



Diplomarbeit

Testgetriebene Entwicklung von Web-Serveranwendungen auf der Basis von Ruby on Rails

Autor: Wienert, Stefan

Studiengruppe: 07/041/02

Betreuender Professor: Prof. Dr. Fritzsche

Datum: 21. September 2011

Inhaltsverzeichnis

1. Einleitung	1
1.1. Motivation	1
1.1.1. Die pludoni GmbH	1
1.1.2. Arbeitsablauf in der pludoni GmbH	3
1.2. Projektbeschreibung und Projektziele	4
1.2.1. Anwendungsfälle	4
1.2.2. Nichtfunktionale Anforderungen	6
1.3. Aufbau der Arbeit	6
2. Automatisierte dynamische Softwaretests	8
2.1. Motivation zum Testen	8
2.2. Arten von Tests	9
2.3. Unittest	11
2.4. Test Doubles – Mocks und Stubs	12
2.5. System- und Akzeptanztests	13
3. Testgetriebene Entwicklung	15
3.1. Motivation	15
3.2. Ablauf	16
3.3. Sonderfälle	19
3.4. Studien zu den Auswirkungen von TDD	20
3.5. Prinzip des Emergent Design – Evolutionäres Softwaredesign	22
3.6. Varianten	23
3.6.1. ATDD – Acceptance TDD	23
3.6.2. Behavior Driven Development	23
3.6.3. Design Driven Testing	24
4. Die Programmiersprache Ruby	25
4.1. Einführung in Ruby	25
4.2. Diskussion	29
4.3. Testframeworks für Ruby	31
4.3.1. Test::Unit und Minitest	31
4.3.2. Cucumber	32

4.4. Ruby on Rails	35
4.4.1. Konzepte von Rails	35
4.4.2. Diskussion	39
5. Code-Metriken	43
5.1. Überblick über Code-Metriken und Skalen	43
5.2. Code-Metriken für Tests	45
5.2.1. Verhältnis von Lines of Test zu Lines of Code	45
5.2.2. Testausführungsabdeckung	46
5.2.3. Mutations/Perturbationstests – Defect insertion	47
5.3. Notwendigkeit von Code Metriken	48
6. Auswahl der Entwicklungsmethodik und -Werkzeuge	49
6.1. Herausbildung einer Entwicklungsstrategie für die Bedürfnisse der plu- doni GmbH	49
6.1.1. Einteilung der Features in Kategorien	49
6.1.2. Weitere Bestandteile der Entwicklungsstrategie	50
6.2. Auswahl der Entwicklungswerkzeuge	51
6.3. Diskussion der Maßnahmen	53
7. Anwendung der Testgetriebenen Entwicklung	54
7.1. Implementierung von Unit-Tests (Modelltests)	54
7.2. Implementierung von Controller-Tests (functional tests)	62
7.3. Testen von externen Abhängigkeiten	66
7.4. System/Akzeptanztests	74
7.5. Testen von Javascript Ereignissen	82
8. Auswertung	84
8.1. Entwicklungsstand ITjobs	84
8.2. Code-Quality-Benchmark mit anderen Ruby-Projekten	88
8.2.1. Vergleich von IT-Jobs-und-Stellen mit eigenen Projekten	89
8.2.2. Vergleich von IT-Jobs-und-Stellen mit anderen Rails Projekten	90
8.2.3. Auswertung und Visualisierung	91
9. Fazit	96
9.1. Ausblick	97

A. Nutzung von Cucumber in Verbindung mit Selenium für Firefox und Guard ohne X-Server	98
B. Eigenschaften erfolgreicher Tests	100
Abbildungsverzeichnis	102
Quellcode-Listings	102
Literaturverzeichnis	105

Acknowledgment

Besonderen Dank gilt meinem Betreuer, Prof. Fritzsche, der sich regelmäßig und ausführlich mit meiner Arbeit beschäftigte und wertvolle Hinweise erteilte.

Mein Dank gilt auch der Fakultät Informatik und Prof. Wiedemann für das zur Verfügung gestellte Büro, der Bibliothek der HTW-Dresden und Prof. Nestler für das Bestellen von tagesaktueller Literatur.

Ich danke auch meiner Frau, die mich während des Schreibens unterstützt hat und bei der Erstellung der Grafiken Tipps gab. Für die orthografische Optimierung sei meinem guten Freund Stefan Koch gedankt.

Diese Arbeit entstand in Zusammenarbeit mit meiner Firma und unter Aufsicht von Jörg Klukas. Ich danke ihm für die Möglichkeit frei zu arbeiten und viele neue Dinge auszuprobieren und für das Vertrauen, dass er in mich gesteckt hat.

Glossar

Im Folgenden werden einige oft-verwendete Begriffe näher erläutert. Innerhalb des Hauptteils dieser Arbeit sind diese Begriffe mit einem [†] gekennzeichnet

Bad Smell oder Code-Smell. Ist ein Anzeichen für eine suboptimale Code-Stelle, die auch ein Hinweis auf ein größeres Design-Problem sein kann. Oft auch ein Kandidat für ein Refactoring. IV, 15, 17

Behavior-Driven-Development (Verhaltensgetriebene Entwicklung). Umformulierung von TDD zur Ausrichtung auf Businessprozesse. Das Vokabular zielt auf die Spezifikation von Erwartungen im Systemverhalten, anstatt Definition nachträglicher Tests.. IV

Code-Metriken Eine Softwaremetrik ist das Ergebnis einer statischen oder dynamischen Codeanalyse zur Generierung von Informationen über den Source-Code. Beispiele: Testabdeckung, Anzahl Codezeilen, Anzahl Bad Smells pro Codezeile.. IV

Code-Qualität beinhaltet die Software-Qualitäten Lesbarkeit, Testbarkeit, Wartbarkeit, Erweiterbarkeit, Geringe Komplexität. IV

CRM Content Management System (Inhaltsverwaltungssystem) ist eine, meist web-basierte Software, die es Nutzern ermöglicht einfach statische Inhalte anzulegen und zu bearbeiten. IV

Entwurfsmuster oder Design Patterns sind bewährte Vorlagen, um häufig wiederkehrende Probleme zu lösen. Weitere Informationen im Buch Design Patterns: Elements of Reusable Object-Oriented Software von Gamme, Helm, Johnson, Vlissides (Gang of Four).. IV

Gem mit in der Ruby-Gemeinschaft Gems (zu deutsch: Edelsteine) werden Bibliotheken Dritter bezeichnet, die in einem zentralem Repository lagern. Das bekannteste dieser Repositories ist rubygems.org. Nahezu alle Ruby-Bibliotheken lassen sich hier finden und innerhalb von Sekunden mittels des Kommandozeilenwerkzeuges gem installieren. Beispielsweise ließe sich [†]Ruby on Rails mittels `gem install rails` installieren.. IV, 32

MVC Model-View-Control, ist ein [†]Entwurfsmuster, dass insbesondere bei GUI- und

Web-Anwendungen beliebt ist. Rails basiert auf dem MVC-Muster. IV, 35, 54, 66

ORM Objekt Relationales Mapping, ist eine Persistenz-Technik, um Objekte transparent in einer Datenbank zu speichern, und andersrum, Tabellenzeilen in Objekte wieder rückzuübersetzen. IV, 35, 37, 54

Refaktorisieren Ist eine Modifikation des Programmcodes ohne Modifikation es externen Verhaltens um nicht-funktionale Eigenschaften des Quellcodes zu verbessern, wie z.B. Lesbarkeit, Wiederverwendbarkeit, Wartbarkeit. IV, 17

RSS RDF Site Summary, ist ein standardisierter XML-Dialekt zur maschinenlesbaren Verteilung und Veröffentlichung von Inhalten. Es existiert in den Versionen 0.9 bis 2.0.1, die sich nur in Details, wie z.B. Einbindung von Rich-Media (Podcasts, ..) und Namespaces unterscheiden. IV

Ruby on Rails ist ein auf der Programmiersprache Ruby basierendes Web-Framework und ist Gegenstand des Kapitel 4.4 text. IV, 15, 22, 29, 32, 39

Software-Fehler oder Defekt, ist ein unerwartetes Verhalten der Software, der zu einem Versagen der Software führen kann. IV

Test oder Testfall ist eine, meist automatisierte, Prüfung des Programmverhaltens bei definierten Eingabeparametern. IV, V, 16

Test-Double haben die Aufgabe, komplexe Objekte in einem isolierten Test zu simulieren, in dem statt komplexer Berechnungen oder externer Zugriffe konstante Werte geliefert werden. Vertreter dieser Test Doubles sind die Mocks und Stubs, siehe Abschnitt 2.4. IV, V, 11, 88

Test-Runner oder Testtreiber) ist eine Software, um \uparrow Tests aufzurufen und deren Ausführung zu überwachen und steuern. IV, V

Test-Suite ist eine Sammlung von mehreren Testfällen für eine Komponente oder das gesamte System. IV, 15

Test-Umgebung ist eine spezielle Konfiguration der Hard- und Software, um Tests unter kontrollierten und bekannten Bedingungen auszuführen. Dies beinhaltet neben dem zu testenden Objekten (Object under Tes) auch sogenannte \uparrow Test-Doubles und den \uparrow Test-Runner.. IV

Testabdeckung auch: Testausführungsabdeckung, Überdeckungsgrad, Testfallabdeckung.

Eine dynamische Code-Metrik, die angibt, welche Codezeilen durch keinen Test abgedeckt wurde. Es wird unterschieden in die Stufen C0, C1 und C2 mit steigender Komplexität der Messung. Details im Abschnitt 5.2. IV

Testgetriebene Entwicklung englisch Test-Driven-Development/Test-Driven-Design, Ausführlich dargelegt in Abschnitt 3. IV, 6, 15, 19, 20, 23, 48, 54

1. Einleitung

1.1. Motivation

Kurz nach der Firmengründung der pludoni GmbH absolvierte der Autor dort sein Pflichtpraktikum, und war bis zum heutigen Tag als Werkstudent tätig. Während dieser Zeit nahmen Programmierer verschiedener Erfahrungsstufen und auch Praktikanten an der Neu- und Weiterentwicklung der Webserver-Software teil. Dies hat zur Folge, dass die Komplexität der Software inzwischen ein Level erreicht hat, dass das Maß an Regressionsfehlern¹ stark anstieg.

Da zum großen Teil keine automatisierten Softwaretests geschrieben wurden, lassen sich diese auch nur schwer aufspüren. Ein Versuch, nachträglich Softwaretests hinzuzufügen wurde evaluiert und als zu aufwändig befunden, da der Code in seinem jetzigen Zustand nur äußerst schwer zu testen ist. Gründe dafür sind suboptimale Codestrukturen (Spaghetticode) die schwer bis unmöglich zu testen sind. Hier müsste zuerst refaktorisiert werden, aber da keine Tests vorhanden sind, ist dies aufgrund der Regressionsfehler riskant. Ein Teufelskreis!

Für ein neues Projekt, it-jobs-und-stellen.de, soll dies nun mit einem anderen Ansatz verlaufen.

Neben der Umstellung auf ein modernes Web-Framework, sollen nun Tests im Einklang zum Code erstellt werden, um auf Knopfdruck umfassende Informationen über den Systemzustand zu erhalten, wie sie ein manueller Test in der Gründlichkeit und Schnelligkeit niemals erreichen kann. Die testgetriebene Entwicklung ist dabei ein populärer Entwicklungsprozess der eigentlich als Ziel hat, gut testbaren, d.h. gut wartbaren und lesbaren Code zu schreiben. Quasi nebenbei erhält man aber auch eine umfassende Test-Suite, und kommt so dem Schritt, funktionierende Software zu entwickeln, einen Schritt näher.

1.1.1. Die pludoni GmbH

Die pludoni GmbH ist ein junges dynamisches Dresdner Unternehmen, dass sich zum Ziel gesetzt hat, lokale Communitys zur gegenseitigen Fachkräfteempfehlung aufzubauen und zu betreuen, sowie Tools für die Vereinfachung der Personalarbeit mittelständischer Unternehmen zu entwickeln. Einige Beispiele für diese Communitys sind

¹ fehlerauslösender Quelltextänderungen

zur Zeit ITSax.de, ITmitte.de und MINTsax.de².



Funktionsweise der Communitys Die Communitys bestehen jeweils aus einer Anzahl mittelständischer Unternehmen einer Branche, Bildungseinrichtungen sowie Vertretern von Städten und Vereinen. Für ITSax.de ist das die IT-Branche. Neben diesem Branchenfokus sammeln sich nur Unternehmen einer spezifischen Region. Bei ITSax.de ist dies Mittel- und Ostsachsen, bei ITmitte.de z.B. Mitteldeutschland, d.h. Thüringen, Sachsen-Anhalt und Westsachsen.

Tabelle 1: Übersicht über pludoni Communitys. Stand September 2011

Community	Branche	Region	Mitglieder
ITSax.de	IT, Software	Sachsen (Fokus Süden und Osten)	63
ITmitte.de	IT, Software	Mitteldeutschland (Thüringen, Sachsen-Anhalt, Westsachsen)	57
MINTsax.de	Maschinenbau, Elektrotechnik	Sachsen	29
OFFICESax.de	Vertrieb, Bürotätigkeiten	Sachsen	in Vorb.
OFFICEmitte.de	Vertrieb, Bürotätigkeiten	Mitteldeutschland	in Vorb.

Diese Unternehmen, die einen jährlichen Mitgliedsbeitrag für das Community-Management³ und Weiterentwicklung der Portale an die pludoni GmbH zahlen, dürfen Ihre für die Region relevanten Jobangebote auf dem jeweiligen Portal einstellen. Was die Communitys von pludoni von der dem bisherigen Online-Jobbörsen unterscheidet, ist das sogenannte **Empfehlungssystem**.

Viele der Personalchefs der beteiligten Firmen haben dieselbe Erfahrung gemacht, dass sie sehr guten Bewerbern absagen mussten, weil z.B. die Stelle schon vergeben wurde, die Fähigkeiten des Bewerbers nicht den Bedürfnissen des Unternehmens entsprachen, eine Einstellung verhinderte. Hier setzt pludoni mit seinen Communitys an, und stellt eine Infrastruktur zur gegenseitigen Empfehlung dieser guten Bewerber bereit. Ausgezeichnete Bewerber erhalten neben der Absage einen Empfehlungscode, mit dem sie

²<http://www.itsax.de/>, <http://www.itmitte.de>, <http://www.mintsax.de/>

³Ein Überblick über die Aufgaben eines Community-Managers finden Sie unter <http://www.pludoni.de/leistungen>

sich auf dem Online-Jobportal bei einer der anderen Mitgliedsfirmen bewerben können. Die Software löst intern den Empfehlungscode auf und bestätigt dieser Firma nun, dass der Bewerber von einem anderen Unternehmen der Community empfohlen wurde.

Dieses Empfehlungssystem überzeugt die beteiligten Unternehmen. Aktuell wurden im letzten Jahr z.B. auf ITmitte.de über 800 Bewerbungen über das Portal versendet, von denen mehr als die Hälfte (440) mit Empfehlungscode versehen waren [Klukas, 2011b]. Dies motivierte mittlerweile über 150 Organisationen bei den drei pludoni Communitys teilzunehmen [Klukas, 2011a].

1.1.2. Arbeitsablauf in der pludoni GmbH

Die pludoni GmbH stellt sich dem Trend der Dezentralisierung von Arbeit, um einerseits Kundenwünsche zur individuellen Betreuung und andererseits Mitarbeiterwünsche nach flexiblen und familienfreundlichen Arbeitsplätzen gerecht zu werden. Ein Großteil der Arbeit findet somit vor Ort beim Kunden, oder in Heim-/Telearbeit statt. Zur persönlichen Abstimmung findet aber mindestens einmal pro Woche ein Meeting statt, in welcher sich 2-4 der Mitarbeiter treffen, um alte Aufgaben abzunehmen und neue zu diskutieren. Die Abnahme erfolgt dabei durch den Teamleiter des Projekts oder dem Geschäftsführer Jörg Klukas.

Zentrales Kommunikationsmittel der pludoni GmbH ist, neben der E-Mail, die Online-Aufgaben- und Fehlerverwaltung, Redmine⁴. Dort werden alle Aufgaben und Fehler erfasst und an die zuständigen Personen verteilt. Neben den technischen Aufgaben der Entwickler, werden auch nicht-technische Aufgaben der anderen Mitarbeiter verwaltet, wie z.B. die Gewinnung neuer Partner (Akquise) oder administrative Aufgaben.

Trotz dieses Tools und Vorgehensweise ist die dezentrale Kollaboration aber immer mit Nachteilen in der Kommunikation behaftet. Dies hat bei den Programmierern teils gravierende Auswirkungen auf die Produktivität. Einerseits, da gleiche Funktionalität doppelt implementiert wird, weil der Überblick fehlt, und somit unnötigerweise neue Fehlerquellen eröffnet werden. Andererseits, weil aufgrund der zeitlich asynchronen Arbeitsleistungen Rückmeldungen der anderen Programmierer oder ein Code-Audit nicht immer gewährleistet werden und Regressionsfehler nicht abgeschätzt werden können, da nicht alle Module bekannt sind.

⁴<http://www.redmine.org> - ein webbasiertes Projektmanagement-Tool auf der Basis von Ruby on Rails. Redmine kann für Benutzer- und Projektverwaltung, Diskussionsforen, Wikis, zur Ticketverwaltung oder Dokumentenablage genutzt werden, *Wikipedia*

Für das kommende Projekt soll nun eine großflächige Testinfrastruktur erstellt werden, um eine aktuelle Dokumentation des Quelltextes zu erhalten und die Code-Qualität verbessern und natürlich um das Risiko für Regressionsfehler zu minimieren.

1.2. Projektbeschreibung und Projektziele

Für den praktischen Teil dieser Arbeit soll anhand der Entwicklung einer Ruby-on-Rails-Anwendung die Testgetriebene Entwicklung erprobt und angewendet werden.

Als Ergänzung zu den lokalspezifischen Communitys mit strenger Mitgliederauswahl soll nun ein neues, allgemeines IT-Jobportal entwickelt werden. Der vorraussichtliche Name wird IT-Jobs-Und-Stellen.de⁵ sein.

Ziel soll es sein, den regionalen Organisationen auch eine Alternative zu den Community-Mitgliedschaften anzubieten, um kurzfristigen Personalbedarf zu decken, analog zu den Branchen prima der Online-Stellenbörsen stepstone.de, monster.de und jobscout24⁶ zu entwickeln. Analog zu Abbildung 1 wurden die Anwendungsfälle wie folgt analysiert.

