

CMM und die Methode der schriftlichen Befragung

Seminararbeit über empirische Forschungsmethoden in der Softwaretechnik

Oliver-Nils Pütz

12. April 2004

Zusammenfassung

Thema dieser Arbeit ist die Auseinandersetzung mit Methoden der empirischen Softwaretechnik und im speziellen mit dem Werkzeug der schriftlichen Befragung, wie sie von Goldenson und Herbsleb bei einer Studie [3] über CMM¹ verwendet wurde. Als Grundlage dienen hauptsächlich zwei von Goldenson, Herbsleb und anderen veröffentlichte Artikel [1][2], in denen die Ergebnisse einer explorativen Studie zu den Auswirkungen von CMM-basierten Prozessverbesserungen vorgestellt werden. Nach einer Übersicht über CMM und die Methode der schriftlichen Befragung werden die zwei Artikel inhaltlich vorgestellt. Anschließend wird die behandelte Studie genauer betrachtet und stellenweise erörtert. Abschließend werde ich einige persönliche Bemerkungen zur Studie machen.

¹Capability Maturity Model

Inhaltsverzeichnis

1 Grundlagen	3
1.1 CMM - Capability Maturity Model	3
1.1.1 Das Stufenmodell des CMM	3
1.1.2 Das Rahmenwerk zum CMM	4
1.1.3 Kritik am CMM	5
1.2 Schriftliche Befragungen	5
1.2.1 Vorbereitung	5
1.2.2 Der Fragenkatalog	6
2 Kurze Darstellung der Artikel	7
2.1 "A Systematic Survey of CMM Experiences and Results" [1] . . .	7
2.2 "Software Quality and the Capability Maturity Model" [2]	7
3 Diskussion der Umfrage von Goldenson und Herbsleb	8
3.1 Zielsetzung und grundlegende Fragestellung der Umfrage	8
3.2 Das Sample – Die Auswahl der Umfrageteilnehmer	9
3.3 Der Fragenkatalog	10
3.4 Durchführung der Befragung	12
3.5 Auswertung der Ergebnisse und Schlußfolgerungen	12
4 Bewertung und Ausblick	13

1 Grundlagen

1.1 CMM - Capability Maturity Model

Das CMM (Capability Maturity Model; dt.: Prozessreifemodell)² ist ein Modell zur Beschreibung einer Organisation bezüglich ihrer Leistungsfähigkeit und Effektivität bei der Softwareentwicklung.

Das CMM wurde 1991 im Rahmen einer Auftragsarbeit für das Department of Defense (DoD) vom Software Engineering Institute (SEI) unter der Leitung von Watts Humphrey entworfen. Das DoD sah für sich die Notwendigkeit, die potentiellen Auftragnehmer bei Ausschreibungen besser beurteilen zu können, um den bestmöglichen und nicht nur den günstigsten Auftragnehmer für ein Projekt auswählen zu können. Scheitern eines Software-Projektes und unzuverlässige Zusagen sollten bereits im Vorfeld ausgeschlossen werden. In diesem Zusammenhang wurde das Capability Maturity Modell entwickelt. Das CMM sollte die Skala sein, in der potentielle Auftragnehmer durch ein "SEI Capability Evaluation"-Verfahren (SCE) eine Bewertung erhielten. Im Laufe der Jahre entwickelte sich im Zusammenhang mit dem CMM ein Rahmenwerk³ zur Verbesserung und Optimierung des Softwareprozesses. Das Rahmenwerk versucht, Methoden und Prozesse festzulegen und ein gewisses Maß an Anleitung bei der Arbeit an der Prozessverbesserung zu geben. Für eine ausführlichere Einführung zum Capability Maturity Model empfehle ich [4]. Es gibt eine gute Übersicht und Erklärung zu den Stufen des CMM sowie dem darauf basierenden Rahmenwerk.

1.1.1 Das Stufenmodell des CMM

Das CMM ist ein deskriptives Modell. Das Modell gibt fünf Stufen in der Leistungsfähigkeit und Ausprägung verschiedener Aspekte des Softwareprozesses an. Jede Stufe wird durch den Zustand des mehr oder weniger vorhandenen Softwareprozesses definiert.

Stufe 1 Initialer Prozeß: Auf dieser Stufe ist der Entwicklungsprozeß unorganisiert und chaotisch. Das Gelingen eines Projektes hängt ausschließlich vom Einsatz einzelner Personen und dem Zufall ab. In der Regel ist ein Erfolg nicht wiederholbar. Zeit- und Kostenplanung ist, wenn überhaupt möglich, nicht einzuhalten. Es existiert keine Qualitätssicherung oder realistische Fortschritts-Überwachung.

Stufe 2 Wiederholbarer Prozeß: Auf dieser Stufe existiert bereits ein Projektmanagement mit Zeit- und Kostenplanung. Erfolge bei Software-Projekten sind wiederholbar hängen jedoch noch vom Einsatz und Wissen einzelner Personen ab.

Stufe 3 Definierter Prozeß: Auf dieser Stufe kommt zum Projektmanagement auch ein definierter Softwareprozeß. Dazu gehören Regeln zur Anfertigung von

²Auch wenn dies ein englischer Begriff ist, für den sich eine deutsche Entsprechung eingebürgert hat, werde ich aus Bequemlichkeit im Folgenden die Abkürzung CMM verwenden.

