Messergebnis: Versto\ss gegen Regel}])^*$ [Gewicht der Heuristik]* [Gewicht des Qualitätsfaktors]
Positive Metriken: <ul style="list-style-type: none"> • GF1-M2.1: Erläuterung von Fachbegriffen • GF1-M3.1: Identifizierbarkeit von Anforderungen • GF2-M4.1: Prüfbarkeit einer Anforderung • GF2-M8.1: Bewertbarkeit • GF2-M9.1: Klassifizierbarkeit • GN2.1-M1.2: Zuordnung zu einer Kategorie • (GN2.1-M1.3: Anzahl Kategorien von technischen Anforderungen im Dokument)¹ • GN2.2-M1.1: Vorhandensein einer visuellen Beschreibung der Benutzerschnittstelle • GN2.2-M1.2: Vorhandensein eines Benutzerprofils • GN2.2-M1.3: Benennung von Eingabegeräten • GN2.3-M1.1: Vorhandensein eines Qualitätsmodells • GN2.3-M1.2: Nennung von Qualitätszielen • GN2.3-M1.3: Metriken für Qualitätsziele 	$[\text{Messergebnis}]^*$ [Gewicht der Metrik]* [Gewicht des Qualitätsfaktors]
Negative Metriken: <ul style="list-style-type: none"> • GF1-M2.2: Anzahl der sprachlichen Fehler, die die Verstehbarkeit beeinflussen • GF1-M3.2: Anzahl mehrdeutiger Begriffe innerhalb einer Anforderung • GF1-M3.3: Mehrdeutigkeit durch sprachliche Fehler • GF1-M4.1: Widersprüchlichkeit einer Anforderung • GF2-M3.1: Redundanz • GF2-M3.2: (Externe) Widersprüchlichkeit einer Anforderung 	$(-1)^*[\text{Messergebnis}]^*$ [Gewicht der Metrik]* [Gewicht des Qualitätsfaktors]

Tabelle 4.3.5: Formeln zu Gewichtung der Messergebnisse

¹ Messergebnisse dieser Metrik werden nicht berücksichtigt, da sie nicht direkt mit der Qualität des Anforderungsdokuments zusammenhängen

Die Abbildungen 4.6 und 4.7 zeigen die gewichteten Messergebnisse für die Projekte. Die einzelnen Teilergebnisse addieren sich zur Gesamtqualität.