1.2.1. Anwendungsfälle

Es gibt im zwei verschiedene Nutzertypen. Zum einen, ein Bewerber, der das Ziel hat einen Job zu suchen, zu finden und sich darauf zu bewerben. Zum anderen ein Mitarbeiter eines Kundenunternehmens, der Jobs auf der Webseite schalten möchte. Verbal formuliert sind die Anwendungsfälle die folgenden:

- Ein Bewerber kann die Webseite nach sichtbaren Stellenangeboten durchsuchen
- Ein Bewerber kann die Detailanzeige einer Stelle betrachten
- Ein Bewerber kann sich auf eine Stelle über das System bewerben. Dabei kann er auch eine Verbindung zu seinem Facebook oder LinkedIn Account herstellen, um so automatisch einen Lebenslauf generieren zu lassen.
- Ein Kunde kann sich am Portal registrieren und dann anmelden, und wird so zum „Arbeitgeber“
- Ein Arbeitgeber kann eine Stellenanzeige für ein Vielfaches von 30 Tagen schalten. Damit wird ein Bezahlvorgang über einen Drittanbieterportal (z.B. paypal)

⁵<http://www.it-jobs-und-stellen.de/>

⁶<http://www.stepstone.de> - <http://www.monster.de> - <http://www.jobscout24.de>

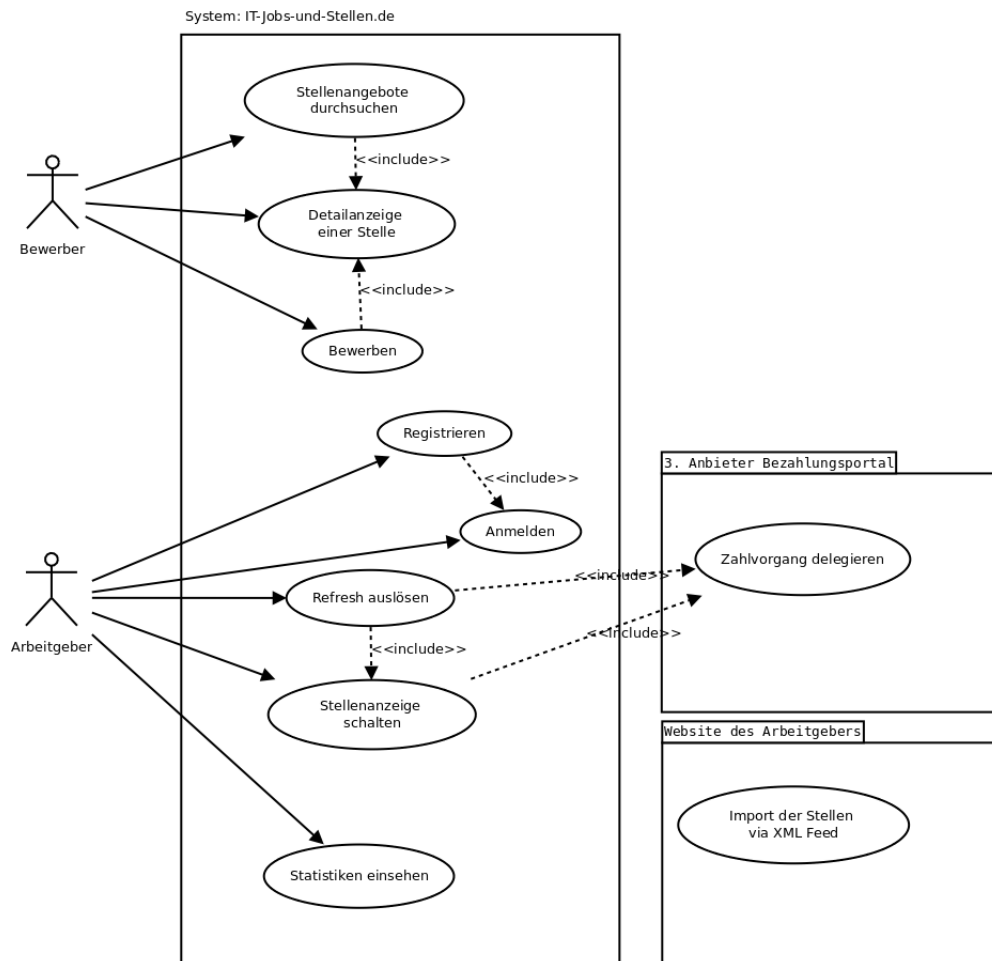


Abbildung 1: Anwendungsfälle

ausgelöst und bei Bestätigung die Stelle sichtbar geschaltet. Er kann dazu zwischen 5 verschiedenen Designs wählen, und Logo und Bild hinzufügen.

- Ein Arbeitgeber kann eine laufende Stelle erneuern (refresh), und so seine Platzierung verbessern. Dies ist möglich, weil die Aktualität einer Stelle Auswirkungen auf die Platzierung in den Suchergebnissen hat. Ein Refresh kostet Geld und löst somit wieder einen Bezahlvorgang aus
- Ein Arbeitgeber kann Statistiken über seine geschalteten Stellenanzeigen einsehen. Dies beinhaltet die Anzahl der Zugriffe pro Tag, Anzahl der Bewerbungen je Stelle
- Ein Arbeitgeber kann seine Stellen mittels eines XML-Feed-Imports automatisiert einlesen. Dazu bezahlt er einen Pauschalbetrag je Kalenderjahr. Das System holt nun alle 6h liest nun alle Stellen aus diesem Feed ein. Das Format ist eine Modifikation von RSS2.0

1.2.2. Nichtfunktionale Anforderungen

Folgende Rahmenbedingungen und zusätzliche Ziele wurden vereinbart.

- Eine hohe C0-Testabdeckung von mindestens 95% als Grundlage für den [†]Testgetriebene Entwicklung-Prozess
- Eine hohe Erweiterbarkeit, um langfristig auch die bereits vorhandenen Communityportale durch das neue System zu ersetzen, welche gegenwärtig auf dem Framework Drupal 5 (PHP) basieren
- Eine moderne Suchfunktion durch einen Suchdaemon realisieren, z.B. Sphinx oder Lucene
- Softwarestack: Ruby 1.9.2 mit Ruby on Rails 3.1, Javascript mit JQuery 1.6
- Die Software soll eine möglichst einfache und eingängige Bedienung haben

1.3. Aufbau der Arbeit

Diese Arbeit ist in 4 Teile mit insgesamt 11 Abschnitten gegliedert. Abschnitt 1 bis 3 haben einleitenden Charakter. In den darauffolgenden Abschnitten 4 bis 7 werden theoretische Grundlagen und Technologien, die für die Aufgabenstellung relevant sind, näher beleuchtet. Dies beinhaltet eine Einführung in die Sprache Ruby, das Framework

Ruby on Rails, sowie Automatisierte Softwaretests und den Prozess der Testgetriebene Entwicklung. In den Abschnitten 8 bis 10 werden dann praktische Erfahrungen und Implementationsdetails dargestellt. Im 11. und letzten Abschnitt wird ein Fazit gezogen und ein Ausblick für weitere Forschungstätigkeit zum Thema dargeboten.

Hinweis zu den Quellcode Beispielen

Hinweis zu Glossareinträgen

2. Automatisierte dynamische Softwaretests

Zum Prüfen der Korrektheit seiner Arbeit macht jeder Programmierer mindestens manuelle Tests. Bei einer Webanwendung hieße dies konkret, den Webserver zu starten und mittels eines Browsers durch die Anwendung zu navigieren, Daten anzulegen und Ausgaben der Anwendung zu kontrollieren. Mit zunehmender Größe einer Anwendung wird es immer aufwändiger, die Software zu testen, da nach jedem Hinzufügen von Funktionalität eigentlich alle Aspekte wieder getestet werden müssen, um Regressionsfehler auszuschließen.

Stattdessen werden automatisierte Softwaretests instrumentalisiert, um auf Knopfdruck alle bisher programmierten Tests auszuführen und so ein Bild über den Zustand der Anwendung zu erhalten. So ist Automatisiertes Testen dem manuellem Testen, das auch Ad-Hoc-Testen genannt wird, in kürzester Zeit zeitlich überlegen [Rappin, 2011].

Dynamische Testverfahren haben gemeinsam, dass sie stichpunktartig testen, damit die Korrektheit des Programmes nicht beweisen können und die Ausführung des Programmcodes mit konkreten Eingabewerten Liggesmeyer [1990][S. 49]. Diese Stichproben werden als **Testdaten** bezeichnet, die optimalerweise repräsentativ, fehlersensitiv, redundanzarm und ökonomisch sind [Liggesmeyer, 1990][S. 51].

Neben den dynamischen Tests, also Tests die Programmcode ausführen, gibt es statische Analyseverfahren, wie formale Verifikation, symbolische Testverfahren oder statische Analysen. Einige statische Anaylsen werden später im Abschnitt 5.2 vorgestellt. Andere statische Testverfahren sind nicht Gegenstand dieser Diplomarbeit.

2.1. Motivation zm Testen

Tests dienen in erster Linie dazu, das Vorhandensein bzw. Nichtvorhandensein von Software-Fehlern zu belegen [Goodliffe, 2006].

Getester Code gilt im Allgemeinen als robuster, korrekter und leichter zu warten [Rappin, 2011]. Im Umkehrschluss bedeutet dies, drastisch formuliert, dass ungetestete Software mit sehr hoher Wahrscheinlichkeit Fehler beinhaltet [Goodliffe, 2006].

Die meisten Programmierer bevorzugen Testen gegenüber dem Debuggen. Testen führe zu einer Minimierung der Debugphase und mache die Software-Entwicklung für Programmierer attraktiver und für Projektleiter leichter zu planen [Orsini, 2007] und insgesamt preiswerter macht [Liggesmeyer, 1990][S.13].

2.2. Arten von Tests

Tests können nach verschiedenen Gesichtspunkten eingeteilt werden. In vielen konkreten Fällen reicht eine eindimensionale Einordnung nicht aus, und Tests können ebenso Teil von mehreren Kategorien sein.

Einteilung nach Sichtbarkeit des Quellcodes Tests werden in **Whitebox** und **Blackbox**-Tests eingeteilt. Whitebox-Tests finden mit Wissen über den zugrundeliegenden Code statt. Ziel eines Whitebox-basierten Testverfahrens ist es, so viele Codeabschnitte wie möglich zu testen. Blackboxtests dagegen ignorieren den inneren Aufbau der Klassen und testen entweder nur Schnittstellen oder das Gesamtsystem, fokussieren also Aktionen und Rückmeldungen des Systems. Das Ziel eines Blackbox-basierten Tests ist die Korrektheit der Software gegenüber den Spezifikationen. Ein Spezialfall sind die sogenannten **Greybox**-Tests, die insbesondere bei der Testgetriebenen Entwicklung auftreten. Da der Test zuerst entwickelt wird, ist noch kein Wissen über den Zielquellcode vorhanden.

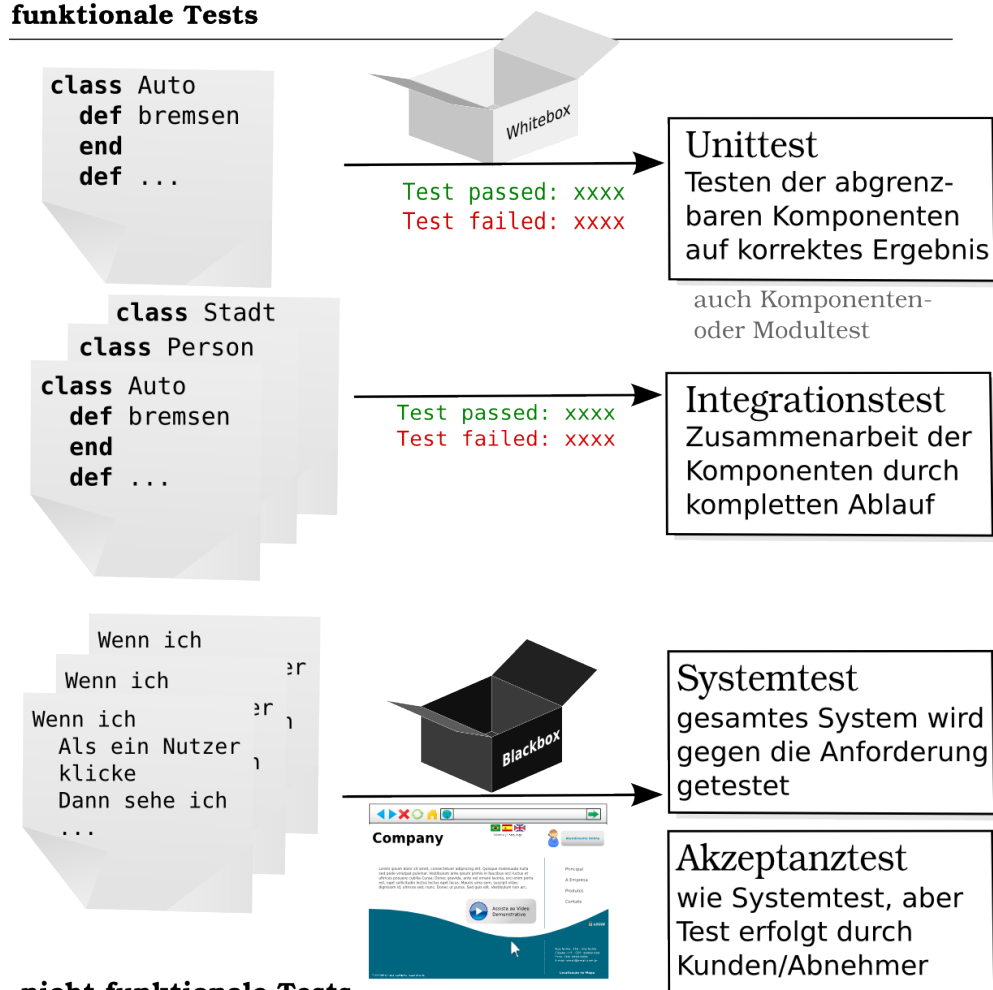
Einteilung nach Testziel (nach Andrew Hunt [1999][S. 238ff])

Unittests Hierbei werden die Einheiten des Programmes auf ihr Verhalten getestet. Dies stellt die Basis für die meist darauffolgenden Integrationstests dar.

Integrationstests das Zusammenspiel zwischen Klassen wird getestet, welche ein gemeinsames Subsystem darstellen (Modul). Hier kann eine Bottom-Up oder Top-Down-Integrationsstrategie verfolgt werden, d.h. ob man mit denjenigen Modulen beginnt, die keine Abhängigkeiten haben und im Ableitungsbaum immer weiter nach oben integriert (bottom-up), oder mit dem zentralen Modul anfängt und nach und nach alle abhängigen oder abgeleiteten Module testet (Top-Down).

Validierung und Verifikation Testet den Fortschritt der Anwendung in Bezug auf die funktionalen Anforderungen. Dies ist meist ein Blackboxtest und testet das System als ganzes (=Systemtest). Ein Spezialfall ist der Akzeptanztest. Hierbei nimmt der Kunde eine Anforderung/Feature ab.

Ressourcennutzung, Performanz, Verhalten im Fehlerfall Die vorherigen Tests finden i.d.R. unter idealen Bedingungen statt. Diese Testkategorie versucht das Applikationsverhalten unter realen Bedingungen zu simulieren. Beim Verhalten im Fehlerfall soll getestet werden, dass der Nutzer nicht durch kryptische Fehler-

funktionale Tests**nicht-funktionale Tests**

Stress/Lasttest Verhalten in Extremsituationen, Zuverlässigkeit	Performanztest Antwortzeit, Skalierbarkeit, Nutzung von Ressourcen	...
Usabilitytest Verständlichkeit und Konsistenz des Nutzer-Interfaces	Sicherheitstest Test auf Sicherheitslücken und Programmierfehler	...

Abbildung 2: Einteilung der Tests

meldungen verwirrt wird, oder z.B. sein Fortschritt gespeichert wurde. Last und Performanztests stellen sicher, dass die Anwendung eine große Zahl von Nutzern oder eine große Menge an Daten verarbeiten kann.

Usability Testing Diese Testmethode kann gegenüber den bisher genannten nicht automatisiert werden, und benötigt immer einen zukünftigen Endanwender. Ziel ist es, die Benutzbarkeit und Handhabung zu testen. Dies wird durch Beobachtung von Kandidaten, meist in einer präparierten Umgebung (Usability Labor) geprüft.

Einteilung nach Testvollständigkeit Tests können auch nach dem Gesichtspunkt eingeteilt werden, wie eine Testvollständigkeit beurteilt wird und notwendige Testfälle generiert werden [Liggesmeyer, 1990]:

Kontrollflussorientierter Test betrachtet den Quelltext und leitet daraus notwendige Unittests ab

Datenflussorientierter Test Beobachtet die Variablendefinitionen, Zugriffe und Entscheidungen anhand der Variablen

Funktionaler Test leitet aus den funktionalen Spezifikationen Testfälle her

Diversifizierender Test Testet verschiedene Versionen eines Programmes gegeneinander. Dies beinhaltet z.B. Mutationentest, der später im Abschnitt 5.2.3 *Mutations/Perturbationstests – Defect insertion* als Metrik zur Beurteilung der Tests erläutert wird.

2.3. Unittest

Da die Testgetriebene Entwicklung in ihrer Reinform auf dem Unittest basiert, soll diese Testgattung im Vorfeld etwas näher beleuchtet werden.

Ziel des Unittests ist es, frühzeitig Fehler im Code zu finden. Der Unit, oder Modultest beschreibt das Testen der Einheiten eines Programms, die im Sinne der Testung nicht weiter zerlegt werden können. Dies können die Funktionen, oder bei einer objektorientierten Sprache, die Klassen sein. Die Objekte unter Beobachtung (Objects under Test) werden beim Unittest werden in strenger Isolation zu den restlichen Units ausgeführt. Abhängigkeiten der Module untereinander und von unterlagerten Diensten werden durch [†]Test-Doubles simuliert. Dies ist notwendig um sicherzustellen, dass

gefundene Fehler von dem betreffenden Modul verursacht wurden, und nicht durch äußere Einflüsse. Diese Isolierung macht das Testen einfacher [Goodliffe, 2006].