³Rahmenwerk ist vielleicht ein wenig übertrieben ausgedrückt. Im wesentlichen beziehe ich mich auf Verfahren oder auch Prozesse, in denen der Reifegrad eines Prozesses festgestellt wird, sowie häufig auftretende Verfahrensweisen bei der Einführung von Veränderungen und sogar die Art der Veränderungen des Softwareprozesses.

Dokumenten und festgelegten Abläufen und Methoden. Das Management überwacht und steuert den Fortschritt eines Software-Projektes. Es werden Schulungen durchgeführt, um das Wissen in der Organisation zu verbessern und dadurch auch Unabhängigkeit von Einzelpersonen zu erreichen. Kosten- und Zeitplanung gelingen sehr gut und zuverlässig, jedoch ist die Qualitätssicherung noch nicht sehr stabil.

Stufe 4 Gesteuerter Prozeß: Zusätzlich zu den Eigenschaften der Stufe 3 gelingt hier auch gute statistische Kontrolle über die Qualität der Softwareprodukte. Der Softwareprozeß wird durch Erhebung von Meßwerten überwacht und in gewissem Maße auch gesteuert.

Stufe 5 Optimierender Prozeß: Auf Stufe 5 kommt zu den Eigenschaften und Charakteristika der Stufe 4 letztendlich die stetige rückgekoppelte Steuerung der Prozessoptimierung. Methoden, Werkzeuge als auch der Softwareprozeß selbst werden überwacht, vermessen und aufgrund der Meßwerte verbessert und verfeinert. Dazu werden Metriken, Prozessüberwachung und instrumentierte Prozessumgebungen herangezogen. Diese Stufe bildet die Spitze des CMM, auf der der Softwareprozeß überwacht und die Prozessautomatisierung und -Verbesserung stetig optimiert wird. Der Softwareprozeß kann an das Projekt angepaßt werden.

1.1.2 Das Rahmenwerk zum CMM

Im Laufe der Jahre haben sich im Umfeld des CMM einige Werkzeug entwickelt, die die Prozessverbesserung zum Ziel haben. Diese Methoden sind klar vom CMM selbst zu trennen, das lediglich ein beschreibendes Zustandsmodell von Softwareorganisationen ist.

Das erste dieser Werkzeug ist das "Software Process Assessment"-Verfahren (SPA). Es ähnelt dem SCE (ursprünglich vom DoD in Auftrag gegeben) und ist ein Verfahren zur Selbstbewertung nach dem CMM. Während einer mehrtägigen Untersuchung, bei der Interviews, schriftliche Befragungen und Beobachtung des Software-Prozesses zum Einsatz kommen, wird der Reifegrad im CMM festgestellt. Desweiteren werden die Stärken und Schwächen der Organisation identifiziert sowie ein Aktionsplan aufgestellt, der angibt, welches die nächsten Schritte zur Prozessentwicklung sind. Der Zweck des SPA war es, Organisationen einen Eindruck zu vermitteln, was bei einem SCE zu erwarten ist. Im Laufe der Zeit diente es aber auch mehr und mehr dem Zweck eine Strategie-Grundlage für "*Software Process Improvement*" Ambitionen (SPI; Prozessentwicklung oder auch Prozessverbesserung) zu schaffen. Dabei werden Schlüsselbereiche zur Entwicklung festgelegt (Key Process Areas, KPA), die als nächstes zu bearbeiten sind, um die Prozessverbesserung durchzuführen.

Eine weiteres Werkzeug die Verbesserung des Softwareprozesses zu unterstützen, ist die Gründung einer Software Engineering Process Group (SEPG). Mitglieder dieser Gruppe sind respektierte Mitarbeiter aus dem Umfeld der Softwareentwicklung. Diese haben die Aufgabe, die Entwicklung der Prozessreife am Laufen zu halten, Neuerungen bzw. Veränderungen anzuregen und Unterstützung bei der Umsetzung zu geben.

1.1.3 Kritik am CMM

Mir sind nicht alle Kritikpunkte zum CMM, die zum Zeitpunkt der Studie von Goldenson und Herbsleb existierten, bekannt. Einige werden auch in der von ihnen angegebenen Literatur ausführlicher behandelt. Ich werde hier nur einige dieser Kritikpunkte auführen.

Das CMM sei zu unpersönlich. Die Personen, die an der Softwareentwicklung beteiligt sind, werden zu sehr als austauschbare Werkzeuge betrachtet. Das CMM unterstütze durch die fehlende Berücksichtigung der Personen diese Haltung.

Auf kleinere Organisationen sei das CMM nur schlecht anwendbar. Es sind Einschränkungen in den vorhandenen Ressourcen, wie Personal und Technologien vorhanden, die keine Berücksichtigung im CMM finden. Die Methoden im Umfeld des CMM sind nur schlecht auf kleinere Organisationen anwendbar.

Das CMM und die darauf basierenden Methoden seien zu bürokratisch. Die Festlegung des Prozesses und die Einhaltung vorgegebener Arbeitsschritte und Regeln behindere die eigentliche Arbeit der Softwareherstellung. Das Anfertigen von Dokumenten und die Erfassung von Daten während des Prozesses beeinflusse die Leistungsfähigkeit negativ.