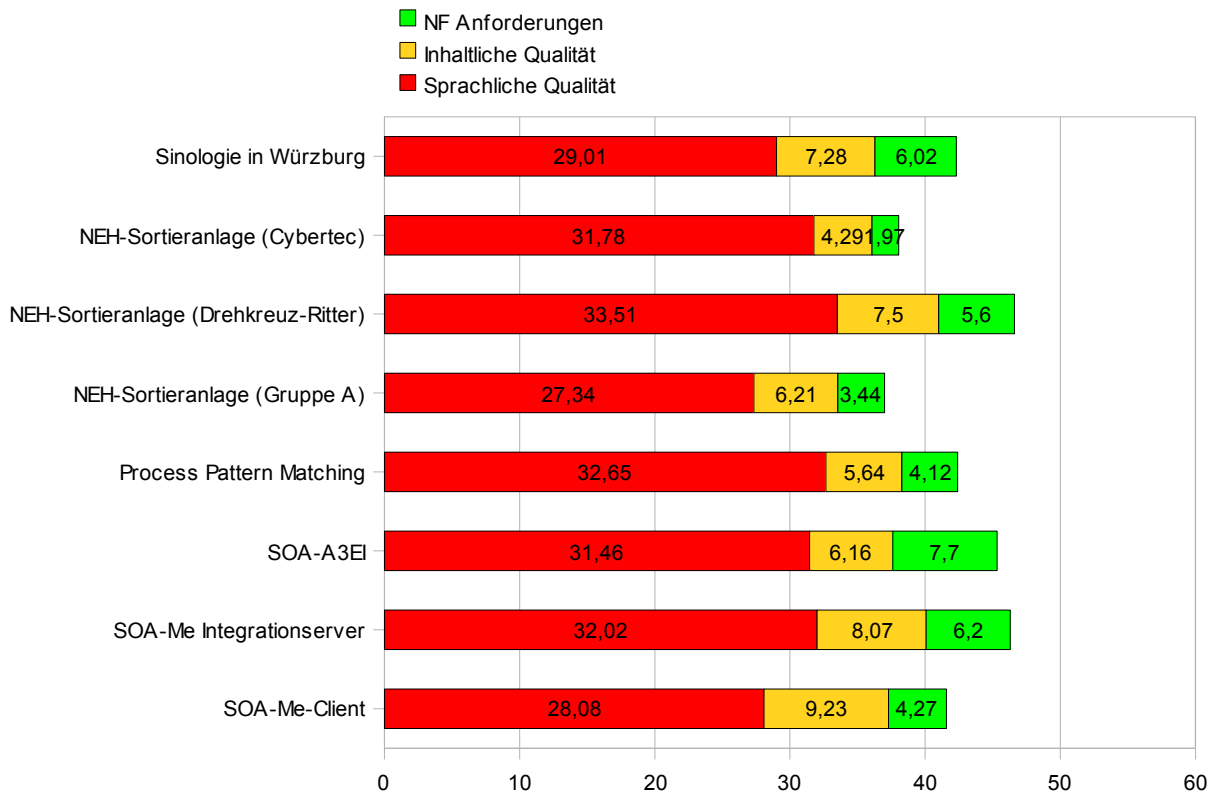


Abbildung 4.6: Gesamtergebnisse vom WS 2006 / 2007

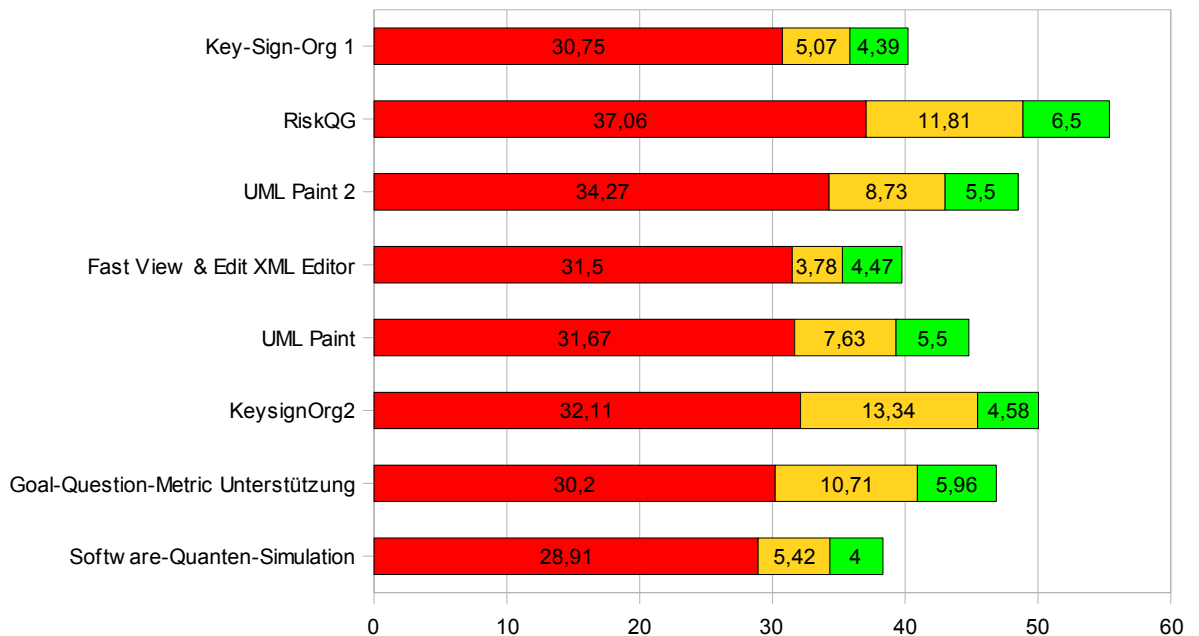


Abbildung 4.7: Gesamtergebnisse WS 2007 / 2008

Die einzelnen Messergebnisse sind auf einer Ordinalskala angeordnet. Je höher der Zahlenwert in einem Messbereich ist, umso besser ist die Qualität eines Projekts in diesem Bereich. Innerhalb eines Projektes können die einzelnen Messergebnisse zu „sprachlicher Qualität“, „inhaltlicher Qualität“ und „nicht-funktionalen Anforderungen“ untereinander nicht verglichen werden.

Das insgesamt beste Messergebnis ergibt sich für das Projekt „RiskQG“ aus dem WS 2007 / 2008: Bezüglich der sprachlichen Qualität fallen hier besonders das sehr vollständige Glossar (GF1-M2.1: 97,47% der Fachbegriffe tauchen im Glossar auf) sowie die sehr gute Identifizierbarkeit der Anforderungen (GF1-M3.1: 91,03% der Sätze enthalten genau eine Anforderung) auf.

Inhaltlich sind die geringen Werte bei Redundanz (GF2-M3.1: 7,69% redundante Anforderungen) und bei (externer) Widersprüchlichkeit (GF2-M3.2: 2,56% widersprüchliche Anforderungen) sowie der sehr hohe Anteil an prüfbaren Anforderungsaspekten (GF2-M4.1: 85,21% prüfbare Anforderungsaspekte) positiv zu erwähnen.

Gute Werte ergeben sich auch für die nicht-funktionalen Anforderungen: So sind viele der technischen Anforderungen in der entsprechenden Kategorie eingeordnet (GN2.1-M1.2: 83,33% der technischen Anforderungen sind einer Kategorie zugeordnet), es existiert eine grafische Beschreibung der Benutzerschnittstelle und die Qualitätsziele sind (vergleichsweise) gut konkretisiert (GN2.3-M1.2: 66,67% konkretisierte Qualitätsziele).

Das insgesamt schlechteste Messergebnis ergibt sich für das Projekt „NEH-Sortieranlage“ der „Gruppe A“ aus dem WS 2006 / 2007: Bei der sprachlichen Qualität fällt insbesondere die vergleichsweise schlechte Einhaltung der Formulierungsregeln (hoher Passiv-Anteil [39,80%], hoher Anteil an Nominalstil [38,78%]), das unvollständige Glossar (GF1-M2.1: 36,96% der Fachbegriffe tauchen im Glossar auf) sowie die hohe Widersprüchlichkeit innerhalb einer Anforderung (GF1-M4.1: 3,5% [Mittelwert über alle Anforderungen]) auf.

Inhaltlich ist die hohe (externe) Widersprüchlichkeit (GF2-M3.2: 14,29% widersprüchliche Anforderungen) und die geringe Prüfbarkeit von Anforderungsaspekten (GF2-M4.1: 63,70% prüfbare Anforderungsaspekte) zu bemängeln.

Bei den nicht-funktionalen Anforderungen sind keinerlei technische Anforderungen einer Kategorie zugeordnet (GN2.1-M1.2: 0% der technischen Anforderungen sind einer Kategorie zugeordnet), es existiert keine grafische Beschreibung der Benutzerschnittstelle, und nur wenige Qualitätsziele sind konkretisiert worden (GN2.3-M1.2: 22,22% konkretisierte Qualitätsziele).

Insgesamt lässt sich feststellen, dass die Projekte, die sich mit ihrer Gesamtpunktzahl in der Spitzengruppe befinden, aus dem WS 2007 / 2008 stammen. Insbesondere bei der inhaltlichen Qualität erreichen die Projekte aus dem WS 2007 / 2008 oft höhere Werte.

Betrachtet man die Qualitätsanforderungen, ergibt sich ein leichter Vorsprung für die Projekte aus dem WS 2006 / 2007: Das Projekt mit dem besten Ergebnis stammt aus diesem Semester. Ferner erreichen zwei Projekte aus dem WS 2006 / 2007 einen Wert über 6, während dies aus dem WS 2007 / 2008 nur für ein Projekt zutrifft. Dieses Ergebnis ist in sofern bemerkenswert, als alle Projekte aus dem WS 2007 / 2008 durch das Vorhandensein einer graphischen Beschreibung der Benutzerschnittstelle einen Vorteil gegenüber den meisten Projekten aus dem WS 2006 / 2007 haben. Jedoch kann dieser Vorteil andere Defizite bezüglich der Qualitätsanforderungen offenbar nicht ausgleichen.