Unittests werden fast immer automatisiert ausgeführt. Verwendung finden dabei meist immer sogenannte Test-Frameworks. Eine der meist-verbreitetsten sind die Frameworks auf Basis von xUnit, die in nahezu allen (objektorientierten) Sprachen Vertreter haben, so z.B. Test::Unit/MiniTest in Ruby, JUnit in Java oder NUnit in C#.

Ein solcher Test besteht in der Regel aus 4 Teilen:

1. Initialisierung der Test-Umgebung und der Objekte
2. Ausführung der zu testenden Aktion, die den Systemzustand ändert
3. Prüfen der spezifizierten Erwartungen (Assertions)
4. Aufräumen nicht mehr benötigter Objekte, File-Pointer, Sockets, Leeren der Datenbank u.ä

Komponententest Ziel des Komponententest ist es, verschiedene Units in Kombination als eine vollständige Komponente zu testen [Goodliffe, 2006].

2.4. Test Doubles – Mocks und Stubs

Beim isolierten Testen wird durch Abhängigkeiten aller Art erschwert. Dies können z.B. Klassen sein, die noch nicht implementiert wurden, externe Ressourcen (Netzwerkzugriffe, Versenden von Mails) oder externe unterliegende Prozesse (Bezahlen in einem Onlineshop, Datenbanken) sein. In diesen Situationen ist es angebracht, auf sogenannte Test-Doubles, meist Mocks und Stubs, zurückzugreifen.

Ein **Stub** ist eine nachahmende Funktion oder Objekt, welches die schwer zu isolierende Klasse während des Testfalls ersetzt. Im Beispiel ein Bezahlprozess einer Bestellung.

```
def test_report_failed_payment
  Payment.stubs(:pay).returns(false)

  bestellung = Bestellung.new()
  bestellung.commit()

  assert bestellung.not_valid?
end
```

Mit dem oben angegeben (Pseudo-) Rubycode würde man z.B. mittels des Mock-Frameworks „mocha“,⁷ ein Stubobjekt erzeugen, welches den Bezahlprozess nachahmt, und die Methode „pay“ ersetzt, so dass sie immer „false“ zurückgibt, d.h. der extern Bezahlvorgang wird in diesem Test fehlschlagen, z.B. wenn der Benutzer falsche Zahlungsdaten angegeben hat. Auf diese Weise könne komplexe Operationen auf ihre Interfaces reduziert werden. Außerdem machen Stubs den Test i.d.R. schneller, da statt der potentiell komplexen und langsamen Operationen statische Werte geliefert werden.

Als Ergänzung dazu gibt es **Mocks**. Ähnlich wie die Stubs ersetzen sie Methoden oder Objekte, um statt komplexer Operationen fixe Werte zurückzugeben. Zusätzlich dienen Mocks selbst als Testfall. Ein Mock wartet darauf, ob die Methode, wie sie definiert wurde, auch tatsächlich aufgerufen wurde.

Hier z.B. ein Mock um einen Netzwerkzugriff zu testen und abzufangen.

```
def test_always_fail
  HTTP.expects(:get).with("http://www.google.com")
end
```

Der gezeigte Test wird immer fehlschlagen, da von einem Mock erwartet wird, dass die nachgeahmte Funktion während des Tests genau einmal aufgerufen wird. Ist dies nicht der Fall, gilt der Test als nicht bestanden. Mocks fungieren somit als zusätzliche Möglichkeit Interna des Programmflusses zu testen.

2.5. System- und Akzeptanztests

Der Unittest ist als Whitebox-Test auf Wissen über den Quelltext angewiesen, und der Zweck ist es, Fehlerfreiheit zu gewähren. Der **Systemtest** dagegen testet das gesamte integrierte System, meist aus der Sicht eines Anwenders, mit dem Ziel, die Software gegenüber den Anforderungen zu validieren. Diese Tests finden unter realen Bedingungen mit realen Daten statt und untersucht, meist sogar auf einer den Parametern der Produktionsumgebung nahe-kommende Hard- und Softwareumgebung.

Der **Akzeptanztest** ist ein Spezialfall des Systemtests. Hier führt der Auftraggeber der Software den Test selbst durch. Er nutzt den Akzeptanztest, um zu entscheiden, ob er die Software akzeptiert, woher der Name rührt. Innerhalb des Kontextes der Agilen Programmierung, dem auch die Testgetriebene Software zuzurechnen ist, dienen Ak-

⁷<http://mocha.rubyforge.org/>

zeptanztests, um den Fortschritt bei der Bearbeitung der „Geschichten“ (user stories) zu überwachen.

Für das Testen von Webserveranwendungen spielt die Simulation eines Browsers eine große Rolle, um die Tests unter möglichst realen Bedingungen durchzuführen. Hierbei bedient man sich der Fernsteuerung eines Browsers mittels einer Middleware, die es ermöglicht innerhalb eines Testfalls einen Webbrowser zu starten. Eine der bekanntesten dieser Middlewares ist das Framework **Selenium**⁸, welches Firefox, Internet Explorer und Google Chrome fernsteuern kann. Dies ermöglicht eine sehr detaillierte Testung auf Browserinkompatibilitäten, da es unter den Browsern gewisse Unterschiede in der Ausführung von Javascript und Darstellung von Elementen gibt. Dieses Tool wird u.a. von Google, Oracle und eBay zum Tests ihrer Anwendungen verwendet und weiterentwickelt [HQ, 2010].

⁸<http://seleniumhq.org/>

3. Testgetriebene Entwicklung

Testgetriebene Entwicklung, im Englischen Test-Driven-Development (TDD) wurde erstmalig von Kent Beck 2002 im Detail erläutert [Beck, 2002]. Zuvor war die Technik „Test-First“ aber schon seit 1999 im Kontext von Extreme Programming (XP) bekannt.

Test-Driven-Development wird als

TDD := Test-First + Refaktorisieren

beschrieben [Ambler, 2002]. So ist es Ziel, dass sich Test schreiben/Implementieren und Refaktorisierungen, d.h. das konstante Verbessern des Systemdesigns und des Quellcodes, abwechseln. Mittlerweile werden die Begriffe Test-First Development und [†]TDD synonym verwendet, allerdings gibt es manchmal Unterschiede in der Herangehensweise in Punkto Design der Software: TDD findet seine Anwendung, wenn nur eine vage Idee der Funktionalität einer Klasse besteht, während TFD kein Design oder Redesign der Klassen vorsieht [Stackoverflow].

Im Folgenden wird die Entwicklungsmethode [†]TDD näher beleuchtet und am Beispiel von [†]Ruby on Rails typische Testwerkzeuge aufzeigen.

TDD is also good for geeks who form emotional attachments to code.

Beck

3.1. Motivation

Das Erstellen einer gut-abgedeckten [†]Test-Suite für ein jedes größeres Softwareprojekt ist eine wichtige Voraussetzung um interne Qualitäten, wie Wartbarkeit und Zuverlässigkeit zu aktivieren. [†]TDD soll nicht dazu dienen, die Software zu auf Korrektheit zu untersuchen⁹. Dies ist aber ein positiver Nebeneffekt. Das Hauptziel ist es, den Code in Einklang mit dem Test zu schreiben, so dass der Test den Code antreibt (Test drives the code). Der messbare Effekt davon, ist ein gut-testbarer Code, welcher in der Regel auch ein gut-wartbarer und verständlicher Code ist. [†]Bad Smells, wie God-Methode und geringe Kohäsion, werden schon im Keim erstickt werden, da diese nur äußerst schwer zu testen sind.

Psychologische Aspekte und Aspekte des Projektmanagements Kent Beck beschreibt die Hauptmotivation für TDD, als das „managing fear during programming“

⁹TODO Kent Beck Verification und ISO 9000:2005 3.8.5

Management von Angst. So hat Angst verschiedene Auswirkungen auf die Entwicklung. Sie mache zögerlich, führe zu weniger Kommunikation und Feedback und mache den Programmierer „mürrisch“ [Beck, 2002, S. xi].

TDD fördert die Entwicklung in kleinen Schritten, und ermöglicht durch bestandene Tests kleine „Belohnungen“ für den Programmierer. Dadurch ist es leichter einen gewissen Arbeitsrhythmus zu erhalten, was stellenweise dem „Flow“¹⁰ ähnelt oder diesen strukturiert ergänzen kann [Brown, 2008].

Falls TDD die Fehlerdichte signifikant verringern würde und nur Code entstünde, der getestet wurde, so hätte dies wohl auch soziale Auswirkungen auf das Entwicklerteam [Beck, 2002, S. x].

1. Die Qualitätssicherung könnte von einer reaktiven, auf eine proaktive Arbeit umstellen.
2. Der Projektmanager kann den Ablauf der Entwicklung besser planen, da weniger überraschende Regressionsfehler im Laufe der Entwicklung auftreten
3. Durch eine niedrige Fehlerdichte kann die Kontinuierliche Integration (Continuous Integration) möglich gemacht werden, und so der Kunde in den Entwicklungsprozess einbezogen werden

3.2. Ablauf

Ziel ist es, vor der Implementation eines Codes, einen Unittest zu implementieren. Davon ausgehend soll der geringstmögliche Code implementiert werden, damit der Test besteht. Alles Dritttest folgt die Refaktorisierung, bei TDD als Designphase genutzt.

Im Detail sind das also folgende Phasen, vgl. Abbildung 3:

1. Schreibe einen neuen [†]Test. Dies kann der erste eines neuen Features sein, oder aber ein Test, um Funktionalität zum aktuellen Feature hinzuzufügen
2. **Red:** Führe alle Tests aus, um sicherzugehen, dass der Test fehlschlägt. Andernfalls ist der Test überflüssig.
3. **Green:** Nachdem der Test fehlschlägt, implementiere nun den einfachsten Code, damit der Test besteht
Dies kann ausdrücklich auch eine Fake-Implementierung sein, also z.B. die Rückgabe

¹⁰Schaffen-, Tätigkeitssrausch

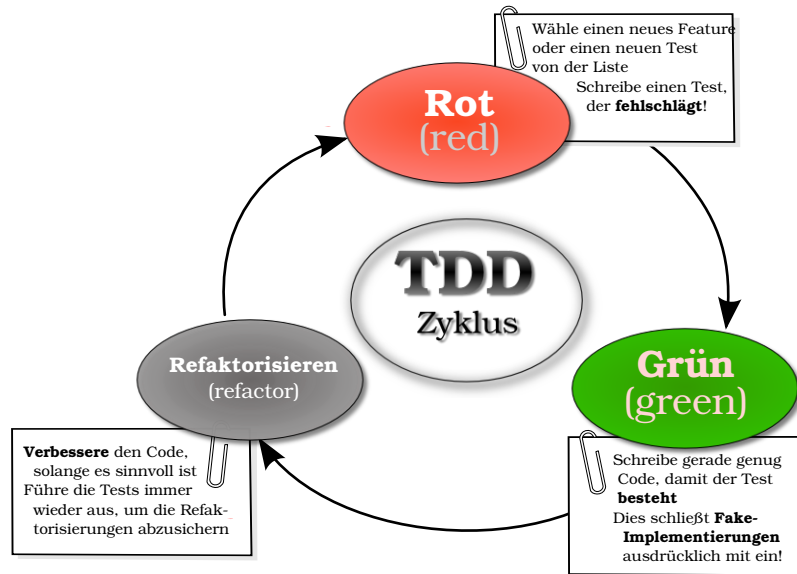


Abbildung 3: Red-Green-Refactor: Der TDD Entwicklungszyklus

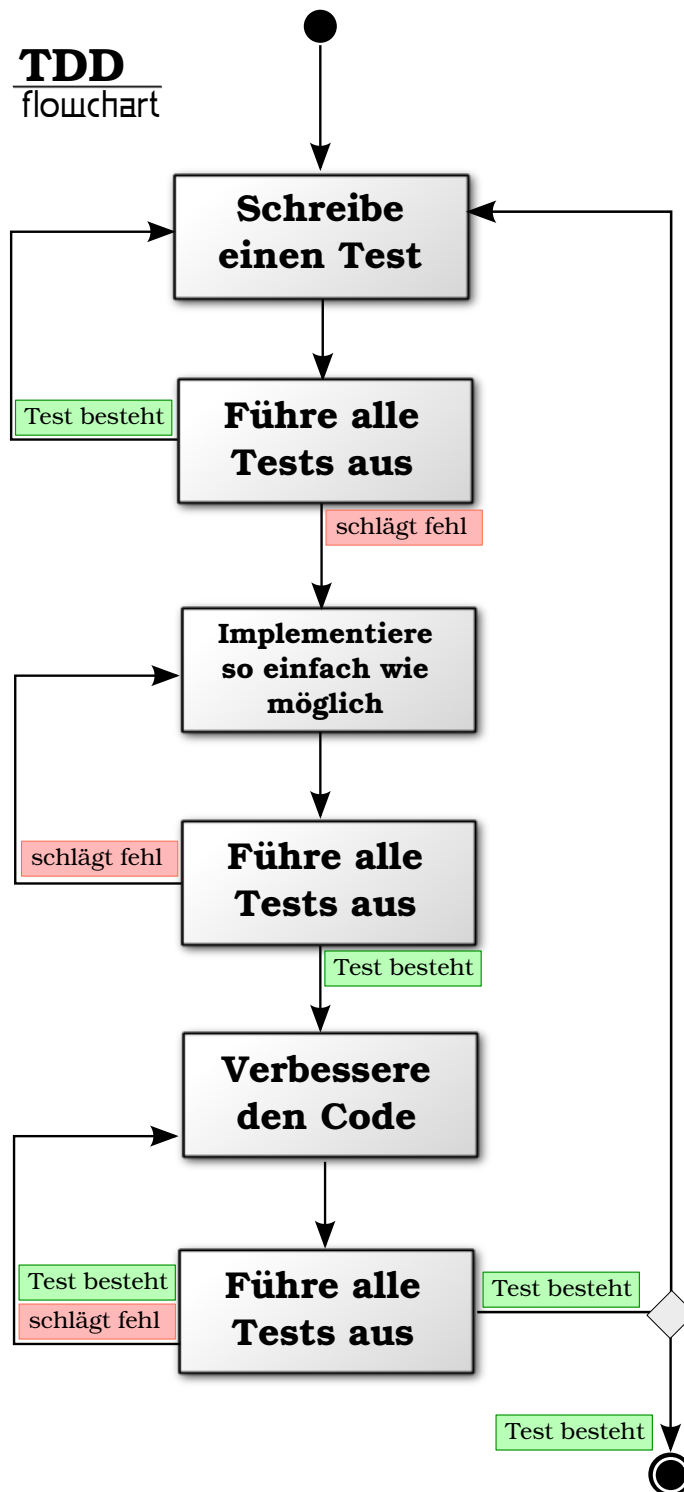
eines konstanten Wertes anstelle einer Berechnung. Wichtig ist, dass diese Phase so schnell wie möglich verlassen wird.

4. **Refactor**: Nachdem der Test bestanden wird, folgt nun die **wichtigste Phase**, die Refaktorisierungsphase.

Da wir bereits einen Test haben, der unser gewünschtes Systemverhalten widerspiegelt, können wir gefahrlos [↑]Refaktorisieren, d.h. meist Duplikation eliminieren. In dieser Phase findet das Design des Codes statt. Man macht sich Gedanken, wie die vorhandenen Klassen optimal refaktoriert werden können, um [↑]Bad Smells zu eliminieren, und welches Entwurfsmuster ggf. angewendet werden kann.

Ein genau spezifizierter Ablauf ist in Abbildung 4 zu finden. Auch dort ist zu sehen, dass die Testerstellung und Refaktorisierungsphase strikt getrennt sind. Innerhalb ersterer solle nur möglichst schnell ein funktionierender Test erstellt und zum Bestehen gebracht werden. Die eigentliche Arbeit findet dann innerhalb der Refaktorisierungsphase statt, in der die wahrscheinlich suboptimale Implementierung verbessert wird, indem iterativ Design hinzugefügt wird.

Jeder Unittest soll prinzipiell nur eine Eigenschaft testen, die Entwicklung erfolgt also



in kleinen Schritten. Dies hat direkte Auswirkungen auf die zu entwickelnden Objekte und Methoden, die ebenfalls übersichtlich werden sollen, und somit dem Objektbegriff, eine Klasse für eine Aufgabe, gerecht werden.

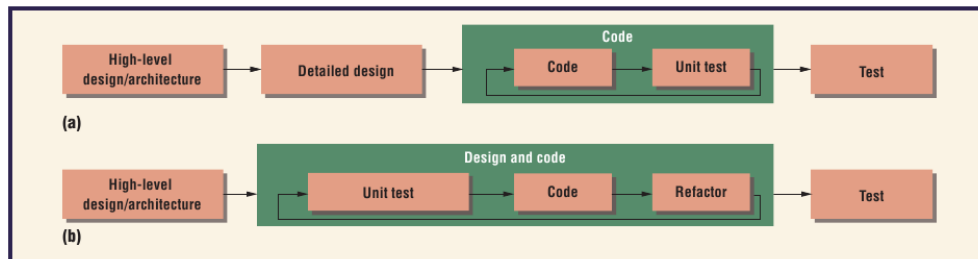


Abbildung 5: Entwicklungsablauf

(a) Traditionelle Entwicklung, b) Testgetriebene Entwicklung

Quelle: [Janzen and Saiedian, 2008]

Aus der Sicht des gesamten Entwicklungsprozesses wird somit die klassische Design-Phase in die Implementation eingegliedert. In Abbildung 5 wird der klassische Entwicklungsablauf dem Ablauf bei einem TDD-orientierten Ablauf verglichen. Demnach wird die (Fein-)Design Phase scheinbar aus dem Ablauf entfernt, und findet sich als Refaktorisierungsphase wieder. Wichtig zu bemerken ist außerdem, dass TDD keineswegs auf eine Analysephase/einen Grobentwurf oder finale (System)-Tests verzichtet, sondern nachwievor auf diese angewiesen ist.

It's about the design, not the tests.

Kent Beck

3.3. Sonderfälle

Der oben gezeigt Ablauf ist für den Normalfall, dem Entwickeln eines neuen Features gedacht. Für einige gesonderte Problemstellungen im Programmieralltag existieren ebenso gewisse Abläufe.