Die Werkzeuge zur Prozessverbesserung zeigen einige unprofessionelle Mängel. In [7] wird die Art der statistischen Auswertung der Ergebnisse eines SPA kritisiert. Im selben Artikel wird auch auf die fehlende Berücksichtigung der Effizienz beim Einsatz vorhandener Technologien hingewiesen.

Die größte Kritik und gleichzeitig auch der größte Schwachpunkt des CMM und der darauf basierender Methoden und Werkzeuge ist das Fehlen von Beweisen. Es existierten 1994 (der Zeitpunkt der Studie auf die ich mich beziehe) keine Beweise für die Gültigkeit des Modells und die Effektivität der Methoden und Werkzeuge. Aber es sind Hinweise in Form von Einzelfall-Studien vorhanden, die allesamt sehr positiv für das CMM ausfallen.

1.2 Schriftliche Befragungen

Herbsleb und Goldenson führten eine schriftliche Befragung durch, bei der der Fragebogen den Teilnehmern der Umfrage postalisch zugestellt wurde. Aus diesem Grund beziehe ich die folgenden Ausführungen hauptsächlich auf diese Form der Befragung.

1.2.1 Vorbereitung

Zu Beginn einer wissenschaftlich Studie sollte die genaue Definition der Ziele und Formulierung von Hypothesen oder Fragestellungen festgelegt werden. Zusammenhänge und Implikationen sollten soweit wie möglich bereits bewußt sein und berücksichtigt werden. Dazu gehört auch die Kenntnis über die Einbettung der Studie in den Kontext bereits vorhandener Studien, die mit dem Gegenstand der Untersuchung zusammenhängen.

Zur Vorbereitung einer Umfrage gehört die Auswahl geeigneter und repräsentativer Teilnehmer und die Erstellung eines Fragenkataloges, der darauf ausgerichtet ist, die gegebene Fragestellung zu beantworten bzw. die Hypothesen zu stützen oder zu verwerfen.

Die Auswahl der geeigneten Teilnehmer der Befragung ist wichtig, um negative Effekte auf die Repräsentativität der Ergebnisse, zu vermeiden. Es muß bekannt sein, was Repräsentativität bei der zugrundeliegenden Zielsetzung überhaupt bedeutet und wie sie durch Auswahl der Teilnehmer beeinflußt wird.

1.2.2 Der Fragenkatalog

Die wichtigsten Punkte bei der Gestaltung eines Fragenkataloges sind, neben dem thematischen Bezug zur Studie, das Vertrauen, das Interesse und die Ehrlichkeit des Teilnehmers. Von diesen Punkten hängt das Ergebnis der Befragung ganz besonders ab. Insgesamt sollte der Katalog übersichtlich und kurz sein, um die Geduld der Befragten nicht zu strapazieren. Er sollte einen Spannungsbogen aufbauen, bei dem am Anfang einfache und unverfängliche Fragen stehen, die das Vertrauen des Teilnehmers aufbauen und sein Interesse wecken sollen; insgesamt also seine Bereitschaft zur Teilnahme. In der Mitte folgen die für die Studie wichtigen Fragen. Nachdem zu Beginn das Interesse geweckt wurde, haben die Fragen in der Mitte des Kataloges die Aufmerksamkeit des Teilnehmers. Die Wahrscheinlichkeit für eine ehrliche und brauchbare Antwort ist hier am größten. Am Ende des Kataloges stehen Fragen zur demographischen Auswertung, deren Beantwortung vom Befragten keine besondere Aufmerksamkeit mehr verlangen.

Die Formulierung von Fragen und deren Antwortenmenge lassen sich zunächst unterscheiden zwischen *offen* und *geschlossen*. Offene Fragen geben die Möglichkeit, die Antwort selbstständig zu formulieren oder eine nur teilweise formulierte Antwort zu Ende zu führen. Geschlossene Fragen geben eine feste Antwortenmenge vor, aus denen der Befragte eine Antwort auszuwählen hat. Diese zwei Typen lassen sich auch noch weiter verfeinern in der Art der Fragestellung oder der Anzahl der vorgegebenen Antworten. Auch inhaltlich lassen sich Fragen unterscheiden zum Beispiel zwischen Fragen nach Wertungen, Verhalten oder Fakten. Fragen lassen sich auch als Instrument einsetzen, um die Befragung weiter zu verfeinern und weitere Unterscheidungen in den Antworten vornehmen zu können. Es können Filterfragen gestellt werden, die dazu dienen, bestimmte Personen ausschließen zu können. Andere Fragen wiederum können Reaktionen provozieren, die auf eine Veränderung in der Haltung bzw. Einstellung des Befragten gegenüber der Umfrage selbst oder dem Gegenstand der Befragung abzielen.

Es wird empfohlen [6], bei schriftlichen Befragungen geschlossene Fragen zu verwenden. Diese ermöglichen eine bessere statistische Auswertbarkeit der Ergebnisse und bieten auch einige psychologisch begründete Vorteile. Geschlossen Fragen verlagern weniger Aufmerksamkeit als offene Fragen. Der Befragte muß sich aktiv mit dem Gegenstand der Frage auseinandersetzen. Im Gegensatz dazu benötigen geschlossene Fragen nur die Erinnerung an den Gegenstand der Frage und die Auswahl der passenden vorgegebenen Antwort.