4.3.4 Zusammenhang zwischen Messergebnis und Projekterfolg

Wie hängt die gemessene Qualität eines Anforderungsdokuments mit dem tatsächlichen Erfolg eines Projekts zusammen? Dieser Zusammenhang soll in diesem Kapitel untersucht werden. Hierzu stehen die Einschätzungen des ursprünglichen Kunden, der das Produkt haben wollte, zur Verfügung, inwieweit ein Projekt auch zu dem ursprünglich geplanten Zweck eingesetzt werden kann.

Beim Vergleich der Messergebnisse und der Einschätzungen der Kunden, ergeben sich zwei Schwellenwerte: Projekte, die ca. 44 Punkte oder mehr haben, werden voraussichtlich auch zu ihrem geplanten Zweck eingesetzt werden, wobei in Projekten, die ca. 40 Punkte oder weniger erreicht haben, häufiger Fehlentwicklungen sichtbar sind. Auch jene Projekte, die voraussichtlich nicht zu ihrem ursprünglichen Zweck eingesetzt werden können, liegen in ihren Messergebnissen alle unterhalb der 40 Punkte-Grenze.

Somit können zwei Hypothesen aufgestellt werden: Zum einen, dass ein Projekt, welches 44 oder mehr Punkte erreicht hat, ein höheres Potential hat, erfolgreich abgeschlossen zu werden und andererseits ein Messwert unterhalb von 40 Punkten ein Indikator für mögliche Probleme bis hin zur Unbrauchbarkeit eines Projekts sein kann.

Um diese Hypothesen zu bestätigen, ist der Datenbestand von 16 Projekten zu klein. Hierzu müssten noch weitere Untersuchungen angestellt werden.

5 Bewertung der Metriken

In diesem Kapitel wird eine Bewertung der Metriken und Heuristiken vorgenommen, um zu beurteilen, inwieweit sie zur Beurteilung der Qualität von Anforderungen geeignet sind. In [Lig02] werden folgende Kriterien genannt, um Metriken sinnvoll nutzen zu können:

1. **Einfachheit**
Interpretationsaufwand des Messergebnisses muss angemessen sein
2. **Eignung (Validität)**
Ausreichende Korrelation zwischen Metrik und zu messender Eigenschaft
3. **Stabilität**
Stabilität des Metrikwertes gegenüber Manipulationen untergeordneter Bedeutung
4. **Rechtzeitigkeit**
Messung muss rechtzeitig genug erfolgen, um noch korrigierend (in den Prozess) eingreifen zu können
5. **Analysierbarkeit**
Messwerte müssen in Relationen zueinander gesetzt werden können
6. **Reproduzierbarkeit**
Es müssen objektive Kriterien existieren. Subjektive Einflussnahme durch den Messenden muss ausgeschlossen werden

Folgende Auflistung gibt an, inwieweit die Metriken des Messplans die oben genannten Eigenschaften erfüllen:

- ✓ Einfachheit: Die Ergebnisse sind prozentualer Art oder absolute Zahlen
- ✓ Eignung (Validität): Metriken wurden konsequent nach der GQM-Methode erstellt
- ✓ Rechtzeitigkeit: Durch die Messung auf Basis des Anforderungsdokuments kann noch korrigierend in den Projektablauf eingegriffen werden
- ✓ Analysierbarkeit: Die Messwerte sind stets vergleichbar (Prozentuale Werte / absolute Zahlen)
- ✓ Stabilität: Gewährleistet durch prozentuale bzw. absolute Zahlen

Die folgenden Eigenschaften sind nur teilweise erfüllt:

- x Eignung (Validität): Formulierungsregeln geben nur Hinweise auf mögliche (inhaltliche) Defekte
- x Reproduzierbarkeit: Die Messergebnisse basieren teilweise auf subjektiven Kriterien und können von daher von Prüfer zu Prüfer variieren

Der potentielle Schwachpunkt der Metriken ist die Reproduzierbarkeit der Messergebnisse. Diese Schwäche wird im Folgenden näher untersucht werden.

5.1 Einschränkungen der Allgemeingültigkeit

Wie stark werden sich die Messergebnisse bei der Analyse eines Anforderungsdokuments je nach Prüfer voneinander unterscheiden? Um dieser Frage nachzugehen, wurde exemplarisch mit zwei Informatikstudenten (beide 7. Semester Bachelor-Studium) und dem Betreuer dieser Arbeit je eine Probemessung durchgeführt. Die Messergebnisse werden im Folgenden miteinander verglichen.

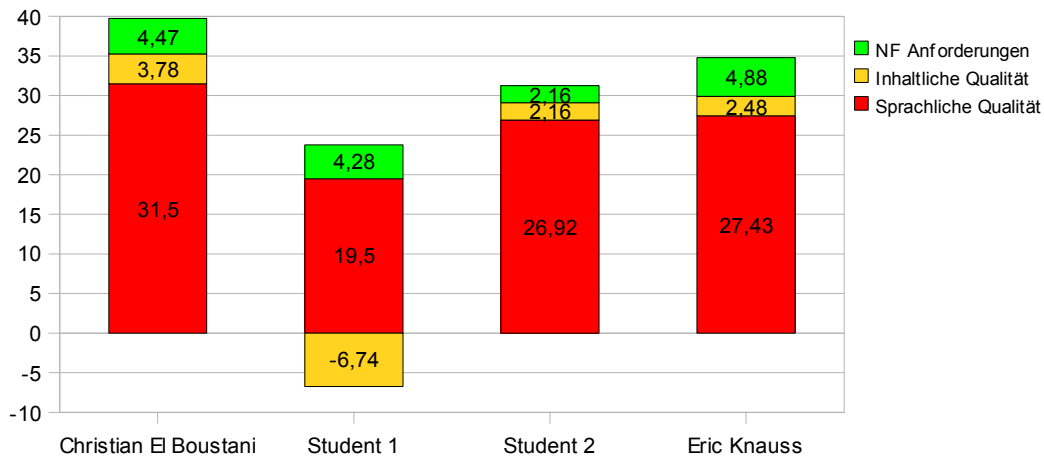


Abbildung 5.1: Vergleichswerte der Messung für Projekt "Fast View & Edit XML"