Fehlerbehebung Falls trotz der Verwendung von [†]TDD Fehler in der Software gefunden werden, so erfolgt:

1. Schreibe einen Test, der den Fehler auslöst bzw. nachbildet
2. Behebe den Fehler im Programmcode
3. Führe alle Tests aus

. Somit wird sichergestellt, dass jeder bisher gefundene Fehler durch einen Test abgesichert wird.

Spikes oder Spike Solution In einigen Fällen ist es nicht ratsam, sofort mit einer testgetriebenen Entwicklung zu beginnen. Gerade wenn Prototypisierung, d.h. schnelle, erforschende, explorative Entwicklung mit dem ausschließlichen Ziel schnell ein lauffähiges Ergebnis zu erhalten gewünscht ist, dann kann auf Tests verzichtet werden. Ein solches, isoliert entwickeltes Experiment wird im [†]TDD-Jargon „Spike“ (zu deutsch: Spitze, Nadel) genannt. Die Idee dahinter ist es zu Lernen, und der produzierte Code wird i.d.R. gelöscht und nach dem Lernprozess nach originärer TDD-Manier entwickelt. Dies soll auch die gewählte Metapher einer Nadel aufzeigen: Schnell eine Nadel durch ein Brett bringen [Shore, 2007]. Ausführliche Informationen über das Wann und Wie eines Einsatzes von Spikes finden sie in dem Buch „The Art of Agile Development“ [Shore, 2007]. Der Autor hat sogar das betreffende Kapitel online verfügbar gemacht¹¹.

Testen von privaten Methoden/Attributen Da die objektorientierte Modellierung das Konzept des Information Hiding und Kapselung vorsieht, soll der interne Aufbau einer Klasse nach außen nicht sichtbar sein. Da Tests aber von außen auf eine Klasse zugreifen, stellt dies ein Problem dar. In einigen Sprachen löst man sich dieses Problem mittels Reflections, um über Umwege auf private Felder zuzugreifen. Dies spielt allerdings nur für das nachträgliche Testen von Legacy-Anwendungen eine Rolle. TDD in der Reinform betrieben habe niemals spezifische Tests von privaten Methoden oder Feldern, da diese ausschließlich durch Refaktorisierung entstanden sein könnten [Caroli, 2008].

3.4. Studien zu den Auswirkungen von TDD

Viele Studien belegen die positiven Effekte, die [†]TDD für die Software-Entwicklung hat.

So zeigt eine Fallstudie, dass TDD zu größerer Produktqualität bei gleichzeitig hoher Flexibilität führt, was ebenso in einer höheren Zufriedenheit bei den Programmieren führt [Wasmus and Gross, 2007].

¹¹http://jamesshore.com/Agile-Book/spike_solutions.html

Einer empirischen Studie von Madeyski zufolge, sei TDD schwierig zu erlernen und in einigen Metriken (Klassen pro Methode, Development Speed, Anteil der bestandenen Akzeptanztests) nicht signifikant besser als traditionelle Test-Last-Methoden. Allerdings hatten die beteiligten TDD-geführten Projekte eine signifikant bessere Testabdeckung und geringere Kopplung unter den Klassen. Test-First sei letztendlich eine mächtige aber kontraintuitive Technik [Madeyski, 2009].

In dem Artikel des IEEE-Software-Journals stellen Janzen und Saiedian eine Studie vor, die akademische und industrielle Javaprojekte die testgetrieben durchgeführt wurden (Test-First), mit denen, bei denen hinterher getestet wurde (Test-Last), vergleicht. Demnach zeigen die Ergebnisse an, dass Programmierer, die einen Test-First Ansatz verfolgen, tendenziell Software in kleineren Einheiten die weniger komplex sind, schreiben, als solche die erst nach der Entwicklung testen [Janzen and Saiedian, 2008].

Einer Studie von Müller zufolge, führt Testen im Allgemeinen zu weniger Methoden und geringerer Kopplung. Der Autor stellt auch eine potenzielle Metrik vor, um statistisch signifikant Projekte, die nach TDD betrieben wurden, von traditionellen Projekten zu unterscheiden: Assignment Controllability¹² [Müller, 2006]. Allerdings rät der Autor zu weiteren Untersuchungen und setzt auch keinen Grenzwert an, ab welchem Grad der Controllability ein Projekt als TDD-Projekt bezeichnet werden kann.

Auswirkungen
auf die Struktur
– Assignment
Controllability

Einer Umfrage unter 25 IT- und Entwicklungsleitern ergab, dass diese zwar die positiven Effekte unterstützen, aber nur 16% TDD in der Praxis einsetzen, und nur 21% Testvollständigkeit messen. Auch verstehen anscheinend etwa die Hälfte der Befragten den Begriff TDD falsch, nämlich als die reine Praxis Tests für alle denkbaren Problemfälle zu schreiben [Inc., 2007].

Verständnis von
TDD in der
Industrie

Das unter den Entwicklungsleitern der Fortune 500 Firmen, die von sich selbst behaupten, TDD zu betreiben, einige von Fehlannahmen ausgehen, wird in dem oben genannten Artikel genannt [Janzen and Saiedian, 2008]. So setzen diese TDD mit automatisierten Tests gleich, oder behaupten sogar TDD sei das Schreiben ALLER Testfälle vor der Implementationsphase, anstelle der eigentlich gedachten kurzen Entwicklungs-Iterationen [Janzen and Saiedian, 2008].

Einer Studie von Microsoft ergab, dass TDD-entwickelnde Teams eine 60% – 90% geringere Fehlerdichte, aber eine 15% – 35% längere Entwicklungszeit hätten, als nicht-TDD Entwickelnde [Nagappan u. a., 2008].

Auswirkungen
auf die
Produktivität

¹²Dies ist ein Maß, inwieweit der lokale Zustand einer Klasse/Methode von außen durch Parameter beeinflusst werden kann

3.5. Prinzip des Emergent Design – Evolutionäres Softwaredesign

Ein Konzept das TDD ermöglicht, ist das des sich Herausbildenden Designs. Gegenüber traditionellen Entwicklungsansätzen erfolgt die Entwicklungsphase (Design) hierbei nicht als eigenständige Phase, sondern ist streng in den Entwicklungsprozess integriert. Immer wenn ein Zyklus beim Refaktorisieren angelangt ist, findet effektiv Design statt. Eine Entwicklung nach TDD sucht den minimalsten Code, der die Anforderungen (Tests) erfüllt. Analog dazu, will ein Emergentes Design die kleinste Menge an benötigten Design suchen, im Gegensatz zu einem Software-Design, das im Vorfeld bedacht wurde. Durch die vielen Iterationen und die darauffolgenden zahlreichen Refaktorisierungen tritt nach und nach das Design hervor, welches optimal für das System ist.

Einige Software-Architekten (Ford, Reeves und Vanderburg) proklamieren, dass praktisches Software-Engineering eigentlich keine ordinäre Ingenieursdisziplin sei. Ingenieure sind für die Planung und das Design verantwortlich, welches in aller Regel im Vorfeld der Implementierung stattfindet. Das klassische Software-Engineering empfindet diesen Prozess nach, was sich z.B. in dem Wasserfallmodell äußert. Diese Autoren äußern nun, dass dies für die meisten Softwareprojekte nicht ideal sei, da sich die Businessanforderungen meist im Laufe einer Entwicklung ändern. Traditionelles Softwaredesign abstrahiere zu früh und spekuliere ohne die letztendlichen Fakten zu kennen. Traditionelle Ingenieurswissenschaftliche Disziplinen hätten außerdem die Beschränkung, dass ein Build-Prozess äußerst teuer ist (man denke an Brücken oder Chips in Flugzeugsystemen), wohingegen dies bei Software in der Regel fast nichts koste. Dadurch kann die Software-Entwicklung ein iteratives Vorgehen nutzen [Ford, 2010b], [Vanderburg, 2010], [Reeves, 1992]. Statt eines großen Designs am Anfang (BDUF – Big Design Upfront) soll das Design durch Entdeckung und Extrahieren aus dem Sourcecode gewonnen werden.

Abstraktion erfolgt beim Emergent Design durch Refaktorisieren, dem Anwenden von Entwurfsmustern und dem Anwenden von sogenannten **Idiomatischen Mustern** (Idomatic Patterns). Im Gegensatz zu den klassischen Entwurfsmustern sind diese Muster nicht allgemein und für jedes Projekt anwendbar, sondern spezifisch für das Projekt, die Sprache oder das Unternehmen. Dies umfassen z.B. wie Transaktionen, Sicherheit und Infrastruktur verwendet wird. Viele Frameworks seien eine Sammlung von idiomatischen Mustern für eine bestimmte Aufgabe, und viele gute Frameworks seien aus alltäglichen Problemlösungen hervorgegangen, u.a. auch [†]Ruby on Rails [Ford, 2010a].

Ein Emergent Design kann auch ohne TDD in evolutionären Entwicklungsprozessen verwendet werden. Allerdings ist die Umsetzung ohne das Vorhandensein einer guten Test-Suite ein risikoreiches Unterfangen. Mit seiner iterativen Herangehensweise passen Emergent Design in Kombination mit TDD dagegen perfekt in den Entwicklungszyklus der Agilen Software Entwicklung.

3.6. Varianten

3.6.1. ATDD – Acceptance TDD

Die Akzeptanztest-getriebene Entwicklung ist eine Modifikation von [†]TDD. Statt der Unittests, stehen hier die Akzeptanztests im Vordergrund

Das ganze lässt sich auch in den übergeordneten Prozess zum Entwickeln eines Features einordnen.

1. Schreibe einen Akzeptanztest/Systemtest um das aktuelle Feature zu implementieren
2. Implementiere die Teilschritte, die notwendig sind, um den Akzeptanztest bestehen zu lassen. Verfahre bei der Implementation nach dem TDD Schema zur Implementierung der benötigten Units wie oben.
 - a) Schreibe einen Unittest
 - b) Prüfe, ob der Test fehlschlägt, andernfalls zurück zu 1.
 - c) Implementiere mit so wenig wie Code möglich, so dass der Test besteht
 - d) Refaktorisiere
3. Nachdem der Akzeptanztest besteht, Prüfe etwaige Refaktorisierungen für die Anwendungsebene.

Somit werden 2 Testebenen erstellt, die Akzeptanz- und die Unittests.

3.6.2. Behavior Driven Development

Die Behavior-Driven-Development ist eine Variation der Testgetriebenen Entwicklung, die die Kommunikation mit dem Kunden priorisiert. Sie sucht ein gemeinsames Vokabular für Business und Technologieinteressen.

Während bei klassischen Unit-Tests, und damit auch bei TDD, die Begriffe „Zusicherungen“ (assertions) und „Tests“ beinhaltet, so hat BDD stattdessen „Erwartungen“ (expectations) und „Spezifikationen“ (specs/specifications), und verwendet oft das Modalverb „sollte“ (should).

3.6.3. Design Driven Testing

Design Driven Testing soll eine Umkehrung von Testgetriebener Software sein und wird Stephens und Rosenberg als Alternative dazu vorgeschlagen [Stephens and Rosenberg, 2010]. Sie kritisieren, dass TDD in Reinform betrieben, lediglich Unittests, aber keine Dokumentation oder höhere Tests höherer Levels produziert. Weiterhin monieren sie, dass TDD zu schwierig und aufwändig sei. Sie schlagen vor, stattdessen die Tests durch das Software-Design steuern zu lassen und sich auf komplexe Code-Abschnitte zu konzentrieren, anstatt wirklich jeden Code durch einen vorausgegangen Test entstehen zu lassen. Sie proklamieren die Nutzung von Akzeptanz- anstelle der Unittests. Code-Qualität soll durch ein gründliches vorheriges Design anstelle nachträglicher massiver Refaktorisierungen bewerkstelligt werden. DDT eignet sich für größere Teams, da Wert auf manuelle Tests gelegt wird und z.B. ein QA-Team eingebunden wird. Da das Projektteam von it-jobs-und-stellen.de ein sehr kleines ist, wird auf diesen Entwicklungsprozess nicht näher eingegangen.

4. Die Programmiersprache Ruby

Ruby ist eine Programmiersprache, die ab 1993 von Yukihiro Matsumoto entwickelt wurde. Dabei ließ er sich von seinen Lieblingsprogrammiersprachen Perl, Smalltalk, Eiffel, Ada und Lisp inspirieren, um eine neue Programmiersprache zu entwickeln, die sowohl funktionale und imperative Programmierung ermöglicht [Team, 2011].

Eine vollständige Einführung in Ruby zu geben würde den Rahmen dieser Diplomarbeit sprengen, weswegen ich mich auf die Herausstellung der Hauptmerkmale und Unterschiede zu anderen Sprachen konzentrieren werde, und was die Auswirkungen auf das Testgeschehen sind.



Ruby ist eine Multiparadigma-Sprache

4.1. Einführung in Ruby

Ruby ist eine interpretierte Sprache, auch Skriptsprache genannt. Dies heisst, dass der Programmcode zur Laufzeit analysiert und ausgeführt wird. Ruby ist auch eine Multiparadigma-Sprache, die Objektorientierung, prozedurale und funktionale Programmierung unterstützt.

Prozedural Funktionen und Variablen können außerhalb von Klassen definiert werden, in dem sogenannten „main“-Objekt

Objektorientierung Alle Datentypen sind Objekte. Alle Variablen beinhalten Referenzen auf ein Objekt. Dies betrifft auch die primitiven Datentypen wie Integer und String

Funktional Anonyme Funktionen und Closures sind Sprachbestandteil. Alle Statements haben einen Rückgabewert. Innerhalb einer Funktion ist dies immer das letzte Statement, falls kein expliziter Rücksprungpunkt gesetzt wurde

Das Ziel von Ruby ist es nicht nur Maschinenlesbarkeit zu sein, sondern vor allem die Lesbarkeit durch Menschen zu verbessern. Dies drückt sich durch eine Syntax, die oft laut als englische Sprache gelesen werden kann, aus, und den Einsatz von Sonderzeichen reduziert. Auch hält Ruby eine Vielzahl von redundanten Keywords bereit (Syntaktischer Zucker), um dem Programmierer mehrere Wege zur Lösung seines Problems zu ermöglichen.

Im Nachfolgenden einige Beispiele für die Verwendung von Ruby, insbesondere die „Alles ist ein Objekt“-Philosophie.

Ruby is simple in appearance, but is very complex inside, just like our human body


```
----- Interaktive Ruby Sitzung (IRB) -----
1  >> 2.even?
2  => true
3  >> "hallo".upcase
4  => "HALLO"
5  >> Date.today + 2
6  => #<Date: 2011-06-30>
7  >> a = 4 + Math.sqrt(9)
8  => 7.0
9  >> if (0..10).include? a
10 >>   puts "a liegt zwischen 0 und 10"
11 >> end
12 => a liegt zwischen 0 und 10
```

Listing: Ruby Beispiele

In den ersten beiden Beispielen sieht man, dass Integer und String Objekte sind, und über Memberfunktionen verfügen. Im ersten Beispiel wird geprüft, ob die Zahl gerade ist. Dabei existiert eine Konvention, dass boolsche Methoden mit einem Fragezeichen am Ende notiert werden. Im Dritten Beispiel wird eine Klassenmethode „today“ auf die Klasse „Date“ ausgeführt, welche ein Datumsobjekt konstruiert und zurückliefert. Da auch die Nutzung von Operatoren letztendlich in Methodenaufrufe gemappt werden, wird die Methode „+(2)“ auf dieses Objekt ausgeführt, und liefert ein neues Datumobjekt, welches sich um 2 Tage unterscheidet, zurück. Im Vierten Beispiel wird der Einsatz von Variablen demonstriert. Das letzte Beispiel zeigt den Einsatz von Kontrollstrukturen. Als Besonderheit seien hier auf die Range „(0..10)“, die ein Intervall für den Integerzahlenbereich von 0 bis einschließlich 10 liefert. Die Methode „include?(a)“ tests nun, ob die Variable „a“ in diesem Intervall existiert. Bei Eindeutigkeit können die Klammern eines Methodenaufrufes weggelassen werden.

Weiterhin erlaubt Ruby die Arbeit mit Lambdas, also anonymen Funktionen. Eine beliebte Verwendungsmöglichkeit ist die Bearbeitung von Arrays und listenähnlichen Strukturen.

```
----- IRB -----
1  >> adder = lambda { |a,b| a + b }
2  >> adder.call(1,2)
3  => 3
4
5  # Sortiere nach Standardvergleichsoperator
6  >> [4,5,7,3].sort()
7  => [3, 4, 5, 7]
8
```

```
9 # Es kann auch eine benutzerdefinierte Sortierfunktion
10 # angegeben werden
11 >> [ "string", "rails", "ruby" ].sort_by{ |item| item.length }
12 => ["ruby", "rails", "string"]
13
14 # Die Quadratzahlen von 1 bis 5
15 >> (1..5).map{|element| element * 2}
16 => [2, 4, 6, 8, 10]
```

Listing: Ruby Beispiel: Blöcke

Typ- und Objektsystem Wie schon erwähnt, sind bei Ruby alle Datentypen ein Objekt. Dies schließt insbesondere Klassen und primitive Datentypen mit ein, wie wir wiefolgt sehen können

```
1 >> 2.class
2 => Fixnum
3 >> Fixnum.class
4 => Class
5 >> Class.class
6 => Class
7
8 >> Fixnum.superclass
9 => Integer
10 >> Fixnum.ancestors
11 => [Fixnum, Integer, Precision, Numeric, Comparable, Object, Kernel]
```

Listing: Klassenhierarchien

Das Literal „2“ ist somit ein Objekt vom Typ Fixnum. Die Klasse Fixnum ist ihrerseits vom Typ „class“. Da Ruby sowohl (Einfach-)Ableitungen als auch sogenannte Includes oder Mixins unterstützt, kann eine Klasse auch eine Menge von Oberklassen haben. Die gezeigte Klasse Fixnum verfügt somit standardmäßig sogar über 7 Oberklassen.

Ruby ist dynamisch stark typisiert, d.h. dass die Zuweisung des Typs einer Variable zur Laufzeit des Programms geschieht. Der Typ einer Variable ergibt sich damit aus ihrem beinhalteten Objekt. Durch die starke Typisierung ist es aber nicht möglich, invalide Operationen auf inkompatible getypte Daten auszuführen, beispielsweise eine Addition von Integer mit String. Rubys Typsystem ist „Duck-typed“, d.h. dass die Semantiken eines Objekts nicht durch seine Klasse und Ableitungshierarchie, sondern seinen Methoden und Attributen bestimmt wird.