Bei Vorgabe fester Antworten muß die Annahme gemacht werden, daß die Teilnehmer im wesentlichen das gleiche Verständnis von den vorgegebenen Antworten haben. Gibt ein Teilnehmer die Antwort 'gut' muß dies als wirklich besser angesehen werden als würde er 'schlecht' antworten. Dies gilt natürlich auch für die Antworten verschiedener Teilnehmer. Im Gegensatz dazu sind Fragen nach numerischen Werte, wie die Programmier-Erfahrung in Jahren, wesentlich einfacher zu interpretieren. Die Antworten können als sicherer angesehen werden, da hier nur wenig Spielraum für Interpretationen ist.

2 Kurze Darstellung der Artikel

Die im Folgenden kurz vorgestellten zwei Artikel präsentieren, mehr oder weniger ausführlich, die Ergebnisse von Untersuchungen, die eine Grundlage für weitere Untersuchungen zum CMM schaffen sollten. Beide Artikel stammen von den Autoren D.R. Goldenson und J.D. Herbsleb, die beide am SEI tätig waren und sich mit Untersuchungen über das CMM und dessen Weiterentwicklung beschäftigten.

2.1 "A Systematic Survey of CMM Experiences and Results" [1]

Dieser Artikel erschien 1996 im Rahmen der 18. Konferenz über Softwaretechnik. Es werden die Ergebnisse einer Studie aus den Jahren 1994/1995 vorgestellt, deren Ziel es war, den typischen Verlauf von SPI Aktivitäten nach einem SPA kennen zu lernen und mehr über den Zusammenhang zwischen Leistungsfähigkeit und Reifegrad zu erfahren. Weniger ausführlich als der zur selben Studie veröffentlichte Technische Bericht [3], informiert dieser Artikel in groben Zügen über die Durchführung der Studie und deren wichtigste Ergebnisse.

2.2 "Software Quality and the Capability Maturity Model" [2]

Dieser Artikel erschien 1997 in der Zeitschrift "Communications of the ACM" und gibt eine Übersicht über die Ergebnisse dreier Studien zum CMM, die am SEI durchgeführt wurden. Eine dieser Studie ist die in 2.1 angesprochene. Wesentlich länger in der Ausführung als [1] werden die Ergebnisse der Studien erörtert und insgesamt der Stand der Forschung des SEI zum Thema CMM dargestellt.

Die erste vorgestellte Studie wurde durch veröffentlichte Einzelfallstudien inspiriert, die einen großen Erfolg CMM-basierter Prozessverbesserung vermuten ließen. Das SEI sammelte daraufhin weitere Daten, die in aktiv Prozessverbesserung betreibenden Organisationen erfaßt wurden. Diese Daten wurden dann zur Darstellung von Entwicklungen mit der Zeit verwendet.

Die zweite Studie, die auch Gegenstand der vorliegenden Arbeit ist, ist eine Umfrage über die weitere Entwicklung von Aktivitäten der Prozessverbesserung nach einem SPA.

Die dritte Studie befasste sich ausschließlich mit Organisationen, die mehrere SPA durchgeführt hatten. Das SEI konnte dabei auf eine Datenbank zurückgreifen, in der die Ergebnisse solcher Verfahren festgehalten wurden. Dazu gehörten: Der festgestellte Reifegrad, die identifizierten Stärken und Schwächen der Organisation, der Umfang des SPA innerhalb der Organisation und das Datum der Durchführung. Auf diese Weise versuchte das SEI zwei Fragen zu beantworten: Wie lange braucht es, damit eine Organisation eine Stufe im CMM aufsteigt? Und: Was sind die Herausforderungen, die Organisationen, die eine Stufe aufsteigen, von denen, die nicht aufsteigen, unterscheiden?

Die Vorstellung dieser drei Studien nimmt die erste Hälfte des Artikels ein. Die andere Hälfte enthält die Ergebnisse und Schlußfolgerungen, die das SEI aus den Ergebnissen zieht. Insgesamt wird das Bild über die Bedeutung von CMM aber auch der Kritik an CMM klarer und es zeichnet sich ab, daß CMM ein brauchbares und zweckdienliches Mittel für Prozessverbesserungen ist. Der Artikel geht auf einige Kritikpunkte zum CMM ein und kann diese mittels der Ergebnisse der Untersuchungen zumindest entkräften. Abschließend gibt der Artikel eine kurze Zusammenfassung der Erkenntnisse und führt weitere noch offene Fragen auf.

3 Diskussion der Umfrage von Goldenson und Herbsleb

In den folgenden Abschnitten werde ich die Studie von Goldenson und Herbsleb durchgehen und einige charakteristische Aspekte ihrer Untersuchung aufführen.

3.1 Zielsetzung und grundlegende Fragestellung der Umfrage

Zum Zeitpunkt der Studie existierten wenige wissenschaftlich fundierte Informationen über die Stärken und Schwächen des CMM und der darauf basierenden Aktivitäten. Zwar waren bereits einige Erfolgsmeldungen bei der Anwendung von CMM in einzelnen Organisationen veröffentlicht worden, doch repräsentative und zuverlässige Studien über das CMM gab es nur wenige. Die Diskussion über CMM war oftmals auf Meinungen und Vermutungen gestützt. Das Fehlen grundlegender Daten über CMM – Zusammenhänge, Faktoren, Anwendbarkeit des Modells – und die vorhandene Kritik nahmen die Autoren zum Anlaß eine Studie durchzuführen, um eine Grundlage für weitere Untersuchungen zum CMM zu schaffen. Für ihre Studie setzten sich die Autoren die drei folgenden Ziele:

1. Die Studie soll aufzeigen, was typischerweise mit CMM-Ambitionen nach einem CMM-basierten Bewertungsverfahren geschieht.
2. Es soll soviel wie möglich über die Gründe herausgefunden werden, warum manche Anstrengungen erfolgreich und warum andere weniger erfolgreich sind.
3. Die Beziehung zwischen Leistungsfähigkeit und Prozessreife einer Organisation soll näher untersucht werden.