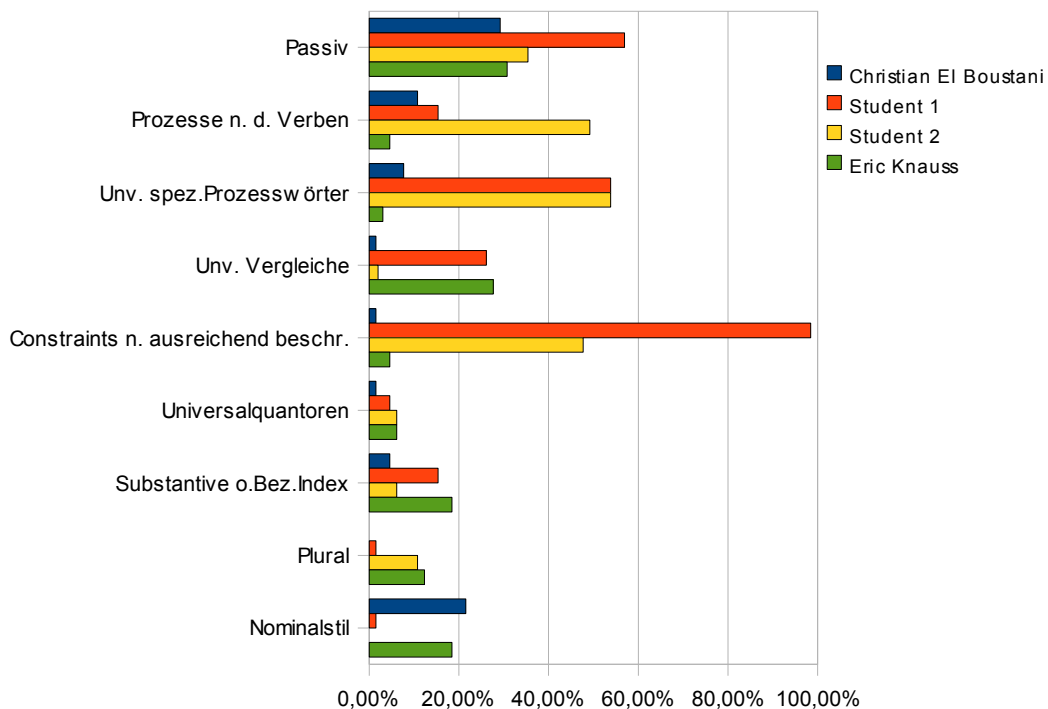


Abbildung 5.2: Vergleichswerte bei der Analyse der Formulierungsregeln

Bei dem Vergleich der Messergebnisse fällt auf, dass insbesondere bei der Analyse der sprachlichen Qualität – speziell bei der Prüfung auf Einhaltung der Formulierungsregeln – signifikante Unterschiede zwischen den einzelnen Prüfern zu beobachten sind. Im Folgenden werden die unterschiedlichen Heuristiken und Metriken isoliert betrachtet und untersucht, welche Faktoren zu den unterschiedlichen Ergebnissen geführt haben könnten.

Allgemein lässt sich jedoch festhalten, dass die Auswertung eines Projekts mit Hilfe des Fragebogen-Programms mit einer Dauer von 120 - 180 Minuten einen großen Aufwand darstellt. Durch diese Auswertungsdauer nähert sich das Verfahren zum einen kommunikativen Prüfverfahren wie z.B. Reviews und Inspektionen an, zum anderen ist zu befürchten, dass psychologische Faktoren wie sinkende Motivation und nachlassende Konzentration die Güte der Messergebnisse negativ beeinflussen könnte. Diese Effekte könnten sich vor allem bei sehr umfangreichen Anforderungsdokumenten einstellen. Auch Zeitdruck kann natürlich zu Fehlern führen und damit die Messergebnisse beeinflussen.

5.1.1 Heuristiken / Metriken zur sprachlichen Qualität

Prüfung, ob Anforderung Passiv enthält

Bei der Prüfung, ob eine Anforderung im Passiv geschrieben worden ist, sind die Kriterien, die die einzelnen Prüfer angelegt haben, sehr unterschiedlich: Teilweise wurde nur konsequent nach Konjugationen des Verbs „werden“ gesucht (was false-positives bei Futur-Formen zur Folge hatte wie z.B. „*Genau wie der bestehende XML Editor wird auch der FVE-XML Editor die bestehenden Schnittstellen des Eclipse Frameworks nutzen*“), andere Prüfer werteten Anforderungen, die keinen Akteur enthielten als Passiv – egal, ob sie grammatikalisch im Passiv formuliert waren oder nicht.

Es lässt sich festhalten, dass sowohl manche passive Formulierungen nicht entdeckt wurden, jedoch auch viele (sogar z.T. eindeutig im Aktiv geschriebene) Anforderungen fälschlicherweise als Befund markiert worden sind. Selbst die Auswertung innerhalb der Ergebnisse eines Prüfers war oft nicht konsequent.

Es liegt nahe, dass die Reproduzierbarkeit der Messergebnisse bezüglich der Passiv-Überprüfung offensichtlich erheblich von der linguistischen Kompetenz des Prüfers abhängt.

Prüfung, ob Prozesswörter nicht durch Vollverben ausgedrückt worden sind

Bei der Überprüfung, ob Prozesswörter durch Vollverben ausgedrückt sind, fällt auf, dass insbesondere die wenig konkrete Definition von „Prozesswort“ eine Fehlerquelle darstellt. Somit wurden teilweise Qualitätsaspekte als Prozesswörter bewertet (z.B. „*Das wesentliche Qualitätsziel des Projektes ist die Usability*“) aber auch Wörter, die den Projektablauf und nicht einen Ablauf bei der Bedienung der Software beschrieben, wie z.B. „*Die grundlegende Erweiterung (...) wird in dem Hinzufügen einer weiteren Ansicht (...) bestehen*“.

Prüfung auf unvollständig spezifizierte Prozesswörter

Bei der Prüfung auf unvollständig spezifizierte Prozesswörter ergibt sich ebenfalls das Problem, dass nicht eindeutig definiert ist, welche Wörter zu den Prozesswörtern zählen. Ferner ist es durchaus subjektiv, wann ein Prozesswort als „unvollständig“ anzusehen ist.

Prüfung auf unvollständige Vergleiche

Bei der Prüfung, ob eine Anforderung unvollständige Vergleiche enthielt, werteten die Prüfer unterschiedliche sprachliche Aspekte als Defekt: Nach [Rup06] sind auch Ausdrücke wie z.B. „effizient“ oder aber auch „effizienter“ als unvollständiger Vergleich zu werten, jedoch wurden solche Ausdrücke nur von manchen Prüfern als Defekt markiert (z.B. „*Dem Nutzer soll ein effizienteres Arbeiten mit XML-Dokumenten ermöglicht werden*“).