”When I see a
bird that walks
like a duck and
swims like a duck
and quacks like a
duck, I call that
bird a duck.”

James Whitcomb
Riley

Ruby verfügt über eine lexikalische und dynamische Bindung¹³, letztere wird allerdings seltener verwendet. Im der Basissyntax verwendet Ruby statische Bindung. Es gibt eine im Ruby-Core enthaltene Bibliothek „Dynamic“ zum dynamischen binden.

Reflektion und Introspection Wichtig anzumerken sei noch, dass Klassen in Ruby nie geschlossen sind, sondern jederzeit erweitert werden können und vorhandene Methoden überschrieben werden können. So ist es z.B. möglich, die String-Klasse um eigene Funktionen zu erweitern. Ruby „merkt“ sich allerdings die überschriebenen Methoden und ein Aufruf der überschriebenen Methoden ist stets mittels „super“ möglich.

```
IRB
1 >> class String
2 >>   def remove_whitespace
3 >>     self.gsub(/\s+/, "")
4 >>   end
5 >> end
6
7 >> "Dies ist ein Test".remove_whitespace
8 => "DiesisteinTest"
```

Listing: Ruby Beispiel offene Klassen

Generische Programmierung und Aspekte der Metaprogrammierung Metaprogrammierung umfasst die Analyse, Transformation und Generierung von Objektprogrammen durch Metaprogramme [Herrmann, 2005]. Dies ermöglicht es Probleme effektiv zu lösen, die andernfalls nur mit erheblichem Aufwand, oder gar nicht zu lösen sind. Ein beliebtes Idiom innerhalb der Ruby-Community ist es, verwendete Methoden auf Basis des Methodennamens zur Laufzeit zu erstellen (Generierung). So verwendet das beliebte Objektrelationale Datenbankframework ActiveRecord dies, um einfache SQL-Statements zu erstellen (Ruby on Rails verwendet standardmäßig ActiveRecord als Schnittstelle zur Datenbank.).

```
IRB
1 >> Person.find_by_first_name("Stefan")
2 # Person Load (0.2ms)  SELECT persons.* FROM persons
3 #   WHERE users.first_name = 'Stefan' LIMIT 1
```

Listing: Demonstration von generischen

¹³ **Static Scoping:** Variablen werden zur Compilezeit gebunden ohne den aufrufenden Code zu berücksichtigen

Dynamic Scoping: Variablen-Bindung kann nur im Moment der Ausführung des Codes festgestellt werden

Die Methode `find_by_first_name` existiert nicht, und wird zur Laufzeit auf Basis des Namens gebaut. Dies ist möglich, da Ruby sogenannte Hooks (Callbacks) bereitstellt. Dies sind z.B. `method_missing`, der aufgerufen wird, wenn ein nicht-existierende Funktion gerufen wird (wie in unserem Beispiel „`find_by_first_name`“), oder „`method_added`“, der aufgerufen wird, wenn in einer Klasse eine Methode definiert wird¹⁴.

Ein weiteres Beispiel ist die Definition der relationalen Beziehungen zwischen den einzelnen Modellen innerhalb von [↑]Ruby on Rails.

```
app/models/job.rb
1 class Job < ActiveRecord::Base
2   belongs_to :user
3 end
```

Listing: Nutzung von Metaprogrammierung zur Erstellung von Objektbeziehungen

Hiermit definieren wir, dass ein Job einem User gehört, es also eine 1:N (oder 1:1)-Beziehung zwischen beiden gibt. Die dafür benötigten Getter und Setter werden mittels des Methodenaufrufs „`belongs_to`“ in die Klasse Job geschrieben.

Diese Beispiele sollten als kurzer Einstieg in Ruby dienen, und einen Querschnitt durch die Besonderheiten der Sprache aufzuzeigen.

Für eine weitere Vertiefung sei das Buch „Programming Ruby 1.9“ empfohlen, das im Detail auf die neueste Version der Programmiersprache eingeht [Dave Thomas, 2009a].

4.2. Diskussion

Ruby als Skriptsprache Dynamisch typisierte Sprachen, wie Ruby, haben gegenüber klassischen statisch-getypisierten Sprachen einige Nachteile. Oft wird der Geschwindigkeitsnachteil angesprochen, den der Prozess des Interpretierens und die dynamische Typisierung verursachen. Der genaue Faktor variiert allerdings je nach Algorithmus und Implementierung, stark. Ein beliebter Benchmark, „shootout.alioth“, vergleicht beliebte Algorithmen der Informatik implementiert in verschiedenen Sprachen miteinander. So ergibt sich z.B. in der Gegenüberstellung von Ruby mit C ein 4 bis 300-fache langsamere Ausführungszeit. Dem gegenüber steht allerdings nur die Hälfte bis 1/7 der Menge an benötigtem Code [Game, 2011], was ein Zeichen für die Ausdruckstärke von Ruby ist. Wichtig ist auch die verwendete Laufzeitumgebung. Neben der Referenzimplementierung von Matsumotu (Ruby MRI), existieren noch JRuby,

¹⁴Einen guten Überblick über die Callbacks die Ruby bereitstellt, und was man damit kann, finden sie hier: [method_added](#)

eine Implementierung auf der Java Virtuellen Maschine, und Rubinius. Die letzten beiden unterstützen auch eine sogenannte Just-in-time (JIT)-Kompilierung zur Verbesserung der Performanz bei längerer Ausführungszeit. Desweiteren gibt es gerade im Bereich Performanz viel Bewegung innerhalb der Ruby-Implementierungen, und fast alle Ruby-Implementierungen nehmen stetig zu [Cangiano, 2010].

Ein Vorteil des Interpretierens, also der Übersetzung zur Laufzeit, ist eine hohe Plattformunabhängigkeit und ein leichter Buildprozess, da das Kompilieren entfällt. Verfechter dynamischer Sprachen erklären weiterhin, dass diese sich ideal für prototypische Implementierungen eignen, da sich Anforderungen ständig ändern können. Weiterhin hätten Programme dynamischer Sprache eine potenziell hohe Wiederverwendbarkeit und eine höhere Lesbarkeit [Meijer and Drayton, 2005] [Ousterhout, 1998].

Allerdings bleiben Fehler, die der Compiler bereits entdeckt hätte, bis zur Ausführung oder schlimmstenfalls noch länger unentdeckt. Dazu gehören z.B. Tippfehler, bei denen der Wert einer nicht deklarierten Variable ausgelesen wird. Im Gegensatz zu z.B. PHP, wirft Ruby aber dann eine Exception.

Auf das Testen hat dies eine direkte Auswirkung. Viele Meinungen belegen, dass eine dynamisch typisierte Sprache mehr Tests benötigt, als eine statisch typisierte [Spiewak and Harrop, 2010].

Schlussfolgerung Die Verwendung von Ruby und anderen dynamischen Sprachen birgt durchaus Risiken, die zu beachten sind. Falls man diese Risiken im Kopf behält, und die Möglichkeiten der Sprache nutzt, um die Lesbarkeit zu verbessern, sind sie gerechtfertigt. Gerade in der Entwicklung kleinerer Entwicklerteams oder Projekten mit engem Budget können dynamische Sprachen ihre Vorteile ausspielen, da sie eine schnellere Entwicklung ermöglicht. Im Gegensatz zu den meisten auf C basierten Sprachen, ist die Syntax von Ruby äußerst leserlich, da nur wenige Sonderzeichen verwendet werden. Auch ist Ruby sehr ausdrucksstark, da die Deklaration entfällt und es viel sogenannten syntaktischen Zucker gibt. Dies kann, richtig angewendet, der Lesbarkeit zuträglich sein.

Sometimes people jot down pseudo-code on paper. If that pseudo-code runs directly on their computers, it's best, isn't it? Ruby tries to be like that, like pseudo-code that runs.

Yukihiro Matsumoto

4.3. Testframeworks für Ruby

4.3.1. Test::Unit und Minitest

Test::Unit (Ruby 1.8.7) und Minitest (1.9.2) sind die Testbibliotheken, die Ruby standardmäßig mitbringt. Beide basieren auf dem xUnit-, bzw. SUnit-Design von Kent Beck, und sind für Nutzer von JUnit oder NUnit leicht nachvollziehbar.

Für eine zu testende Klasse wird eine analoge Testklasse erstellt. Diese trägt per Definition denselben Namen wie die zu testende Klasse mit einem „Test“ am Anfang. Um z.B. eine Klasse „job“ zu testen, wird eine Datei `test_job.rb` (Ruby Standard) oder `job_test.rb` (Rails Standard) erstellt. Dort wiederum wird eine Klasse mit Namen `TestJob` definiert.

Ein Beispieltest sieht z.B. so aus:

```
1 [language=Ruby,label=Testen mit Test::Unit,caption=Testen mit Test::Unit in Ruby]
2 require "job"
3
4 class TestJob < Test::Unit::TestCase
5   def setup
6     @job = Job.create
7   end
8
9   def teardown
10    Job.delete_all
11  end
12
13  def test_job_exists
14    @job.title = "Ruby on Rails in Entwickler"
15    @job.add_location_to_title( "Dresden")
16
17    assert_equal( "Ruby on Rails Entwickler in Dresden", Job.first.title)
18  end
19 end
```

Listing: Testen mit Test::Unit in Ruby

Unsere Klasse `TestJob` erbt von der `TestUnit` Basisklasse. Sie beinhaltet die Methoden „setup“ und „teardown“, die jeweils vor, respektive nach jedem einzelnen Testfall aufgerufen werden. In der Setup-Methode nehmen wir z.B. das Anlegen eines Jobs vor, in der Teardown Methode löschen wir alle Jobs in der Datenbank, um einen sauberen Test zu gewährleisten

Danach können nun beliebig viele Testmethoden folgen, deren Namen mit `test_` beginnen müssen. Jede Testmethode besteht in der Regel aus einer Initialisierung (kann in die `setup`-Methode ausgelagert werden), der Ausführung einer zu testenden Aktion und dem Prüfen der danach geltenden Eigenschaften mittels Assertions. Diese Zusicherungen sind z.B.:

`assert_equal(a, b)` prüft die beiden Parameter auf Gleichheit

`assert(a)` Prüft, ob `a` wahr ist

`assert_raise(exception, &block)` Prüft, ob innerhalb des übergebenen Codestückes eine Exception vom Typ `exception` geworfen wird

Natürlich lassen sich beliebige weitere Zusicherungen definieren. [†]Ruby on Rails z.B. definiert Zusicherungen, um zu testen, ob ein Objekt eine gültige Instanz hinsichtlich der definierten Validierungen ist (zum Thema Validierungen später mehr).

Testdatengenerierung Nachdem Testdaten einmal in zentraler Form definiert wurden, erledigt [†]Ruby on Rails das Management, d.h. Laden und Löschen dieser selbständig. Diese Art der Testdatenbereitstellung wird als `Fixtures` bezeichnet. Alternativ ist der Einsatz sogenannter `Factories` möglich, um schnell Objekte mit bestimmten Eigenschaften zu erstellen. In jedem Fall setzt Rails die Datenbank nach jedem einzelnen Test zurück, oder kapselt Tests innerhalb von Transaktionen, insofern die verwendete Datenbank dies unterstützt.

Zur Generierung von größeren Mengen an zufälligen Daten einer bestimmten Domäne (z.B. für Stresstests) existieren Lösungen. Mittels der [†]Gems „populator“ und „faker“ lassen sich beispielsweise eine beliebige Menge an gültig-anscheinenden Personendaten (Name, Vorname, Adresse, E-Mail-Adresse, Passwort,...) oder Blind-Texte generieren¹⁵.

4.3.2. Cucumber

Cucumber ist ein relativ neues Framework (2008), um mittels einer domainspezifischen Sprache verständliche automatisierte Tests zu schreiben. Dabei gibt es 2 Ebenen. In der obersten werden Tests in Englisch, Deutsch oder einer anderen der mehr als 30 unterstützten Sprachen spezifiziert. In der darunterliegenden werden diese Schritte in

¹⁵Eine sehr gute Erklärung der Nutzung ist im Railscast #128 zu finden <http://railscasts.com/episodes/126-populating-a-database>

echten Testcode implementiert. Im Folgenden sei ein Trivialbeispiel einer Anwendung, die Addieren unterstützt gezeigt.

Listing 8: Cucumber: Additionsfeature in Deutsch

```
# language: de
Funktionalität: Addition zweier Zahlen
  Hier würde eine grobe Beschreibung des Businessvalues
  und der Rahmenbedingungen stehen
  Szenario: Addition von ganzen positiven Zahlen
    Wenn ich "1" für a und "2" für b eingebe
      Und auf "Addieren" klicke
      Dann sehe ich "3"
```

Wenn man nun die Datei mittels Cucumber ausführt, so wird darauf hingewiesen, dass die Testschritte noch nicht implementiert sind. Eine Beispielimplementation (ohne Verwendung einer GUI-Anwendung) der Testschritte wäre:

Listing 9: Cucumber: Implementierung der Additionstestschritte in Ruby

```
Wenn /^ich "([^"]*)" für a und "([^"]*)" für b eingebe$/ do |arg1, arg2|
  @addierer = Addierer.new(arg1, arg2)
end

Wenn /^auf "([^"]*)" klicke$/ do |arg1|
  @result = @addierer.add()
end

Dann /^sehe ich "([^"]*)"$/ do |arg1|
  assert_equal( arg1.to_f, @result)
end
```

Wie man sehen können werden die Testschritte über Regex gemacht. Jeder Schritt kann nun eine beliebige Implementierung besitzen. Meist ist es entweder eine Initialisierung, eine Aktion oder eine Erwartung, ausgedrückt durch die Schlüsselwörter „Angenommen“, „Wenn“ und „Dann“, bzw. „Given“, „When“ und „Then“ im Originaldialekt. Die Einteilung in klare Testschritte fördert die Wiederverwendbarkeit der Testschritte selbst.

Der Vorteil von Cucumber ist nun, dass diese Feature-Datei zusammen mit dem Kunden durchgesprochen werden kann, und am Ende eine funktionale Validierung durchzuführen. Auch hilft es, nur diejenigen Features zu implementieren, die letztendlich einen Wert für das Geschäft haben werden.

In der Regel werden mit Cucumber Webanwendungen getestet. Dabei kann z.B. auch ein echter Browser ferngesteuert werden, um den Test so authentisch am echten Nutzungsprozess wie möglich zu orientieren.

Neben Ruby wird auch die Implementierung der Testschritte in JVM und .NET-Sprachen unterstützt.

4.4. Ruby on Rails

Für das Projekt IT-jobs-und-stellen.de soll das Webframework Ruby-on-Rails verwendet werden. Rails wurde 2006 von der Firma 37signals unter der Leitung von David Heinemeier Hansson entwickelt und erlangte seitdem eine wachsende Popularität. Rails inspirierte viele andere Frameworks, wie z.B. cakePHP, Groovy on Grails, Symfony und ASP.NET MVC.



Viele professionelle Websites, die meist als Startup begannen, setzen bis heute auf Rails. Darunter z.B. Yellow Pages, die Gelben Seiten der USA, Github, eine sehr beliebte Community für OpenSource Programmierer, Groupon, dem führenden Unternehmen bei Online-Gutscheinen und XING, einer deutschen Online-Community für Business-Kontakte [on Rails.org, 2011].

Im Folgenden werden die Grundprinzipien und -konzepte von Ruby on Rails näher erläutert.

4.4.1. Konzepte von Rails

Rails ist ein Applikationsframeworks für Webanwendungen, und basiert auf dem [†]Model-View-Controller (MVC) (Model-View-Controller Entwurfsmuster), welches eine 3-Schichten Architektur darstellt. Jede Schicht hat fest definierte Aufgaben. Diese bilden normalerweise ein Dreigespann, bei Rails „Ressource“, genannt. Im folgenden werden die Schichten kurz erläutert, und am Beispiel einer Ressource „Job“

Model In Klassen dieser Schicht werden Zugriffe auf die Persistenzschicht vorgenommen. Meist geschieht dies durch Ausführung von SQL-Befehlen auf eine relationale Datenbank. Innerhalb von Rails ist dies aber meist nicht notwendig, da das [†]Objekt-relationale Mapping (ORM)-ActiveRecord häufig verwendete SQL-Befehle abstrahiert. Auch die Geschäftslogik soll per Definition zu großem Teil in dieser Schicht erfolgen.

Für einen Job ist das ein Modell, welches die Datenbanktabelle „jobs“, anspricht, und z.B. die Attribute „titel“, „datum“ und „beschreibung“ besitzt. Dabei können auf diesem Level auch datenbankunabhängige Constraints definiert werden, z.B. dass ein Job nur dann gespeichert werden soll, wenn der Titel mindestens 20 Zeichen lang ist, und das Datum mindestens das heutige ist.

Controller Klassen dieser Schicht vereinigen Methoden, die von außen per HTTP

erreichbar sind. Diese Methoden kommunizieren mit den korrespondierenden Models und bestimmen, welche View im einzelnen ausgeliefert wird. Weitere Funktionen eines Controllers sind Authentifizierung und Autorisierung (Wer darf was).

Standardmäßig stellt Rails die CRUD¹⁶-Operationen bereit, welche in Form eines REST¹⁷

View Eine View ist in der Regel ein Stück HTML Code welches einem Model zugeordnet ist, das bei einer bestimmten Aktion dem Clienten ausgeliefert wird. Neben HTML ist auch Javascript oder XML eine mögliche Auslieferungsform. Für den Job wäre das eine View für die Liste aller Jobs, einen Job im Detail anzeigen sowie das Formular zum Anlegen und Bearbeiten eines Jobs.