Bezeichnend ist, daß das CMM als Grundlage zur Bewertung der Leistungsfähigkeit von Firmen bereits jahrelang Verwendung fand, obwohl keine Kenntnisse über die Gültigkeit des Modells, also der Zusammenhang zwischen Reifegrad und Leistungsfähigkeit, existierten (dies wird in Frage 3 sehr deutlich ausgedrückt). Das Modell selbst klingt plausibel und ist aus der mehr als dreißig Jahre alten Theorie und Praxis zur Qualitätskontrolle hervorgegangen. Aber die Anwendung in produzierenden, kommerziellen Organisationen ohne Beweise für die Brauchbarkeit ist bemerkenswert.

Insgesamt drücken diese Ziele den explorativen Charakter der Studie aus. Es soll eine Datengrundlage zur Bildung von Hypothesen für die Durchführung weiterer Studien geschaffen werden.

3.2 Das Sample – Die Auswahl der Umfrageteilnehmer

Zur Auswahl der geeigneten Teilnehmer verwendeten Herbsleb und Goldenson die PAIS (Process Appraisal Information System) Datenbank, in der die Ergebnisse früherer Bewertungsverfahren von Organisationen gespeichert sind, die im Laufe der Jahre von autorisierten CMM-Trainern durchgeführt wurden. Die Verwendung dieses Datenbestandes gab ihnen die Möglichkeit, den Entwicklungsstand einer Organisation während eines Bewertungsverfahrens und nach einem ein- bis zweijährigen Zeitraum zu vergleichen, ausgehend von der Annahme, daß durch das CMM angeregte Änderungen des Softwareprozesses in dieser Zeit umgesetzt werden können und ihre Auswirkungen zeigen. Bei der Auswahl beschränkten sie sich auf Organisationen (bzw. SPA) aus den USA und Kanada, um einen zu großen administrativen Aufwand zu vermeiden. Außerdem grenzten sie den Zeitraum des Bewertungsverfahrens auf die Jahre 1992/1993 ein. Damit sollte ausreichend Zeit vergangen sein (mindestens 1 und maximal 2-3 Jahre; die Erstellung des Samples fand im September 1994 statt), damit innerhalb der Organisationen eingeführte Prozessveränderungen ihre Wirkung zeigen konnten.

Einen Aspekt, für das das CMM kritisiert wurde, wird von Goldenson und Herbsleb nicht explizit bei der Auswahl berücksichtigt: Die Größe der beteiligten Organisationen (Anzahl der Angestellten). Leider geben sie in ihrem technischen Bericht [3] keine Informationen über die Verteilung der Organisationsgrößen, obwohl sie darauf Hinweisen, daß die Anzahl der Mitarbeiter bei verschiedenen Fragen einen Zusammenhang zu den erhaltenen Antworten vermuten ließ. So sei das Gefühl, daß bestimmte Aspekte der Prozessverbesserung die eigentliche Arbeit behindern, wesentlich öfter von Umfrageteilnehmern aus kleineren Organisationen verneint worden als von Teilnehmern aus größeren Organisationen.

Aus den ausgesuchten Organisationen mußten anschließend geeignete Personen ausgewählt werden, deren Ansichten und Antworten in die Ergebnisse einfließen sollten. Dabei besteht die Gefahr zu einseitige Ergebnisse zu erzielen, wenn nur Programmierer oder nur Projektmanager oder nur Personen aus anderen Tätigkeitsprofilen befragt werden. Goldenson und Herbsleb wählten drei (ursprünglich vier) am Softwareprozeß (bzw. an der CMM-basierten Bewertung und Umsetzung der Anregungen) beteiligte Tätigkeitsprofile aus, die sie für wichtig genug befanden, um ihre Antworten zu berücksichtigen.

1. Ein Senior-Entwickler, der an der praktischen Entwicklung von Software beteiligt ist. Er sieht die Auswirkungen von Prozessveränderungen auf die Praxis der Softwareherstellung, wie zum Beispiel den Einfluß auf das Betriebsklima unter den Programmierern. Seine Aufmerksamkeit innerhalb des Prozesses gilt unter anderem gutem Design und guter Implementierung.
2. Ein Projektmanager, der den Ablauf eines Projektes überwacht und steuert. Er sieht den Einfluß auf das gesamte Projekt, bei dem Prozessveränderungen eingeführt werden; zum Beispiel die verbrauchten Ressourcen, die verursachten Kosten, Einhaltung der Zeitplanung. Seine Aufmerksamkeit gilt der Einhaltung von zugesagten Terminen und dem veranschlagten Kostenrahmen eines Projektes.
3. Ein Mitglied der SEPG⁴, das direkt zuständig ist für die Einführung von Veränderungen im Prozeß und den Fortschritt des SPI überwacht, leitet und vorantreibt.
4. Ein Senior-Manager, unter dessen 'Schirmherrschaft' das CMM-basierte Bewertungsverfahren stattfand. Er unterstützt innerhalb der Leitung der Organisation die Bestrebungen, den Softwareprozeß weiter zu entwickeln und zu verbessern. Seine Aufmerksamkeit gilt dem Fortschritt des SPI.