Prüfung, ob Constraints nicht ausreichend beschrieben worden sind

Bei der Prüfung, ob Constraints nicht ausreichend beschrieben worden sind, fällt insbesondere die hohe Anzahl an Defekten des Prüfers „Student 1“ auf: Nahezu jede Anforderung enthält seiner Wertung nach unvollständige Constraints. Diese Einschätzung ist jedoch problematisch, da nicht jede Anforderung tatsächlich auch einen Constraint enthält: Die Anforderung „*Die Usability ist dabei die zentrale Anforderung*“ legt beispielsweise nur die Priorität eines Qualitätsziels fest, enthält jedoch keine Einschränkung (Constraint), welche näher beschrieben werden sollte.

Prüfung, ob Anforderung Universalquantoren enthält

Bei der Prüfung, ob eine Anforderung Universalquantoren enthält, wurden teilweise auch Anforderungen als Defekt markiert, die nach [Rup06] keine Universalquantoren enthalten.

Prüfung, ob Anforderung Substantive ohne Bezugsindex enthält

Bei der Prüfung, ob eine Anforderung Substantive ohne einen Bezugsindex enthält, werteten die Prüfer teilweise auch den Namen der zu erstellenden Software („*Höchste Priorität besitzt für den Kunden die Entwicklung eines **Editors**, (...)*“) als Befund.

Prüfung, ob Anforderung Plural enthält, wo auch Singular richtig wäre

Bei der Prüfung, ob eine Anforderung Plural enthält, wo jedoch auch Singular richtig wäre, fällt der hohe Anteil an false-positives auf: Häufig wurde Anforderungen als defekt markiert, die offenkundig überhaupt kein Plural enthielten. Andererseits wurde oft jede Art von Plural als ein Defekt markiert, selbst wenn er grammatikalisch notwendig war, wie z.B. „*Um den Überblick über das Dokument zu erleichtern, sollen bekannte **Features** wie Code-Folding und Syntax-Highlighting implementiert werden*“. Es scheint hier abermals einen Zusammenhang zwischen der linguistischen Kompetenz der Prüfer und der Anzahl der gefundenen Defekte zu geben. Ferner muss auch gesagt werden, dass die Entscheidung, ob Singular auch richtig wäre, nicht immer einfach zu treffen ist: Im Zweifelsfall muss dies auch im Dialog mit dem Kunden abschließend geklärt werden. Dennoch kann ein Prüfer zunächst eine Einschätzung abgeben, die dann in kommunikativen Prüfverfahren weiter untersucht wird.

Prüfung, ob Anforderung Nominalstil enthält

Bei der Prüfung, ob eine Anforderung im Nominalstil geschrieben ist, ergab sich eine Überschneidung mit der Heuristik „Prozesse durch Vollverben ausgedrückt“: War ein Prozesswort nicht als Verb benutzt worden, wurde es durch das entsprechende Substantiv ausgedrückt (z.B. „(...) *soll das Produkt (...) sowohl zur **Erstellung** als auch zur **Bearbeitung** von XML-Dokumenten eingesetzt werden*“). Es stellte sich für die Prüfer die Frage, ob die hervorgehobenen Wörter aus dem obigen Beispiel als Befunde für die Heuristik „Prozesse nicht durch Vollverben“ oder aber als „Nominalstil“ - oder sogar beides - anzusehen sind. Ferner wurden teilweise auch Fachbegriffe (z.B. „*Syntax-Highlighting*“) als Nominalstil gewertet.

GF1-M2.1: Erläuterung von Fachbegriffen

Bei der Beurteilung, wie viele Fachbegriffe im Glossar des Anforderungsdokuments erklärt worden sind, fällt auf, dass unterschiedliche Auffassungen, welche Wörter bereits als Fachbegriff zählen, Einfluss auf die Messergebnisse hat: So wurden z.B. Qualitätsaspekte teilweise als Fachbegriffe gezählt, teilweise nicht. Die Abweichungen in den Einschätzungen, wie viele Fachbegriffe eine Anforderung enthält, betragen oft bis zu + 2 / - 2 Fachbegriffe.

Nur geringe Abweichungen gab es bei der Einschätzung, welche der identifizierten Fachbegriffe im Glossar auftauchten. Hier war z.B. entscheidend, ob ein Prüfer etwa Synonyme als erläuterten (z.B. „*View*“ im Glossar, „*Sicht*“ im Text Anforderung) Fachbegriff hat zählen lassen oder nicht.

GF1-M2.2: Anzahl der sprachlichen Fehler, die die Verstehbarkeit beeinflussen

Keine Abweichungen zwischen den Prüfern.¹

GF1-M3.1: Identifizierbarkeit von Anforderungen

Bei der Beurteilung der Identifizierbarkeit von Anforderung – also konkret, ob ein Satz im Anforderungsdokument mehr als eine Anforderung enthält – sind die Ergebnisse relativ konsistent. Schwachpunkt ist hier die Einschätzung, ob eine Anforderung aus mehreren Aspekten besteht oder aber, ob es sich bereits um mehrere verschiedene Anforderungen handelt.

GF1-M3.2: Anzahl mehrdeutiger Begriffe innerhalb einer Anforderung

Bei der Beurteilung, wie viele mehrdeutige Begriffe eine Anforderung enthält, fällt auf, dass in vielen Fällen die Messergebnisse von drei Prüfern identisch sind, während das Messergebnis des vierten Prüfers abweicht. Die Abweichung beträgt max. +2 mehrdeutige Begriffe pro Anforderung.

Bei dieser Metrik spielt der Wissensstand des Prüfers eine Rolle: So wäre etwa der Begriff

¹ Hypothese, dass sich sprachliche Fehler auf Verstehbarkeit auswirken, hat sich nicht bestätigt – dadurch keine Befunde bei dieser Metrik

„Browser“ für einen Prüfer, der nur einen einzigen Browser kennt, ein (vergleichsweise) eindeutiger Begriff, ein anderer Prüfer jedoch würde einen solchen Begriff als mehrdeutig markieren.