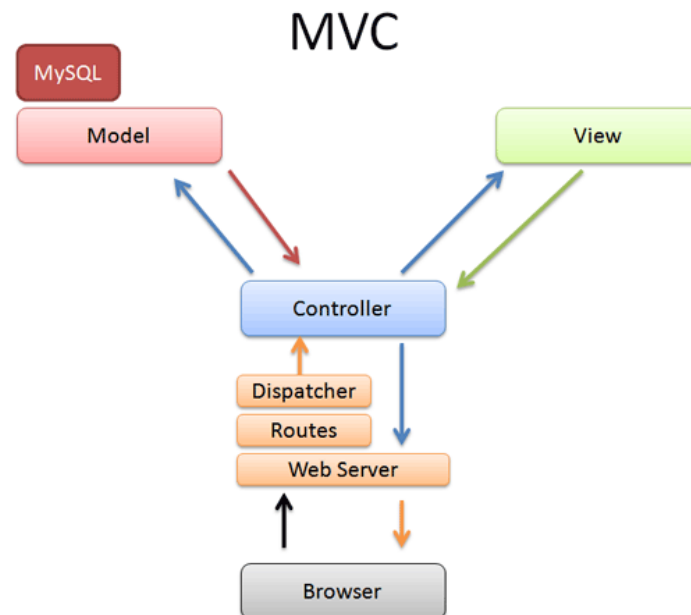


Abbildung 6: MVC Modell von Rails

Quelle: betterexplained.com

¹⁶Create Read Update Delete

¹⁷Representational State Transfer die HTTP-Methoden GET, POST, PUT, DELETE werden in Kombination mit einem definierten URL-Schema direkt auf die Aktionen **Auflisten**, **Anzeigen**, **Bearbeiten**, **Löschen**, **Neu anlegen** gemappt. http://en.wikipedia.org/wiki/Representational_State_Transfer

In Abbildung 6 ist der Ablauf einer Anfrage an den Server dargestellt. Die Anfrage des Browsers an die Website `http://localhost/jobs/12` wird über den Webserver, z.B. Apache2, an die Railsanwendung gestellt. Innerhalb von Rails wird dieser Anforderungsstring anhand der Routen, die die Anwendung anbietet, gematcht. In unserem Falle würde `/jobs/12` auf den Controller `jobs` aufgelöst werden. Innerhalb dieses Controllers wird eine Methode (Aktion) `show` erwartet. Diese Methode wird nun ihrerseits eine Anfrage an das Model `Job` stellen, den Job mit der ID 12 aus der Datenbank zu holen. Danach wird ein HTML Template zur Detailanzeige des Jobs generiert.

Neben diesem architektonischen Konzept verfolgt Rails noch andere Strategien, um das Entwickeln produktiver zu gestalten.

Convention over Configuration Rails ist so konzipiert, um als Framework komplett out-of-the-box zu funktionieren. Außer die Datenbankeinstellung wird keine Konfiguration im Vordergrund benötigt. Diese Methodologie zieht sich auch durch das Ökosystem durch. Die meisten externen Bibliotheken, bei Ruby Gems genannt, funktionieren bereits nach wenigen Kommandos. Dies macht das prototypische Entwickeln äußerst effektiv. Weiterhin ist die Struktur eines Railsprojektes extrem fest definiert. So gibt es u.a. einen Ordner `app`, mit den Model, Controller und View Dateien und einen Ordner `test`, der wiederum in `unit`, `functional`, `integration`, und `performance`, unterteilt ist. So finden sich Railsprogrammierer auch in fremden Projekten sofort zurecht.

Don't repeat yourself (DRY) Hier ist das Ziel, die Duplikation soweit wie möglich zu reduzieren, um bei Änderungen nur an einer Stelle ansetzen zu müssen. Ein Beispiel ist die Definition der Spalten der ORM's. Im Gegensatz zu anderen ORM-Frameworks ist dies bei Rails nicht notwendig. Rails erstellt automatisch Getter und Setter für die in der Datenbank definierten Tabellenspalten. Hintergrund ist, dass die Definition über Name und Typ der Attribute bereits in der Datenbank vorliegt, und eine Wiederholung im Quellcode dem DRY-Prinzip widerspräche.

REST Representational State Transfer ist eine Software-Architektur für HTTP-WebServices.

Dabei werden neben den Standard HTTP-Methoden GET und POST auch die selten benutzten Verben DELETE und PUT verwendet, um Aktionen auf einer Ressource zu definieren. Das Ziel ist ein sehr einfaches Design der URLs. Hier folgt die Auflistung der vier CRUD Operationen plus Auflistung von REST am Beispiel einer Ressource `jobs`, eines Webservice

Ruby on Rails is a breakthrough in lowering the barriers of entry to programming. Powerful web applications that formerly might have taken weeks or months to develop can be produced in a matter of days.

Tim O'Reilly,
Founder of
O'Reilly Media

GET /jobs.html Auflisten aller Jobs, Ausgabe als HTML Format

GET /jobs/12.xml Job mit der ID 12 anzeigen, Formatiere als XML

POST /jobs Einen Job anlegen. Alle benötigten Parameter, wie Titel, Beschreibung oder Datum sollten im POST-Body der HTTP-Anfrage enthalten sein

PUT /jobs/12 Den Job mit der ID 12 aktualisieren. Die Eigenschaften, die aktualisiert werden, müssen wiederum als Parameter mit übergeben werden

DELETE /jobs/12 Lösche den Job mit der ID 12

Rails macht das Arbeiten im Kontext dieser Architektur sehr einfach, und es gilt als die bevorzugte Methode in der Community, APIs zu bauen.

Codegeneratoren Rails bietet viele Codegeneratoren an, um schnell benötigte Klassen und Datenbanktabellen anzulegen. Im Railsprojekt reicht z.B.:

```
1 rails generate scaffold job title:string description:text \
2   start_date:datetime active:boolean user:references
```

Listing: Nutzung der Codegeneratoren von der Kommandozeile

Damit wird das Model Job, eine Erstellung der Tabelle „jobs“, ein Controller „jobs“ mit den REST-Standardaktionen und entsprechenden BeispielvIEWS, sowie Testfälle für Unit- und Funktionale Tests angelegt. Weiterhin sei zu bemerken, dass durch die Anweisung `user references` eine Spalte `user_id` angelegt wird und eine 1:n-Beziehung zum Modell „user“ hergestellt wird.

Full-Stack Webframework Rails bringt out-of-the-box alles mit, was zur Webentwicklung benötigt wird. Im Gegensatz zu anderen Webframeworks wurde für Datenbankankbindung, Templatesystem, Javascriptframework, Testframework und Webserver-API bereits eine Vorauswahl getroffen. Im aktuellen Rails 3.1 sind dies ActiveRecord, ERB, JQuery, sowie Test::Unit und Rack. Die meisten dieser Teil-Frameworks lassen sich zwar leicht austauschen, Rails selbst aber proklamiert „opinionated“, also rechthaberisch/eigensinnig, zu sein, und den Entwickler Standards vorzugeben [Hansson, 2011].

Eine Einführung und Programmierung in Rails soll nicht Bestandteil dieser Diplomarbeit werden. Für eine weitere Einarbeitung seien die folgende Quellen insbesondere empfohlen:

Rails for Zombies Dies ist ein moderner, interaktiver Onlinekurs. Greg Pollack und das Team von RailsEnvy verpackt die Lektionen in humorige interaktive Lerner-

Rails will have strong defaults. They might change over time but Rails will remain opinionated.

fahrungen. Jeweils eingeleitet durch ein Video muss der Teilnehmer Aufgaben direkt im Sourcecode lösen. Die Teilnahme ist kostenlos.

<http://railsforzombies.org/>

Agile Webdevelopment with Ruby on Rails Das quasi-Standardwerk, u.a. geschrieben vom [†]Ruby on Rails-Begründer David Hansson. Das Buch wird meist parallel mit einer neuen Rails-Version in einer neuen Auflage gedruckt, aktuell die Dritte [Dave Thomas, 2009b].

Rails Guides Die von Ruby-on-Rails herausgegebenen “Rails Guides,, sind eine gut strukturierte, kostenlose Online-Dokumentation, die nahezu alle Aspekte von Rails beleuchten.

http://guides.rubyonrails.org/getting_started.html

4.4.2. Diskussion

Nach einem kurzen Überblick über Rails, sollen nun die Eigenschaften des Frameworks diskutiert werden, und welche Auswirkungen sich dadurch auf das Testen ergibt.

Vorteile Dank der Modularität können als Persistenzgrundlage sowohl relationale Datenbank, wie MySQL, SQLite und Oracle, aber auch andere Formen, wie NoSQL-Datenbanken transparent verwendet werden. Dank einer einfach zu verstehenden Syntax, ist das Schreiben von SQL in 95% der Fälle überflüssig und zudem auch sicherer. Hier ein Beispiel, wie das Anlegen und Auslesen einer Instanz “Vertragsart,, erfolgt. Parallel dazu als Kommentar die SQL Kommandos, die ActiveRecord im Hintergrund ausführt.

```
365f7e
1 c = ContractType.new
2 c.name = "Vollzeit"
3 c.save
4 => false
5 # SQL (0.1ms) BEGIN
6 # SQL (0.3ms) SELECT 1 FROM 'contract_types'
7 # WHERE ('contract_types'.name = BINARY 'Vollzeit') LIMIT 1
8 # SQL (0.1ms) ROLLBACK
9 >> c.errors
10 => {:name=>["Name bereits vorhanden. Der Name muss einmalig sein"]}
11 >> c.name = "Praktikum"
12 >> c.save
13 => true
```

```
14 # SQL (0.1ms) BEGIN
15 # SQL (0.4ms) SELECT 1 FROM 'contract_types'
16 # WHERE ('contract_types'.name = BINARY 'Praktikum') LIMIT 1
17 # SQL (0.4ms) describe 'contract_types'
18 # AREL (0.4ms) INSERT INTO 'contract_types' ('name') VALUES ('Praktikum')
19 # SQL (7.1ms) COMMIT
```

Listing: Abstraktion von SQL-Befehlen

Hierbei ist auch schön zusehen, wie in Rails standardmäßig Transaktionen verwendet werden. Auch sieht man, wie eine Validierung funktioniert. Für das Modell VertragsArt wurde eine Einmaligkeit des Attributs name vereinbart. Dies prüft Rails vor dem Speichern und bricht das Speichern ab, falls eine Prüfung fehlschlug.

Die Implementation dieser Prüfung ist denkbar einfach.

```
app/models/contract_type.rb
1 class ContractType < ActiveRecord::Base
2   validates :name, :uniqueness => true, :presence => true
3
4   has_and_belongs_to_many :jobs
5 end
```

Listing: Nutzung von Validierungskriterien und DSL für Modellbeziehungen

Wir vereinbaren so, dass das Attribut Name einmalig sein muss (uniqueness), und ausgefüllt sein muss (presence). Die andere Zeile mit has_and_belongs_to_many definiert eine n:m Beziehung mit dem Modell Job, d.h. ein Job hat mehrere Vertragsarten. Der Rest geschieht dann durch Metaprogrammierung.

I needed to be way more productive...

David Heinemeier Hansson

Mit minimalem Code lassen sich so komplexe Probleme abbilden. Dies macht Rails zu dem hochproduktivem Framework, als das es entwickelt wurde.

Rails bietet eine gute Ausgangsbasis um sichere Websoftware zu entwickeln. Das Verwenden eines Datenbankframeworks macht SQL-Injections unmöglich. Cross-Site-Request-Forgery und Session-Angriffe werden erschwert, da Session und Cookie-Variablen standardmäßig verschlüsselt werden, und bei Nutzung der Formulargeneratoren ein CSRF-Token generiert wird, um Replay-Angriffe zu unterbinden.

Sicherheit

Nachteile Interaktion mit Legacy-Software ist nicht immer möglich. ActiveRecord reserviert ein paar Spaltennamen, wie `type` und `class`. Eine Benennung der Spalten sollte der Ruby-Namenskonvention entsprechen, also nur Buchstaben Zahlen und Unterstriche enthalten. Ansonsten können die Spalten nur über Umwege angesprochen werden.

Oft wird angeführt, dass Ruby als Skriptsprache und Rails als darauf aufbauendes Framework eine schlechte Performance hat, und dadurch ungeeignet für große Webanwendungen ist.

Performance

Andererseits gibt es Anzeichen dafür, dass eine clevere Architektur und Caching für skalierende Anwendung entscheidender ist, als die letztendliche Ausführungszeit.

Das dies möglich ist, zeigen z.B. Groupon, der führende Online-Coupon-Anbieter mit mehr als 50 Mio Abonnenten¹⁸ und Twitter, die jeweils Rails verwenden [on Rails.org, 2011].

Da Rails ein komplettes Webframework ist, wurden bei der Auswahl der einzelnen Komponenten bereits Entscheidungen getroffen. Erst jüngst gab es Kritik aus Teilen der Community, da ab der Version 3.1 „CoffeeScript“ und „SASS“ Bestandteil einer Rails-Distribution sind. Beide sind Zwischensprachen, die in Javascript respektive CSS kompilieren, und diese um Funktionalität erweitern.

Auswahl der
Komponenten

Rails und Tests Rails bietet ausgezeichnete Voraussetzungen zum Softwaretest. Dafür sprechen, dass...

- benötigte Bibliotheken bereits mitgeliefert werden. Dies umfasst einen Test-Runner, vorkonfigurierte Test-Datenbanken (auf Basis von SQLite) und das Testframework Minitest,
- die Verwendung stark erleichtert wird, da Rails beim Nutzen der Codegeneratoren analoge Testdateien gleich mitgeneriert,
- neben den mitgelieferten Tools das Rails Ökosystem eine Vielzahl von Testtools bereitstellt, u.a. Rspec (BDD¹⁹-Testframework), Rcov (Testabdeckung), diverse Mockbibliotheken (mocha, FlexMock, RR, Rspec Mocks), Tools zum Generieren und Bereitstellen von Testdaten (Fixtures, Factories, Faker) und Codemetriken (metric-fu)

¹⁸<http://www.socialshopping.com/Groupon/news/Groupon-hits-50m-Subscribers-Shopping-site-sensation-201101210398/>

¹⁹Behavior Driven Development, siehe Glossar

- das Testen einen sehr hohen Stellenwert in der Ruby und Rails-Community hat. Nahezu alle namhaften Ruby-Programmierer schreiben umfassende Tests [DeVries and Naberezny, 2008]. Das Resultat ist, dass auch fast alle OpenSource Bibliotheken (Gems) bei Ruby eine „solid suite of tests“ haben [DeVries and Naberezny, 2008].

Dabei werden mehrere verschiedene Testarten unterstützt und definiert.

Unittests oder Modelltests (model test). Zielstellung: Hier wird die Logik einer Modelklasse untersucht

Testziel: alle (komplexeren) Methoden die das Modell anbietet. Am Beispiel Job kann das die Aussage sein, wann ein Job gültig ist, d.h. welche Bedingungen für die einzelnen Attribute gelten sollen.

Testart: Whitebox

Funktionale Tests Untersuchungsgegenstand sind die Controller, also die Schnittstellen zum Nutzer.

Testziel: Getestet wird meist der Arbeitsablauf innerhalb eines Controllers, also Weiterleitungen, Benachrichtigungen und welches Template gerendert wird.

Testart: Whitebox Es können auch oberflächliche View-Tests unternommen werden, also die Aussage ob ein bestimmtes HTML-Element auf der Seite zu sehen ist.

Integrationstests es wird ein Browser simuliert der von außen auf die Applikation zugreift Zielstellung: Testen komplexer Interaktionen zwischen verschiedenen Teilen der Software

Beispiel: Ein User loggt sich ein und legt einen neuen Job an

Testart: Graybox oder Blackbox

Performanz-Tests Eine Testart, die alle Methoden aus Unittests und Funktionalen Tests beinhaltet

Zielstellung: Herausfinden von Performanz-Flaschenhälsen in allen Ebenen.

Beispiel: Es werden 1000 Jobs generiert und geprüft, ob die Anzeige schnell genug läuft

Testart: Graybox

5. Code-Metriken

Eine Codemetrik ist eine Maßzahl, die zum Vergleich dient und ein Qualitätsmerkmal für ein Stück Code oder ein Programm darstellt. Sie ist wird den Software-Metriken und Produkt-Metriken zugeordnet.

A function whose inputs are software data and whose output is a single numerical value that can be interpreted as the degree to which software possesses a given attribute that affects its quality

IEEE [1998]

Dem Verwenden von Code-Metriken liegt der Wunsch zugrunde, komplexe Codeteile auf einfache Zahlen automatisiert beurteilen zu lassen, um potenziell suboptimale Codestellen zu finden, welche Defekte verursachen könnten. Aus Business-Sicht stellen Code-Metriken eine Methode dar, Entwicklungsfortschritt zu messen und qualitativ zu beurteilen.

5.1. Überblick über Code-Metriken und Skalen

Hier seien nun einige der geläufigsten Code-Metriken vorgestellt.

Lines of Code (LOC) ist eine häufig verwendete, und die am leichtesten zu bestimmende Größe. Sie repräsentiert den Umfang eines Programmes. Diese Größe erhält dann eine größere Aussage, wenn man sie ins Verhältnis z.B. der Klasse oder eines Codefiles setzt. So kann man mit „LOC / Klasse“ schon diejenigen Klassen finden, die wahrscheinlich zu komplex sind.

Es werden alle Zeilen einer Datei gezählt, die nicht leer und keine Kommentare sind. Kommentare wiederum können als eigene Metrik verwendet werden, um den Grad der Quelltextdokumentation zu bestimmen.

Zyklomatische Komplexität ist ein Indikator für die Komplexität auf Basis des Kontrollflussgraphen eines Programms. Gemessen wird die Anzahl der linear unabhängigen Programmpfade. Sie ist für einen Graphen definiert durch:

$$M = E - N + 2P$$

E	Anzahl der Kanten
N	Anzahl der Knoten
P	Anzahl der verbundenen Komponenten

In einem normalen Programm ist die Zyklomatische Komplexität die Anzahl der Entscheidungspunkte + 1 [McCabe, 1976, S. 314].

Ein daraus abgeleitetes Testverfahren „Basis Path Testing“ schlägt vor, dass die Anzahl der Tests mindestens genauso groß sein sollte, wie die Grad der Komplexität [McCabe, 1976, S. 318]. Dadurch erreicht man Branch-Coverage (C1) siehe dazu weiter unten.