Die vierte Rolle wurde im Laufe der Studie außer acht gelassen, da nach Ansicht der Autoren der Bezug dieses Profils zur Thematik CMM nicht ausreichend eng ist, um relevante Daten zu erhalten.

Jedem Profil kann eine andere Sichtweise auf die vom CMM inspirierten Veränderungen zugeschrieben werden. Jedes hat eigene Interessenspunkte am Softwareprozeß und sieht andere Probleme und Verbesserungsmöglichkeiten.

Um Teilnehmer mit diesen Tätigkeitsprofilen zu erhalten, wurden die zuvor ausgesuchten Organisationen gebeten, zu jeder Rolle eine geeignete Person aus ihren Reihen auszuwählen, die an der Umfrage teilnehmen sollte.

Die auf diese Weise eingegrenzte Personengruppe wurde weiter beschränkt durch die Verfügbarkeit von Kontaktinformationen und dem Fehlen wichtiger Bewertungsdaten zu den ausgewählten Organisationen in der PAIS-Datenbank. Letztendlich konnten 61 Bewertungsverfahren und aus diesen wiederum 167 Personen ausgewählt werden, denen als Teilnehmer der Umfrage ein Fragebogen zugesendet wurde.

3.3 Der Fragenkatalog

In der äußeren Form hält sich der Fragenkatalog (eine Kopie ist in [3] zu finden) an die Grundregeln zum Entwurf eines Fragebogens. Er beginnt mit Angaben zu Titel, Autoren und Kontaktadressen. Es folgt eine an die Teilnehmer der Umfrage gerichtete Einleitung, in der nochmals der Gegenstand der Fragen genannt

⁴SEPG (Software Engineering Process Group) ist eine Gruppe innerhalb der Organisation, die die Umsetzung von Methoden und Prozessen innerhalb einer Organisation verfolgt und anregt und damit direkt für die Prozessverbesserung verantwortlich ist

wird. Außerdem werden die Teilnehmer auf die Wichtigkeit gerade ihrer Antworten hingewiesen. Es wird eine geschätzte Bearbeitungszeit von 15 Minuten genannt.

Die Einleitung schließt mit dem Hinweis, daß alle Antworten vertraulich behandelt werden. Ein Hinweis, der nicht nur wichtig ist, um das persönliche Mißtrauen der Teilnehmer einer Umfrage abzubauen⁵. Die Teilnehmer dieser speziellen Studie kommen alle aus einer kommerziell orientierten Umgebung, in der Aspekte wie Aktienkurse oder Betriebsgeheimnisse von großer Bedeutung sind. Fragen zur Leistungsfähigkeit können den derzeitigen Aktienkurs einer Aktiengesellschaft gefährden⁶.

Die Liste der Fragen ist in thematische Blöcke unterteilt, die das durchgeführte Bewertungsverfahren, die bisher umgesetzten Prozessveränderungen, die bewerteten Organisation und den Hintergrund des Teilnehmers zum Thema haben. Insgesamt werden 64, in allen Fällen geschlossen formulierte Fragen gestellt zu denen zumeist ein fester Satz von Antworten vorgegeben wird (bis auf zwei Ausnahmen, bei denen auch freie Antworten gegeben werden können). In sieben Fällen wird nach einer numerischen Antwort gefragt; zum Beispiel die Anzahl der Mitarbeiter, die Anzahl der Jahre in der jetzigen Position/Organisation, der Erfahrung sowie der geschätzte Arbeitszeitanteil für Prozessverbesserung und Qualitätsmanagement über das gesamte vorausgegangene Jahr.

Einige Gründe für die Verwendung geschlossener Frageformen und die Vorgabe eines festen Satzes von Antworten sind in Abschnitt 1.2 aufgeführt. Diese Gründe haben auch bei der Studie von Goldenson und Herbsleb ihre Gültigkeit. Aber den vorgegebenen Antworten muß eine besondere Aufmerksamkeit zu Teil werden. Es sollte keine Lücke vorhanden sein, die den Befragten vor das Problem stellt, eine andere Antwort auswählen zu müssen. Auch wird das Ergebnis der Umfrage durch die vorgegebenen Antworten stark eingeschränkt, was von großer Bedeutung für die Gültigkeit der gesamten Untersuchung ist.

Wie bereits erwähnt, wird zu einem Großteil der Fragen ein fester Satz von Antworten vorgegeben. Die Fragestellung ist in diesen Fällen als Aussage formuliert, zu der eine Bewertung oder Einschätzung abgegeben werden soll. Ein Beispiel für diese Frageform ist die Aussage "*Prozessveränderungen waren leichter umzusetzen, als erwartet*", die sich auf die Charakterisierung des Fortschrittes der Prozessverbesserung nach dem Bewertungsverfahren bezieht und bewertet werden soll. Als Antworten bzw. mögliche Bewertungen stehen zur Auswahl: *Zustimmung, starke Zustimmung, eher keine Zustimmung, überhaupt keine Zustimmung* (in der deutschen Sprache finde ich keine eindeutige Entsprechung zu *strongly disagree*).