GF1-M3.3: Mehrdeutigkeit durch sprachliche Fehler

Keine Abweichungen zwischen den Prüfern.¹

GF1-M4.1: Widersprüchlichkeit einer Anforderung

Bei der Beurteilung, wie viele widersprüchliche Aspekte ein Anforderung enthält, wurden teilweise auch allgemein schwer verständliche Anforderungen als widersprüchlich klassifiziert, wobei die Beurteilung, ob ein Widerspruch vorliegt, auch allgemein subjektiv ist. Konsistent bis auf geringe Abweichungen waren jedoch die Einschätzungen, wie viele Aspekte eine Anforderung enthält.

5.1.2 Metriken zur inhaltlichen Qualität

GF2-M3.1: Redundanz

Bei der Bewertung ob eine Anforderung redundant zu einer oder mehreren Anforderungen ist, fiel auf, dass der Grad der Übereinstimmung zwischen den Anforderung, der zur Feststellung einer Redundanz notwendig war, von Prüfer zu Prüfer variierte: Während mancher Prüfer schon bei kleinsten inhaltlichen Überschneidungen eine Redundanz feststellte, werteten andere Prüfer nur Anforderungen, die vollständig identisch waren, als redundant.

GF2-M3.2: (Externe) Widersprüchlichkeit einer Anforderung

Bei der Bewertung, welche Anforderungen sich gegenseitig widersprechen, ergeben sich signifikante Abweichungen zwischen den Prüfern.

Teilweise wurde etwa bei der Beurteilung der Priorität einer Anforderung nicht zwischen funktionaler und nicht-funktionaler Anforderung unterschieden und somit ein Widerspruch erkannt wie zwischen folgenden Anforderungen *„Die wesentliche funktionale Anforderung stellt die neue Sicht bzw. Seite innerhalb des Editors dar“* und *„Das wesentliche Qualitätsziel des Projektes ist die Usability“*, andererseits wurden z.B. die Inhalte der folgenden Anforderungen *„Die Umsetzung des Produkts soll als eine Erweiterung des bestehenden Eclipse XML Editors (...) implementiert werden“* und *„Der Kunde favorisiert die Realisierung des Produkts als Plug-In für die integrierte Entwicklungsumgebung 'Eclipse' (...)“* nicht von allen Prüfern als widersprüchlich gewertet.

Die Einschätzung welche Anforderungen sich widersprechen, scheint ein im besonderen Maße subjektives Maß zu sein.

¹ Hypothese, dass sich sprachliche Fehler auf die Eindeutigkeit auswirken, hat sich nicht bestätigt – dadurch keine Befunde bei dieser Metrik

GF2-M4.1: Prüfbarkeit einer Anforderung

Bei der Überprüfung, wie viele Anforderungsaspekte prüfbar sind, ergaben sich relativ konsistente Ergebnisse. Geringe Abweichungen ergaben sich bei der Beurteilung, wie viele Aspekte eine Anforderung enthält, was die Abweichungen in den Gesamtergebnissen erklärt.

GF2-M8.1: Bewertbarkeit

Bei der Überprüfung, wie viele Anforderungsaspekte auf ihre Wichtigkeit bewertet sind, ergaben sich relativ konsistente Ergebnisse.

GF2-M9.1: Klassifizierbarkeit

Keine Abweichungen zwischen den Prüfern, da rechtliche Verbindlichkeiten auf Grund von fehlenden Verträgen zwischen Kunde und Entwickler keine Rolle in dem analysierten Projekt spielen.

5.1.3 Metriken für nicht-funktionale Anforderungen

GN2.1-M1.2: Zuordnung zu einer Kategorie

Bis auf einen Fall, in dem ein Prüfer die Metrik falsch verstanden hatte, gab es keine Abweichungen zwischen den Prüfern.

GN2.1-M1.3: Anzahl Kategorien von technischen Anforderungen im Dokument

Bis auf einen Fall, in dem ein Prüfer die Metrik falsch verstanden hatte, gab es keine Abweichungen zwischen den Prüfern.

GN2.2-M1.1: Vorhandensein einer visuellen Beschreibung der Benutzerschnittstelle

Keine Abweichungen zwischen den Prüfern.

GN2.2-M1.2: Vorhandensein eines Benutzerprofils

Keine Abweichungen zwischen den Prüfern.

GN2.2-M1.3: Benennung von Eingabegeräten

Keine Abweichungen zwischen den Prüfern.

GN2.3-M1.1: Vorhandensein eines Qualitätsmodells

Keine Abweichungen zwischen den Prüfern.

GN2.3-M1.2: Nennung von Qualitätszielen

Bei der Überprüfung, wie viele der allgemeinen Qualitätsziele konkretisiert worden sind, fällt auf, dass insbesondere die beiden studentischen Prüfer deutlich weniger Qualitätsaspekte als konkret gewertet haben als die anderen Prüfer.

Schwachpunkt dieser Metrik ist die Einschätzung, ab wann ein Qualitätsziel als konkretisiert anzusehen ist.

GN2.3-M1.3: Metriken für Qualitätsziele

Keine Abweichungen zwischen den Prüfern.

5.2 Optimierungsmöglichkeiten

Wie lassen sich nun die unter 5.1 beschriebenen Effekte bzgl. der Reproduzierbarkeit der Metriken reduzieren?

Tabelle 5.2.1 gibt einen Überblick über die Störfaktoren und mögliche Optimierungsmöglichkeiten bzgl. der Metriken und Heuristiken.