Anzahl Bad Smells ist eine aggregierte Metrik über alle suboptimale Codestellen. Da es viele verschiedene Bad Smells gibt, können ebensoviele Metriken davon abgeleitet werden. Für eine Übersicht genügt aber auch einfach die Summe. Welche Bad Smells für die Entwicklung entscheidend sind hängt von der gewählten Sprache, dem damit einhergehenden Programmierparadigma und manchmal auch den verwendeten Frameworks. Hier seien einige der häufig gebrauchten Smells für Ruby vorgestellt [Rutherford, 2010]:

Geringe Kohäsion insb. „Feature Envy“ (deutsch: Neid), ist für alle objektorientierten Programme anzuwenden. Eine Klasse weiß zuviel über die internen Strukturen einer anderen Klasse, und implementiert Funktionalität, die eigentlich in jene Klasse gehören würde. Im Beispiel würde diese Berechnung in die Klasse Checkout gehören.

```
@checkout.total = @checkout.total_price * MWST
```

Nichtssagender Name gilt für alle Programmiersprachen. Falls Bezeichner weniger als 3 Zeichen lang sind, oder Funktionen den Namen „do“ oder „run“ haben. Ausnahmen könnte man z.B. für die Schleifenvariable *i* rechtfertigen

Gesetz von Demeter bzw. die Verletzung desselben. Objekte sollten nur mit den Objekten in ihren unmittelbaren Nähe kommunizieren, und nicht etwa in Nachrichtenketten, wie z.B.:

```
@job.user.address.street
```

Beim Law of Demeter ist eine solche Kette bis maximal Länge 1 erlaubt.

Diese Smells können mit dem Tool reek²⁰ festgestellt werden.

Duplikation ist das Auftreten von gleichen oder ähnlichen Codeteilen an mehr als einer Stelle im Programm. Fortgeschritten Analysemethoden betrachten die Baumstruktur und finden ähnliche Teile unabhängig von Bezeichnernamen.

5.2. Code-Metriken für Tests

Tests haben (auch) die Aufgabe, ein Programm oder Codestück auf Korrektheit zu untersuchen. Die Tests allerdings haben ihrerseits i.d.R. keine Tests. Um also die Nützlichkeit der eigenen Tests zu bestimmen, kann man sich aber zumindest auf Code-Metriken stützen. Tests sind in erster Linie natürlich auch Code und können mit den oben genannten Metriken beurteilt werden. Zudem gibt es aber einige weitere exklusive Methoden, Qualität von Tests zu messen.

5.2.1. Verhältnis von Lines of Test zu Lines of Code

Neben den Lines of Code kann auf dieselbe Weise die Anzahl der Codezeilen der Testklassen ermittelt werden. Daraus ermittelt sich das Verhältnis:

$$R = \frac{\text{Lines of Code}}{\text{Lines of Test}}$$

$R \ll 1$ Falls es deutlich weniger Testzeile (LoT), als Codezeilen (LoC) geben sollte
so ist dies ein Indiz für zu wenige Tests

$R > 1$ Eine große Anzahl an Tests ist zwar wünschenswert, aber dies macht
keine Aussage über den Vollständigkeit oder die Qualität der Tests

²⁰<https://github.com/kevinrutherford/reek/wiki/Code-Smells>

Sollte dieser deutlich kleiner als 1 sein, so ist dies ein Symptom für zu wenige Tests. Diese Zahl ist von dem Testframework und dem Programmframework stark abhängig. Gute Projekte sollten mehr Test-Code, als Programmcode besitzen, um so die Zahl der Defekte gegen 0 zu reduzieren [Andrew Hunt, 1999][S. 238].

5.2.2. Testausführungsabdeckung

Die Testabdeckung, misst den Grad inwieweit ein Programm getestet wurde. Die Angabe erfolgt in Prozent, wobei 100% bedeuten, „das Programm wurde durch die Tests komplett ausgeführt“, und 0% „Das Programm wurde durch die Tests überhaupt nicht berührt“. Dabei wird die vorhandene Test-Suite ausgeführt und währenddessen der entsprechende Quellcode beobachtet. Es wird festgehalten, welche Anweisungen ausgeführt wurden. Allerdings gibt es 3 Abstufungen, diese Abdeckung zu beobachten (Mit steigender Komplexität des Messverfahrens):

- C0** (Anweisungsüberdeckung, Statement Coverage) ist die am einfachsten zu bestimmende Abdeckung. Dabei wird geprüft, ob jede Zeile des Quellcodes während der Codeausführung mindestens einmal ausgeführt wurde
- C1** (Zweigüberdeckung, Branch Coverage) prüft zusätzlich, ob jeder Zweig jeder Zeile ausgeführt wurde. Dies ist wichtig, falls man ternäre Ausdrücke²¹ verwendet
- C2** (Pfadüberdeckung, Path Coverage) prüft, ob jeder mögliche Codepfad durchlaufen wurde. Ein Codepfad sei eine einmalige Abfolge von Zweigen innerhalb einer Funktion von Eintritt bis Rücksprung [Cornett, 1996]. So werden z.B. bei 10 Bedingungen 1024 Pfade generiert, denen bei einer 100% Abdeckung auch 1024 Tests entgegenstehen müssten.

Anmerkung: In der Literatur startet in einigen Fällen die Nummerierung bei C0 [Powell, 2008], in anderen Fällen aber bei C1 [Cornett, 1996].

Für Ruby 1.8.7 gibt es das Tool `rcov`²², für Ruby ab 1.9.1 `simple-cov`²³, welche beide die C0 Testabdeckung bestimmen können. Zum aktuellen Zeitpunkt sind keine weiteren Tools bekannt, um C1 oder C2 Abdeckungen zu bestimmen.

²¹if-then-else in einer Zeile: `int a = (1==1) ? 5 : 3`

²²<http://relevance.github.com/rcov/>

²³<https://github.com/colszowka/simplecov>

Wieviel Testabdeckung ist sinnvoll oder notwendig Beim Messen der Abdeckung stellt man sich schnell die Frage, wieviel Testabdeckung notwendig ist. Zuerst sei die Art des Messverfahrens, also C0 bis C2, wichtig. Je komplexer das Messen erfolgte, desto geringer kann also die Testabdeckung am Ende ausfallen [Powell, 2008].

Falls dem TDD-Prozess minutiös gefolgt wurde, so ist die C0 Testabdeckung immer 100% [Beck, 2002]. Für ein Rails Projekt sei es auch relativ leicht, 100% oder nahe 100% zu erreichen [Rappin, 2011]. Die Zahl „100%“ sei für sich genommen nutzlos, aber sie zu erreichen, sei für den Prozess der Testgetriebenen Entwicklung nützlich [Rappin, 2011, S. 270]. Vielen Autoren bringen aber zum Ausdruck, dass es von der Situation abhängt [Elssamadisy, 2007] wieviel Testabdeckung sinnvoll ist. Test-Anfänger sollten sich zuerst überhaupt ans Testen gewöhnen, und erfahrene Entwickler sollte wissen, dass es keine einzige einfache Antwort auf diese Frage gebe [Elssamadisy, 2007]. Zudem gebe eine hohe Abdeckung keinen Aufschluss darüber, dass gut getestet wurde. Aber eine niedrige Zahl zeigt deutlich auf Missstände hin. Einem pragmatischen Ansatz von Savoia folgend, kann man aus dem Verhältnis der Zyklomatischen Komplexität mit der Testabdeckung eines Codestückes suboptimale Teile finden. Je mehr Verzweigung eine Methode hat, desto höher sollte ihre Testabdeckung sein [Savoia, 2007]. In einem Artikel empfiehlt Cornett eine Liste von Zielen, die es je nach Budget und Zeit zu erreichen gilt, beginnend damit, dass mindestens eine Funktion in 90% der Quelltextdateien durch die Tests aufgerufen wird bis zum finalen Schritt einer 100% Zweiüberdeckungs-Testabdeckung [Cornett, 1996].

Zusammenfassend kann man sagen, dass es keine eindeutige Antwort gibt. Eine niedrige C0 Abdeckung von 50% oder weniger zeigt allerdings deutliche Missstände beim Testverfahren an.

5.2.3. Mutations/Perturbationstests – Defect insertion

Eine weitere Methode, um die Qualität von Testcode zu messen, ist der Mutationstest. Dies ist ein diversifizierendes, fehlerbasiertes Testverfahren [Liggesmeyer, 1990]. Hierbei werden (automatisiert oder manuell) nacheinander alle Zeilen des Programmcodes geändert, und geprüft, ob danach ein Test fehlschlägt [Beck, 2002].

Für Ruby gibt es ein Werkzeug, „Heckle“,²⁴, welches dieses Verfahren implementiert. Im Detail werden Bedingungen negiert, konstante Zahlen und Funktionsaufrufe verändert, Zuweisungen verändert usw. [Sadists, 2010]. Dabei wird immer eine Änderung (Muta-

²⁴<http://ruby.sadi.st/Heckle.html>

tion) vorgenommen, und dann alle Tests ausgeführt. Sollten dennoch in einer Mutation kein Test fehlschlagen, dass ein Test fehle, so ist die Annahme des Testverfahrens.

5.3. Notwendigkeit von Code Metriken

Code-Metriken geben dem Programmierer automatisiert und schnell ein Feedback über die Qualität seiner Arbeit. Sie helfen dabei, Probleme frühzeitig zu erkennen und die Wartbarkeit durch gezielte Refaktorisierungen nachhaltig zu verbessern. Auch psychologische Auswirkungen dürfen nicht unterschätzt werden. Alleine der Fakt, dass Codemetriken in einem Unternehmen regelmäßig verwendet werden, motiviert den Programmierer keinen sogenannten „Big Ball of Mud“²⁵ zu schreiben. Insbesondere in kleinen Projektteams, die keine dedizierte Qualitätssicherung haben, sind Code-metriken als kostengünstiges Kontrollinstrument unerlässlich. Studien zeigen, dass der konsequente Einsatz von Code-Metriken und Analysebenchmarks die Fehlerdichte und Entwicklungskosten stark verringern kann [Baggen u. a., 2011, S.10f].

Für die Testgetriebene Software dient insbesondere in der Anfangsphase die Testabdeckung als Kontrollinstrument, um zu prüfen, ob der TDD-Prozess korrekt umgesetzt wird [Nagappan u. a., 2008, S. 300]. Außerdem sollte der zeitliche Verlauf der Metriken beobachtet werden, um Trends abzuschätzen und frühzeitig gegensteuern zu können. Für in [↑]TDD erfahrene Programmierer mag die Beobachtung der Testabdeckung nicht notwendig sein, für Einsteiger allerdings eine effektive Kontrollmöglichkeit.

Nach Erfahrungen in der pludoni GmbH sind Code-Metriken ein wichtiges Feedback-instrument, und unterstützen damit das Schreiben sauberen Codes. Wichtig ist, dass die Metriken regelmäßig berechnet werden, entweder als Cronjob oder nach jedem Einchecken in den Hauptentwicklungszweig der Versionsverwaltung, und in regelmäßigen Abständen von den Programmierern und Team-Leiter gelesen und besprochen werden. Allerdings besteht bei einer zu hohen Beobachtung der Metriken die Gefahr, eines Hawthorne-Effektes, d.h. dass die unter Beobachtung stehenden Programmierer ihr Verhalten den Code-Metriken anpassen, um optimale Ergebnisse zu erhalten [Langr and Ottinger, 2011, 52. Karte], und so nur eine scheinbare Verbesserung erzielen würden.

²⁵Ein Antipattern, in dem ein System keinerlei offensichtliche Architektur zu haben scheint

6. Auswahl der Entwicklungsmethodik und -Werkzeuge

6.1. Herausbildung einer Entwicklungsstrategie für die Bedürfnisse der pludoni GmbH

Viele der gängigen Entwicklungsstrategien, wie V-Modell oder Rational Unified Process, finden ihre Anwendung in großen Projektteams. Für mittelgroße Projektteams gibt es seit ca. 10 Jahren die agilen Prozesse. Sie haben einen eher pragmatischen Ansatz, mit dem Ziel gemeinsam mit dem Kunden eine funktionierende Software zu bauen. Zu eigen machen sie sich dabei kurze Releasezyklen, welche regelmäßig Feedback geben. Damit wird der klassische GAU am Ende des Projektes, wenn die Wünsche des Kunden mit den tatsächlichen Umsetzungen doch nicht einher gehen, vermieden. Aber viele dieser Methoden, wie z.B. SCRUM, benötigen eine Schulung für das gesamte Team, die nicht immer finanzierbar ist. Für die Arbeit von sehr kleinen Teams mit weniger als 4 Mitgliedern, wird nun eine Entwicklungsstrategie auf Basis der Testgetriebenen Entwicklung mit Ruby on Rails vorgestellt, die auf die Bedürfnisse der pludoni GmbH zugeschnitten ist.

Diese Bedürfnisse umfassen

- kurze Feedbackzyklen von 1 Woche
- Arbeit meist aus der Ferne ohne direkte Kommunikation mit den anderen Teammitgliedern. Daraus folgt ein äußerst selbstständiger Arbeitsstil
- möglichst fehlerfreie Software
- Kontinuierliche Integration
- pragmatisches Testen, 100% Testabdeckung ist nicht erforderlich. Wichtige Systemlogiken, wie Bezahlvorgang und Suche müssen dagegen getestet werden. Offensichtliche CRUD²⁶-Methoden müssen nicht getestet werden

6.1.1. Einteilung der Features in Kategorien

Grundsätzlich teilt die pludoni GmbH Features in zwei Kategorien ein:

- A. Features, welche in der Ansicht für Kunden und Besucher der Website sichtbar sind → Detailansichten, Listen, Bezahlvorgänge, ...

²⁶Create Read Update Delete - Die 4 Standardmethoden, die auf Ressourcen ausgeführt werden können

B. Features, welche nur dem Admin sichtbar sind, oder welche im Backend ausgeführt werden → Reporting, Statistiken, Indizierung der Datenbank, Cron-Scripte, Caching, ...

Features der **Kategorie A** sollen in Zukunft Akzeptanztestgetrieben entwickelt werden. Die Entwicklung verläuft nach dem Schema, dass in Abschnitt 3.6.1 vorgestellt wurde. Die Akzeptanztests sollen in Cucumber geschrieben werden. Die Websiteinteraktion soll mithilfe von Rack:Test im Falle einer Rails-Anwendung, und mit Capybara andernfalls simuliert werden. Ziel ist es, das bei Webanwendungen übliche wiederholte manuelle Ausprobieren mit dem Browser, auf ein Minimum zu reduzieren. Jeglicher Vorgang, den der Kunde am Browser testet, lässt sich auch als ein Akzeptanztest formulieren. Ein automatisierter Test hat zudem den Vorteil zu einem späteren Zeitpunkt leicht wiederholt zu werden.

Der Vorteil dieser Outside-In Entwicklung ist, dass er auf den Kunden ausgerichtet ist. Die Verwendung der domänspezifischen Sprache Cucumber fördert zudem die Implementierung von Business-relevanten Features gemeinsam mit dem Kunden. Das gesamte Vokabular orientiert sich an der Anforderungsanalyse und an Businessprozessen, die auch der möglicherweise nicht-technische Kunde verstehen kann.

Für die von außen nicht-sichtbaren Features der **Kategorie B** sollen aus Kostengründen normale Unittests entwickelt nach der klassischen Testgetriebenen Entwicklung genügen. Die zusätzliche Abstraktionsebene der Akzeptanztests ist nicht notwendig.

6.1.2. Weitere Bestandteile der Entwicklungsstrategie

Der oben genannte Teil bezieht sich in erster Linie auf den Kunden. Weitere Praktiken, die für den Programmieralltag wichtig sind, umfassen:

Kontinuierliche Integration Das Vorhandensein einer großen Test-Suite ermöglicht es, diese beim Einchecken in den Hauptzweig komplett auszuführen. Damit lässt sich sicherstellen, dass auf dem Hauptzweig eine immer lauffähige Version vorhanden ist. Die Verwendung einer Versionsverwaltung, z.B. git, ist obligatorisch.

In großen Projekten ist es üblich, komplexe Testpläne zu erstellen. Anscheinend sind aber automatisierte Tests, die bei jeden Einchecken durchgeführt werden, effektiver als rein formale Testpläne [Andrew Hunt, 1999][S. 238].

Code-Metriken Ein tägliches Messen des Code-Zustandes mittels Code-Metriken ermöglicht es den Programmierern, sich selbst und gegenseitig auf die Finger zu schauen. Sollte ein Programmierer nämlich schlechten Code im Hinblick auf Komplexität und Code-Smells abliefern, so macht die Code-Analyse dies sichtbar. Dies dient in erster Linie nicht, um den Programmierer zu maßregeln, sondern ihm dabei zu helfen, den TDD-Prozess zu lernen und seinen Programmierstil ständig zu verbessern. Die Erfahrungen zeigen, dass die Programmierer meist selbst unzufrieden mit schlechtem Code, den sie geschrieben haben, sind. Code-Metriken können dabei helfen, dem Programmierer schnell ein Feedback zu seinem Code zu geben, wie es ein Code-Audit durch Andere in der Geschwindigkeit und Effizienz nie könnte.

Regelmäßige Paar-Programmierung Die in Abschnitt 1.1.2 angesprochene Dezentralisierung der Zusammenarbeit erschwert eine regelmäßige Paarprogrammierung (Pair-Programming [PP]). Nichtsdestotrotz sollten in regelmäßigen Abständen Features zu zweit entwickelt werden. Insbesondere beim Lösen schwieriger Aufgaben und beim Anlernen neuer Teammitglieder ist Pair-Programming eine effektive Methode [Hulkko and Abrahamsson, 2005][S. 9]. Erfahrungsgemäß führt Pair-Programming zu besser dokumentierten Code, kann die Anzahl der Fehler verringern und zu einer höheren Arbeits-Effektivität führen [Hulkko and Abrahamsson, 2005].

6.2. Auswahl der Entwicklungswerkzeuge

Für die zukünftige Entwicklung vorrangig von Webanwendungen, werden folgende Werkzeuge berücksichtigt.

Werkzeuge für Tests Die formale Beschreibungssprache **Cucumber** dient als Schnittstelle für die vom Programmierer entwickelten Testschritte. Diese könnten in einem von vielen Testframeworks geschrieben werden. Die Entscheidung viel hierbei auf Minitest (Test::Unit in Ruby 1.9), da die Syntax und Prädikate denen von JUnit und NUnit sehr ähneln, und so den Übergang zu Ruby leichter machen. Da es auch das Standard-Testframework von Ruby on Rails ist, ist so eine gute Unterstützung durch gängige Werkzeuge garantiert.

Für die Simulation eines Browsers gibt es ebenfalls verschiedene Ansätze. Als Basis fungiert dabei **Capybara**, welches unterschiedliche Browsersimulationen abstrahiert. Damit lassen sich z.B. **Selenium** ansteuern, welches wiederum Mozilla Firefox, Inter-

net Explorer oder Google Chrome fernsteuern kann. Dies ist allerdings sehr langsam, da ein kompletter Browser gestartet und ferngesteuert wird. Daher ist die Nutzung von Selenium nur für das Testen von möglicherweise problematischen Interaktionen und das Testen von Javascript notwendig. Für alle anderen Fälle kann man auf Rack-Test zurückgreifen, welches extrem schnell eine Rack-Anwendung²⁷ simuliert. Falls in Zukunft mehr Geschwindigkeit in der Testausführung, insbesondere bei den Selenium-Tests, gewünscht wird, so kann man auf Parallelisierung auf mehreren Computern zurückgreifen.