⁵Zwei Beispiele: *Wie sehr behindern persönliche Interessenkonflikte und Konkurrenzkämpfe [‘der Angestellten‘ oder auch ‘innerhalb der Organisation‘] den Fortschritt der Prozessverbesserung?* (Im Original: *How much does "turf guarding" inhibit the progress of software process improvement?*)

⁶Beispiel: *To the best of your knowledge, what now is the software process maturity level of the organization that was assessed?*; als Antworten sind die CMM-Stufen und Zwischenschritte vorgegeben.

3.4 Durchführung der Befragung

Der Fragebogen wurde den ausgewählten Teilnehmern postalisch zugestellt. Der Betreuung der Teilnehmer kam eine große Aufmerksamkeit zu, um durch Hilfe bei Fragen und regelmäßige Erinnerung der Teilnehmer das Interesse und damit auch die Rücksenderate hoch zu halten. Goldenson und Herbsleb konnten, bedingt durch den ausgewählten Personenkreis, von einer hohen Grundbereitschaft zur Teilnahme und regem Interesse an der Thematik sowie den Ergebnissen der Studie unter den Teilnehmern ausgehen. Diese hatten bei ihrer Arbeit direkt mit CMM und dem Thema Prozessverbesserung zu tun. Interesse dürften auch die beteiligten Organisationen selbst gehabt haben, da die Studie es ihnen ermöglicht hätte, ihre Aktivitäten und Erfolge mit denen anderer Organisationen zu vergleichen. Zusammen mit einer engagierten Betreuung und regelmäßiger Erinnerung über E-Mail und Telefon konnte eine Rücksenderate von 83% (138 von 167 beantwortete Fragebögen) erreicht werden.

Die 138 Antworten konnten zur Bewertung von 56 der 61 (92%) ausgewählten Bewertungsverfahren verwendet werden. Die rechnerischen Unterschiede zwischen der Anzahl der Teilnehmer und der Anzahl der Bewertungsverfahren kommt daher, daß in einigen Firmen ein Teilnehmer mehrere Rollen innehatte.

3.5 Auswertung der Ergebnisse und Schlußfolgerungen

Im wesentlichen ist die Umfrage über CMM positiv ausgefallen. Die wichtigsten Erkenntnisse, die Goldenson und Herbsleb aus den Ergebnissen der Umfrage gewinnen konnten:

1. Die meisten Teilnehmer stimmen zu, daß das SPA zur Feststellung des Reifegrades ein zutreffendes oder sehr zutreffendes Ergebnis geliefert hat und dazu beitrug, den Softwareprozess zu verbessern.
2. Die meisten Teilnehmer stimmen zu, daß nach dem SPA Fortschritte bei der Umsetzung der durch CMM angeregten Veränderungen zur Prozessentwicklung gemacht worden sind.
3. Organisationen mit einem höheren Reifegrad zeigen eher bessere Leistung bei Qualität, Produktivität, Zeitplanung und einigen anderen Indikatoren.
4. Ein bis drei Jahre scheinen nicht ausreichend zu sein, um auf die nächsthöhere Stufe im CMM zu wechseln.

Bei diesen Erkenntnissen ist zu beachten, daß sie aufgrund der Einschätzungen der Teilnehmer der Umfrage gewonnen werden. Punkt 3 basiert im wesentlichen auf den Selbsteinschätzungen der Teilnehmer bezüglich der gefragten Leistungsindikatoren. Punkt 4 basiert auf der Selbsteinschätzung des Reifegrades der Organisationen der Teilnehmer.

Neben diesen Hauptkenntnissen ließen sich noch andere Zusammenhänge erkennen. So scheint die Art der Tätigkeit im Software-Prozess keinen erkennbaren Einfluß auf die Einschätzung des CMM zu haben. Es zeigten sich keine charakteristischen Unterschiede in den Antworten der verschiedenen ausgewählten Rollen. Tabelle 1 zeigt die Verteilung der Rollen auf die eingegangenen

Anworten. Die Verhältnisse sind erstaunlich gleichmäßig. Wären Unterschiede vorhanden gewesen, wäre eine gute Vergleichbarkeit gegeben gewesen.

Rolle	Anzahl
Senior-Entwickler	47
Projektmanager	47
SEPG-Mitglied	44
Gesamt	138

Tabelle 1: Personenzahlen der verschiedenen Rollen, die zur Umfrage beitrugen

Zur Frage nach Faktoren für eine erfolgreiche Prozessverbesserung konnten ebenfalls einige Erkenntnisse gewonnen werden. Im Fragenkatalog wurden Kritiken und vermutete Erfolgsfaktoren aufgegriffen und deren Auswertung zeigte relevante Zusammenhänge. Zum Beispiel zeigte sich ein positiver Zusammenhang zwischen erfolgreicher Prozessverbesserung und der Beteiligung des Managements. Die ausdrückliche Berücksichtigung der Prozessverbesserung bei der Festlegung der Ziele der Organisation hängt ebenfalls positiv damit zusammen. Ein negativer Einfluß zeigte sich bei Konkurrenzverhalten und interner Organisationspolitik.