Metrik / Heuristik	Störfaktor(en)	Optimierungsmöglichkeiten
Anforderung enthält kein Passiv?	Linguistische Kompetenz der Prüfer	<ul style="list-style-type: none"> Grammatikkenntnisse der Prüfer schulen Schlüsselwortanalyse durchführen (z.B. nur „werden“-Formen suchen)
Prozesse durch Verben?	<ul style="list-style-type: none"> Unterschiedliche Wertung der Wörter als Prozesswörter Überlappung mit Heuristik „Kein Nominalstil?“ 	<ul style="list-style-type: none"> Schlüsselwortliste mit Prozesswörter vorgeben Festlegung, dass Qualitätsziele nicht mitgezählt werden
Keine unvollständig spez. Prozesswörter?	Unterschiedliche Wertung der Wörter als Prozesswörter	Schlüsselwortliste mit Prozesswörtern vorgeben
Keine unvollständigen Vergleiche?	Unterschiedliche Wertung von sprachlichen Ausdrücken als Vergleich	<ul style="list-style-type: none"> Beispielsätze den Prüfern zeigen Regel vorgegeben: Bezugsindex muss vorhanden sein, damit es sich um einen Vergleich handelt [Rup06]
Constraints ausreichend beschrieben?	Jede Art von allgemeiner Beschreibung als Defekt gewertet	Schlüsselwortliste vorgeben („Signalwörter“ s. [Rup06])
Kein Einsatz von Universalquantoren?	Unterschiedliche Wertung der Wörter als Universalquantoren	Schlüsselwortliste mit Universalquantoren vorgeben
Keine Substantive ohne Bezugsindex?	Wertung des Namens der Software als Defekt („das Programm“ , „der Editor“)	Blacklist vorgeben: Folgende Begriffe nicht als Defekt werden
Kein Plural, wo auch Singular richtig?	<ul style="list-style-type: none"> Linguistische Kompetenz der Prüfer Jede Art von Plural wurde als Defekt erkannt 	Grammatikkenntnisse der Prüfer schulen
Kein Nominalstil?	Überlappung mit Heuristik „Kein Nominalstil?“	Blacklist mit Wörtern, die nicht mitgezählt werden dürfen z.B. Prozesswörter (s. Prozesswortliste) oder Qualitätsziele
GF1-M2.1: Erläuterung von Fachbegriffen	Fachkenntnisse des Prüfers: Welche Begriffe sind für ihn selbstverständlich?	Verschiedene Prüfer mit unterschiedlichen Fachkenntnissen prüfen lassen
GF1-M2.2: Anzahl der sprachlichen Fehler, die die Verstehbarkeit beeinflussen	-	-
GF1-M3.1: Identifizierbarkeit von Anforderungen	Besteht Satz aus einer Anforderung mit mehreren Aspekten oder aus mehreren Anforderungen?	-

Metrik / Heuristik	Störfaktor(en)	Optimierungsmöglichkeiten
GF1-M3.2: Anzahl mehrdeutiger Begriffe innerhalb einer Anforderung	Fachkenntnisse des Prüfers: Wie viele Bedeutungen zu einem bestimmten Begriff kennt er?	Prüfer mit ähnlichen Kenntnissen für die Prüfung wählen
GF1-M3.3: Mehrdeutigkeit durch sprachliche Fehler	Keine Störfaktoren erkennbar	-
GF1-M4.1: Widersprüchlichkeit einer Anforderung	Allgemein subjektiv	-
GF2-M3.1: Redundanz	Größe der Schnittmenge zwischen zwei Anforderungen – ab wann gilt Redundanz?	Festlegung: Ab X gleichen Aspekten gilt Redundanz oder auch festlegen: Erläuterungen und Begründungen zu einer Anforderung gelten nicht als redundante Anforderungen etc.
GF2-M3.2: (Externe) Widersprüchlichkeit einer Anforderung	Allgemein subjektiv	-
GF2-M4.1: Prüfbarkeit einer Anforderung	Anzahl der Anforderungsaspekte unterschiedlich gezählt	-
GF2-M8.1: Bewertbarkeit	Keine Störfaktoren erkennbar	-
GF2-M9.1: Klassifizierbarkeit	Keine Störfaktoren erkennbar	-
GN2.1-M1.2: Zuordnung zu einer Kategorie	Keine Störfaktoren erkennbar	-
GN2.1-M1.3: Anzahl Kategorien von technischen Anforderungen im Dokument	Keine Störfaktoren erkennbar	-
GN2.2-M1.1: Vorhandensein einer visuellen Beschreibung der Benutzerschnittstelle	Keine Störfaktoren erkennbar	-
GN2.2-M1.2: Vorhandensein eines Benutzerprofils	Keine Störfaktoren erkennbar	-
GN2.2-M1.3: Benennung von Eingabegeräten	Keine Störfaktoren erkennbar	-
GN2.3-M1.1: Vorhandensein eines Qualitätsmodells	Keine Störfaktoren erkennbar	-
GN2.3-M1.2: Nennung von Qualitätszielen	Ab wann ist ein Qualitätsziel konkretisiert?	Festlegung: Kurze Erklärung, was Qualitätsziel in aktuellen Projekt bedeutet, reicht aus
GN2.3-M1.3: Metriken für Qualitätsziele	Keine Störfaktoren erkennbar	-

Tabelle 5.2.1: Optimierungsmöglichkeiten der Metriken / Heuristiken für die Messung

6 Fazit

In dieser Arbeit sollten Anforderungen von verschiedenen Softwareprojekten qualitativ und quantitativ miteinander verglichen werden. Hierzu wurde mittels der GQM-Methode ein Messplan erstellt und auf die Anforderungsdokumente von Projekten angewendet. Die Ergebnisse wurden zusammengefasst, um zu einer Kennzahl für die Gesamtqualität eines Anforderungsdokuments zu kommen und die Projekte untereinander vergleichen zu können.

Bezüglich der Qualität der Anforderungsdokumente lässt sich sagen, dass sie über die Zeit vom WS 2006 / 2007 bis zum WS 2007 / 2008, zugenommen hat. Die Durchschnittspunktzahl bei Projekten aus dem WS 2007 / 2008 ist höher als bei den Projekten aus dem WS 2006 / 2007. Ferner stammt das beste Anforderungsdokument aus dem WS 2007 / 2008, während das schlechteste Dokument aus dem WS 2006 / 2007 stammt.

Inwiefern ein gutes Messergebnis bzgl. eines Anforderungsdokuments mit dem Erfolg eines Projekts zusammenhängt, kann nicht abschließend geklärt werden, wobei die Vermutung nahe liegt, dass ein Projekt, dessen Anforderungsdokument einen gutes Messergebnis erhalten hat, ein höheres Potential hat, erfolgreich abgeschlossen zu werden. Allgemein ist der Zusammenhang zwischen Anforderungsqualität und Qualität des Produkts eines Projekts schwer, sicher nachzuweisen [Rec+06-2].