Wenn auch nur gering signifikant zeigten auch kleine Organisationen einen proportionalen Zusammenhang zwischen Reifegrad und Leistungsindikatoren. Die geringe Signifikanz resultiert aus dem geringen Anteil kleiner Organisationen bei höheren Reifestufen.

Dies sind nur einige beispielhafte Zusammenhänge, die von Goldenson und Herbsleb in ihrem technischen Bericht [3] aufgeführt werden. Es lassen sich aber mit Sicherheit noch weitere nicht in [3] aufgeführte Zusammenhänge finden.

4 Bewertung und Ausblick

Im ganzen betrachtet halte ich die Studie für gut durchgeführt und ausgewertet. Sie gibt interessante Hinweise zu praktischen Aspekten bei der Prozessverbesserung aufgrund des CMM. Der Zusammenhang zwischen Leistungsfähigkeit und Prozessreife, die Grundidee des CMM, wird näher beleuchtet und gibt damit einige Hinweise auf die Gültigkeit des CMM. In fast jedem Fall werden die zugrundeliegenden Daten angegeben und dadurch begründet. Ich wurde jedoch darauf hingewiesen, daß die Wortwahl an einigen Stellen unsachlich sei. Mangels sprachlicher Kenntnisse möchte ich dazu jedoch keine Beispiele angeben.

Da die in dieser Arbeit behandelte Studie bereits 10 Jahre alt ist, halte ich es für unzumutbar einen Ausblick auf weiterführende Untersuchungen zu geben. Außerdem wurde diese Untersuchung als explorativen Studie betrachtet und Goldenson und Herbsleb geben in ihren Veröffentlichungen ([1][2][3]) bereits selbst einige Ausblicke für weitere Untersuchungen zum CMM.

An einigen Stellen möchte ich jedoch Kritik üben. Die erste Kritik ist die Möglichkeit die Erkenntnisse nachvollziehen zu können. Ich halte die gezogenen Schlußfolgerungen zwar für plausibel, aber es könnten noch weitere interessante Zusammenhänge in den Ergebnissen der Umfrage zu finden sein. Eine auf Seite 16 in [3] zeigt, daß in einigen Organisationen die Ergebnisse des SPA nur gering als Grundlage zur Prozessverbesserung dienen. Es wäre interessant, die Antworten zur Prozessreife und den Leistungsindikatoren dieser Organisationen für sich zu betrachten und mit den restlichen Organisationen zu vergleichen. Es könnte sein, daß kein signifikanter Unterschied zwischen diesen beiden Gruppen vorhanden ist. Aber es dieser Punkt wird zumindest nicht direkt angesprochen.

Die zweite Kritik bezieht sich auf die Methode der Auswahl der Teilnehmer. Auf der einen Seite wurde durch Betreuung und Erinnerung der Teilnehmer viel dafür getan, daß die Rücksenderate hoch ausfällt. Auf der anderen Seite wurde die Auswahl der Teilnehmer zum Teil den Teilnehmern selbst überlassen. Die ausgesuchten Organisation wurden gebeten, Angestellte mit guten Kenntnissen zu CMM und Prozessverbesserung auszuwählen. Dafür wurden Ansprechpartner innerhalb dieser Organisationen, die im PAIS-Datenbestand ebenfalls angegeben sind, kontaktiert und um eine Auswahl gebeten. Damit wird die Teilnehmerauswahl Personen überlassen, deren Position innerhalb dieser Studie nicht berücksichtigt werden. Es ist vorstellbar, daß eine Selbstauswahl eingetreten sein könnte, bei der nur positiv eingestellte Personen befragt wurden. Dieser Fehler in der Methodik ist eigentlich sehr schwerwiegend, da die Gültigkeit von Umfragen zu einem großen Teil von der Auswahl repräsentativer Teilnehmer abhängig ist.

Vielleicht hätte diese Untersuchung von einem externen Institut durchgeführt werden sollen, um dadurch die Möglichkeit der voreingenommenen Einflußnahme auf die Untersuchung zu vermeiden.

Literatur

- [1] J.D. Herbsleb, D.R. Goldenson: *A Systematic Survey of CMM Experiences and Results*; Proceedings of ICSE-18, 1996, S.323f
- [2] J.D. Herbsleb, D.R. Goldenson et. al.: *Software Quality and the Capability Maturity Model*; Communications of the ACM, Juni 1997/Vol.40 No.6, S.30f
- [3] J.D. Herbsleb, D.R. Goldenson: *A Systematic Survey of CMM Experiences and Results*; Technical Report CMU/SEI-95-TR-009, 1995; <http://www.sei.cmu.edu>
- [4] M.C. Paulk, B. Curtis, M.B. Chrissis, Ch.V. Weber: *Capability Maturity Model for Software, Version 1.1*; Technical Report CMU/SEI-93-TR-024, 1993; <http://www.sei.cmu.edu>
- [5] Ch. Deephouse, D.R. Goldenson, M. Kellner: *The Effects of Software Processes on Meeting Targets and Quality*; Proceedings of the 28th Annual Hawaii International Conference on System Sciences, Januar 1995; IEEE Computer Society Press
- [6] Stephan Barth: *Die schriftliche Befragung*; <http://www.stephan-barth.de/schriftl.htm>; Siehe Literaturangaben: RICHTER(1970)
- [7] T.B. Bollinger, C. McGowan: *A Critical Look at Software Capability Evaluations*; IEEE Software 8, Juli 1991, S.25f