Die Einschätzung, dass es mittels der Anforderungsschablone (Abbildung 4.1) möglich ist, „hochwertige Anforderungen zu formulieren“ [Hor07], kann mit den Ergebnissen dieser Arbeit bestätigt werden. Auffallend war, dass jede Art von vorgegebenem Schema bezüglich der Formulierungen einer Anforderung signifikant ihre Qualität erhöhte. So waren Anforderungen aus Use-Cases, die durch die tabellarische Struktur beschränkt waren, qualitativ besser als Anforderungen im Fließtext des Anforderungsdokuments. Somit wird deutlich, dass bestimmte strukturelle Vorgaben innerhalb des Dokuments sich positiv auf die Qualität der Anforderungen auswirken können. Die Auflage, ein Qualitätsmodell anlegen zu müssen bzw. die Struktur des Qualitätsmodells bereits ins Dokumenten-Template aufzunehmen, könnte sich als sinnvolle Maßnahme herausstellen, um die Qualität der Qualitätsanforderungen zu steigern.

Die in dieser Arbeit aufgestellten Metriken geben Qualitätsbeauftragten ein Mittel an die Hand, um Stellen im Anforderungsdokument ausfindig zu machen, die in späteren Projektphasen zu Problemen führen könnten: Wenn z.B. zu einer Anforderung keine prüfbareren Erfüllbarkeitskriterien angegeben sind, so ist die Gefahr gegeben, dass Entwickler und Kunde unterschiedliche Auffassungen darüber haben, wann die Anforderung erfüllt ist. Derartige Mängel können mittels des Messplans dieser Arbeit aufgedeckt werden und gezielt im Dialog mit dem Kunden geklärt werden.

Es gibt jedoch Grenzen der Messbarkeit: So lassen sich wesentliche inhaltliche Aspekte wie die Vollständigkeit oder die Aktualität einer Anforderung nicht isoliert anhand des Anforderungsdokuments messen. Zwar können sprachliche Befunde Indikatoren für *mögliche* Defekte liefern, aber sie geben keine Garantie, dass tatsächlich auch ein inhaltlicher Defekt vorliegt. Jedoch waren auch die hier angewendeten Prüfverfahren nicht frei von subjektiven Einflüssen: Ob etwa eine Anforderung verständlich ist oder nicht, oder ob ein Wort bereits als Fachwort zu bewerten ist oder noch zum „alltäglichen Vokabular“ gehört,

variiert stark von Stakeholder zur Stakeholder [Lig02]. Letztendlich muss auch zur abschließenden Klärung dieser Aspekte der direkte Dialog mit den jeweiligen Stakeholdern (also v.a. den Kunden) gesucht werden. Somit scheint es also zu den zeitaufwändigen kommunikativen Prüfverfahren keine Alternativen zu geben. Der in dieser Arbeit aufgestellte Messplan kann also die inhaltliche Validierung von Anforderungen nicht ersetzen, stellt jedoch ein sinnvolle vorbereitende Maßnahme für kommunikative Prüfverfahren dar.

7 Ausblick

Optimierungspotential ergibt sich bei der Effizienz der hier angewendeten Metriken: So könnte etwa eine automatisierte Schlüsselwortanalyse die Messung einiger Aspekte beschleunigen. Interessant wäre in diesem Zusammenhang auch, welche Metriken sich überhaupt automatisieren lassen und welche Auswirkungen dies auf die Messergebnisse hätte.

Wie schon im Fazit geschrieben, erfordert die Anwendung bestimmter Metriken den Kontakt zu Stakeholdern. Jene Metriken, die nicht in den Messplan eingeflossen sind, wie z.B. Metriken zur Vollständigkeit oder zur Aktualität einer Anforderung, könnten im Rahmen zukünftiger Projekte angewendet und anhand der Messergebnisse die zu Grunde liegenden Hypothesen verifiziert werden.

Aufschlussreich wäre auch, welche Auswirkungen ein nach den Messungen „schlechtes“ Anforderungsdokument auf den Projektverlauf – insbesondere auf den Zeitplan – hätte. Abschließend wäre auch interessant zu ermitteln, welche Methoden der Anforderungserhebung ein Projektteam anwendet, wie viel Zeit in das Requirement-Engineering investiert wird und wie sich diese Faktoren auf die Qualität der Anforderungen auswirken. Untersuchungen zeigen, dass eine Aufwendung von weniger als 5% für das Requirements-Engineering zu einer signifikanten Überschreitung des Gesamtbudgets führen (80% - 200%), jedoch eine Aufwendung von bis zu 14 % nur zu einer Budgetüberschreitung von ca. 25% - 60% führen [Rec+06-2]. Es gilt zu untersuchen, ob diese Zahlen bestätigt werden können.

8 Literaturverzeichnis

- [Bro08]: Brockhaus AG, Der Brockhaus multimedial premium 2008, 2008
- [Sch07]: Kurt Schneider, Abenteuer Software Qualität. dpunkt.Verlag, 2007, S. 31 - 37, 51 - 58, 68 – 78
- [Rup06]: Chris Rupp, Requirements-Engineering und Management, 4. Auflage. Hanser, 2007, S. 26 - 30, 140 - 166, 227 - 234, 255 – 275
- [Rec+06]: Recknagel, Rupp, Messbare Qualität in Anforderungsdokumenten, 2006, [http://www.sophist.de/DownloadDB.nsf/0/e7f9468eee05bddfc125711800459ae9/\\$FILE/Messbare%20Qualit%C3%A4t%20in%20Anforderungsdokumenten.pdf](http://www.sophist.de/DownloadDB.nsf/0/e7f9468eee05bddfc125711800459ae9/$FILE/Messbare%20Qualit%C3%A4t%20in%20Anforderungsdokumenten.pdf)
- [Sch-AuE07]: Kurt Schneider, Vorlesung: Anforderungen und Entwurf SS2007, 2007
- [Sol+99]: Rini van Solingen, Egon Berghout, The Goal/Question/Metric Method. McGraw-Hill Publishing Company, 1999, S. 53 – 58
- [Hor07]: Nicko Horst, Quantitative und qualitative Analyse von Anforderungen in Software-Projekten, Leibniz Universität Hannover, FG Software Engineering, Studienarbeit, 2007
- [Dav05]: Alan M. Davis, Just Enough Requirements Management. Dorset House Publishing, 2005, S. 190 – 192
- [Lig02]: Peter Liggesmeyer, Software-Qualität. Spektrum, 2002, S. 214 – 215
- [Rec+06-2]: Matthias Recknagel, Chris Rupp, Metriken für Anforderungen - Wie gut sind Ihre Anforderungen wirklich?, 2006, http://www.itq.ch/pdf/sqm/sqm_rupp_recknagel_prmetrics&requirement.pdf