Für ein unmittelbares Feedback sind auch automatische Test-Runner erwünscht. Hierbei gibt es z.B. **autotest** und **guard**. Diese Programme beobachten den Projektbaum, und führen bei Änderung der Dateien automatisch die relevanten Tests aus. Um die Geschwindigkeit, und damit den Feedbackzyklus zu verbessern, können diese Programme so gesteuert werden, dass sie nur den Testfall ausführen, an dem gerade gearbeitet wird.

Durch **spork** lässt sich eine RoR-Anwendung starten und im Hintergrund halten, so dass eine erneute Testausführung deutlich schneller von statten geht, als wenn die komplette Anwendung neu geladen werden müsste.

Werkzeuge für Code-Metriken Für die Generierung von Code-Metriken dient das Ruby-Gem „metric-fu“, welches seinerseits über verschiedene Code-Metriken Zusammenfassungen bildet und diese auch zeitlich darstellen kann. Darunter fallen z.B. die Zyklomatische Komplexität, den Grad der Duplikationen, verschiedene Code-Smells, Nutzung der Versionsverwaltung und Testabdeckung.

Die Testabdeckung wird durch simple-cov berechnet, welches eine C0 Code Coverage bestimmt.

Texteditoren Innerhalb der Entwickler der pludoni GmbH besteht ein Konsens für die Verwendung von vim, da hier bereits eine große Basis an Plugins gesammelt wurde, die die Entwicklung von Webanwendungen und insbesondere Ruby on Rails erleichtern. Ein weiterer Vorteil ist es, dass vim ohne ein graphisches Interface auskommt, und so direkt von der Shell auf dem Webserver ausgeführt werden kann. Nichtsdestotrotz sei es zukünftigen Entwicklern freigestellt, eine IDE, wie Eclipse mit dem Plugin RadRails oder Netbeans zu verwenden.

²⁷Rack ist ein minimales Interface zwischen Webserver und Webanwendung – Ruby on Rails ist eine solche Rack Anwendung

6.3. Diskussion der Maßnahmen

TODO Zusammenfassung evtl ins Fazit oder Auswertung

Viele dieser Maßnahmen dienen dazu, den Feedbackzyklus so kurz wie möglich zu halten. Für eine Testgetriebene Entwicklung ist es unerlässlich, dass die Testausführung schnell abläuft. Andernfalls, so die Erfahrung, führt dies zu einer verminderten Ausführungsrate, und ist damit hinderlich für die Entwicklung von sauberen Code.

Durch die gewählte Entwicklungsstrategie kann sogar die Anforderungsanalyse und Abnahme testgetrieben durchgeführt werden.

7. Anwendung der Testgetriebenen Entwicklung

In den nachfolgenden Abschnitten wird exemplarisch an dem Objekt „Job“, also der internen Repräsentation einer Stellenanzeige, die Testgetriebene Entwicklung mit praktischen Beispielen näher erläutert. Besonderes Augenmerk soll dabei auf den Entwicklungsfluss von \uparrow TDD gelegt werden. Zu dessen Verdeutlichung ist am Dokumentenrand die jeweilige Phase innerhalb des TDD-Zyklus zu finden (Red, Green, Refactor), dem die im Text gezeigten Codeabschnitte zuzuordnen sind.

Ziel dieses Kapitels wird es sein, einen Überblick über die Art und Weise zu erhalten, mit denen die verschiedenen Teilbereiche einer Webanwendung testgetrieben entwickelt werden können. Die ersten beiden Abschnitte richten sich an zwei der Grundbausteine einer \uparrow MVC (vgl. Abschnitt 4.4.1). Webanwendung, den Modell- und Controllertests. Im Dritten Abschnitt sehen wir, wie Test Doubles verwendet werden können, um Zugriffe auf externe Datenlieferanten zu simulieren. Danach betrachten wir die Anwendung aus Anwendersicht, und widmen wir uns der Implementierung von Akzeptanz- und Systemtests (vgl. Abschnitt 2.5). Zum Schluss gibt es einen Ausblick auf das Testen von Javascript-Ereignissen.

7.1. Implementierung von Unit-Tests (Modelltests)

Ein Rails-Modell, wie in 4.4.1 auf S. 35 beschrieben, repräsentiert die Daten der Anwendung, und die Regeln, wie diese zu verändern sind. Bei Rails werden sie hauptsächlich dazu verwendet, um mit der zugrundeliegenden Datenbanktabelle zu interagieren. Per Konvention von Rails findet hier die Hauptarbeit, also die Business-Logik, statt.

Fast jeder Unittest bei Rails beinhaltet das Testen auf Validierungskriterien seines korrespondierenden Modells, d.h. wann eine Instanz dieses Modells gültig ist und damit gespeichert werden darf (man denke z.B. an Pflichtfelder für ein Modell „Nutzer“, oder die Validierung des Formates seiner E-Mail-Adresse). Weiterhin sollten natürlich alle weiteren, selbstdefinierten, Methoden getestet werden.

Diese Validierungen werden durch das \uparrow ORM-Framework ActiveRecord, welches Rails standardmäßig nutzt, bereitgestellt. Bevor wir weiter auf die

1. Der Anfang Während der Analyse wurden die benötigten Attribute bestimmt. In Abbildung 7 sei z.B. ein Fragment des Grobdesigns, in dem die Basisattribute der

Job
+title: String +description: Text +link: String +is_visible: bool +user_id: ref +company_site_id: ref -start_date -end_date +...

Abbildung 7: Attribute des Modells „Job“

Tabelle dargestellt werden. Neben den einfachen Attributen, wie Title, Description und Link, existieren auch Referenzen auf andere Objekte (d.h. dies stellen Fremdschlüssel zu anderen Tabellen dar), wie z.B. Schlagwörter (Tags), ein Besitzer einer Stellenanzeige (User) und so weiter.

Einer der häufigsten Wege, ein Modell und dessen Datenbankschema zu generieren, ist die Nutzung des mitgelieferten Codegenerators. Mittels des Kommandos:

```
rails generate model MODELLNAME spalte1:datentyp1 spalte2:datentyp2 ...
```

generieren wird ein Modell mit dem angegebenen Modellnamen. Dazu geben wir paarweise die gewünschten Spaltennamen und deren Datentypen an (string, text, datetime, references, integer, boolean, decimal, ...).

```
~/it-jobs$ rails generate model job title:string link:string \
description:text user:references visible:boolean ...

invoke active_record
create db/migrate/20110828160636_create_jobs.rb
create app/models/job.rb
invoke test_unit
create test/unit/job_test.rb
create test/fixtures/jobs.yml
```

Mit der Anweisung uns ein Modell „job“, mit den nachfolgenden Attributen zu generieren, hat Rails uns nun schon ein Stück Arbeit abgenommen.

Es wurden erstellt:

- Eine Migration (`db/migrate/2011xxxxxx_create_jobs.rb`). Dies stellt eine datenbankunabhängige Repräsentation einer Änderung an der Struktur unserer Datenbank dar. In diesem ist es die Erstellung einer Tabelle „jobs,, (beachte: Plural!), mit den Spalten Titel, Link als String, Description als Textfeld, eine User_ID als Referenz auf ein anderes Modell usw.
- Die Modelklasse (`app/models/job.rb`). Trotz unserer Definition der Spalten und deren Typen über die Kommandozeile, ist diese Klasse leer. Da wir ActiveRecord verwenden, definieren wir die Attribute, die unser Modell hat, nicht in der Modellklasse, sondern ausschließ in der Datenbank. Die Migration erspart uns die manuelle Arbeit, selbst in unserer Datenbank Spalten anzulegen. Bei Initialisierung eines Modells lädt ActiveRecord die Spalteninformationen aus der Datenbank, und generiert dafür Getter und Setter Methoden.
- Die dazugehörige Testklasse (`app/unit/job_test.rb`)
- und Fixtures-Datei (`test/fixtures/jobs.yml`), zur Definition von Testdaten.

Die Migration liegt nun zwar vor, aber es existiert noch keine Datenbank und demnach auch noch keine Tabelle mit dem Namen „jobs“. Dazu weisen wir nun Rails an, alle offenen Migrationen auszuführen. Standardmäßig erstellt Rails dann selbstständig eine SQLite Datenbank unter „db/development.sqlite3“. Danach können wir die Rails-Test-Suite auch schon ausführen:

```
$ rake db:migrate && rake test

== CreateJobs: migrating =====
-- create_table(:jobs)
--> 0.0020s
== CreateJobs: migrated (0.0021s) =====

(in /home/zealot64/TEST)
Loaded suite /usr/lib/ruby/gems/1.8/gems/rake-0.8.7/lib/rake/
  rake_test_loader
Started
.
Finished in 0.043818 seconds.

1 tests, 1 assertions, 0 failures, 0 errors
```

Es wurde also schon ein Testfall erfolgreich ausgeführt, nämlich ein Dummytestfall von Rails:

```
test/units/job_test.rb
1 #test/unit/job_test.rb
2 require 'test_helper'
3
4 class JobTest < ActiveSupport::TestCase
5   # Replace this with your real tests.
6   test "the truth" do
7     assert true
8   end
9 end
```

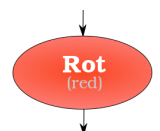
Listing: Listing Test

2. Testen auf Validierung Ein Feature von Rails umfassen die sogenannten Validierungen. Diese stellen sicher, dass eine Instanz eines Modells nur dann gespeichert ist, wenn es gewissen Kriterien entspricht. Viele der Validations sind vergleichbar mit den Datenbank-Constraints einiger Datenbanken. Rails nutzt diese standardmäßig nicht, da es auch andere Persistenzsysteme unterstützt, wie z.B. Key-Value-Store oder sogenannte NoSQL Datenbanken. So stellt Rails die Konsistenz und referenzielle Integrität innerhalb der Applikationsschicht sicher.

Nun möchten wir sicherstellen, dass eine Stellenanzeige nur dann gespeichert wird, wenn sie einen Titel beinhaltet. Der Test dazu würde wie folgt lauten:

```
test/units/job_test.rb
1 require 'test_helper'
2
3 class JobTest < ActiveSupport::TestCase
4   test "ein Job muss einen Titel haben" do
5     job = Job.new
6     job.title = nil
7     assert !job.save
8   end
9 end
```

Listing: Test auf Vorhandensein eines Titels



Zuerst instanziiieren wir einen Job, und geben ihm explizit einen leeren Titel, um das Testziel nochmal herauszustellen. Danach rufen wir die „save“-Methode auf, die prüft, ob alle Validierungskriterien erfolgt sind, und speichert das Objekt persistent in der Datenbank im Erfolgsfall. Dann gibt „save“ ein „true“ zurück, andernfalls, d.h. wenn die Validierung fehlschlug, „false“. Der Ablauf ist in der Abbildung 8 noch einmal erläutert.

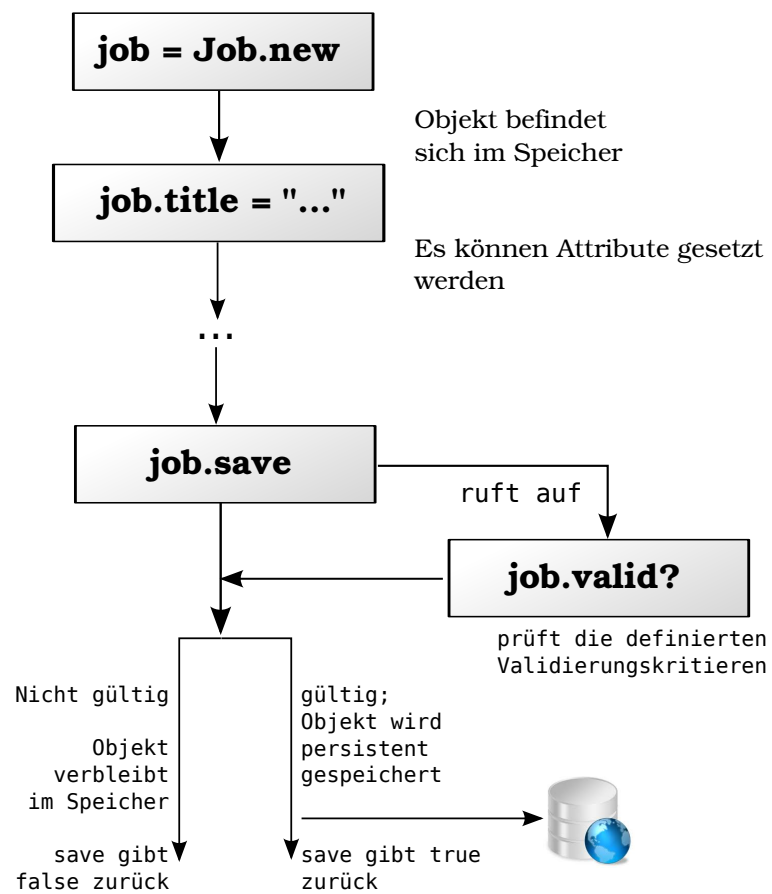


Abbildung 8: Funktionsweise von save bei ActiveRecord Objekten

Da wir noch keine Validierungskriterien implementiert haben, schlägt dieser Test fehl, da das Objekt gespeichert wurde.

Unser nächstes Ziel ist es nun, mit so wenig Code wie möglich den Test bestehen zu lassen. Das können wir mittels der eingebauten wie schon erwähnten Validierungen:

```
app/models/job.rb  
1 class Job < ActiveRecord::Base  
2   validates :title, :presence => true  
3 end
```

Listing: Implementierung der Validierung in die Klasse Job

„validates“ ist eine Funktion aus der ActiveRecord Bibliothek, die zwei Parameter entgegennimmt: Der erste ist die Spalte, auf der sich die Validierung bezieht, als Zweites folgt eine Liste an Validierungskriterien. Hier ist das Kriterium „presence“, also das Vorhandensein eines nicht-leeren Attributs. Weitere Kriterien sind z.B. Format, Länge, Minimum, Maximum, oder selbst definierte Kriterien.

Nach erneuter Ausführung der Testsuite, besteht der Test nun. Jetzt folgt die Refaktorisierungsphase. Der Programmcode lässt sich nicht weiter vereinfachen. Aber der Testcode ist ausdrücklich nicht von Refaktorisierungen befreit, und eine Refaktorisierung wäre z.B.:

```
test/unit/job_test.rb  
1 test "ein Job muss einen Titel haben" do  
2   job = Job.new :title => nil  
3   assert !job.save  
4 end
```

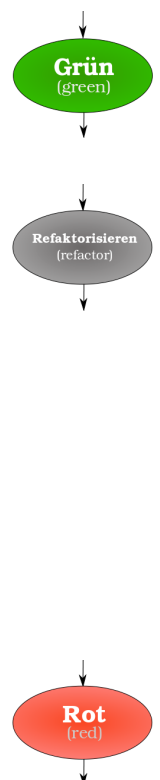
Listing: refaktorisierter Test

Nun wollen wir dasselbe für das Feld E-Mail tun, hierbei aber nicht nur das Vorhandensein prüfen, sondern auch das Format.

```
test/unit/job_test.rb  
1 test "ein Job muss eine gültige E-Mail haben" do  
2   job = Job.new :email => "invalid_email"  
3   assert !job.save  
4 end
```

Die Implementierung wäre dann:

```
app/models/job.rb  
1 class Job < ActiveRecord::Base  
2   validates :email, :format => /^[w\d_~]+@[w\d_~]\. [w\d]{2,3}$/  
3   ...  
4 end
```



Eine Refaktorisierung ist aufgrund der Einfachheit der Beispiel hier nur gering möglich. Man könnte z.B. den regulären Ausdruck, der das Format der E-Mail Adresse beschreibt in eine neue Klasse oder zumindest eine Konstante auslagern. Wir wählen eine Konstante, die beim Laden von Rails bereitgestellt wird.

```
_____ config/initializers/job.rb und app/models/job.rb _____  
1 # config/initializers/regex.rb  
2 REGEX_EMAIL_FORMAT = /^[w\d_~]+@[w\d_~]+\.[w\d]{2,3}$/  
3  
4 # app/models/job.rb  
5 class Job < ActiveRecord::Base  
6   validates :email, :format => REGEX_EMAIL_FORMAT  
7   ...  
8 end
```

Listing: Auslagerung des Regulären Ausdrucks in einen Initialisierer

Ein erneutes Ausführen der Tests betätigt den Erfolg der Refaktorisierung.

3. Refaktorisierungen der Testklasse Nun fehlt aber noch die Definition eines Positiv-Beispiel für einen gültigen Job.

```
_____ test/unit/job_test.rb _____  
1 ...  
2 test "ein vollstaendiger Job muss gueltig seinn" do  
3   job = Job.new :title => "Rails Entwickler", :email => "info@stefanwienert.net"  
4   assert_valid job  
5 end
```

Dieser Test besteht sofort, macht also genau genommen keine weitere Aussage über unser System. Nach der „reinen“ Testgetriebenen Leere sollte dieser entfernt werden. Es ist allerdings eine gute Strategie, bei Validierungen mindestens ein Beispiel zu präsentieren, dass angenommen wird. Nichtsdestotrotz können wir nun Refaktorisieren. Insbesondere unsere Testfunktionen enthalten unnötige Redundanzen:

```
_____ test/unit/job_test.rb _____  
1 test "ein Job muss einen Titel haben" do  
2   job = Job.new :title => nil  
3   assert !job.save  
4 end  
5 test "ein Job muss eine gültige E-Mail haben" do  
6   job = Job.new :email => "invalid_email"  
7   assert !job.save  
8 end
```



```
9 test "ein vollstaendiger Job muss gueltig seinn" do
10   job = Job.new :title => "Rails Entwickler", :email => "info@stefanwienert.net"
11   assert_valid job
12 end
```

Listing: Alle bisherigen Testmethoden in der Klasse JobTest

In allen drei Methoden wird ein Job instanziiert, und lediglich verschiedene Attribute überprüft. Auch haben unsere ersten beiden Tests keine gültige Aussage mehr, da der jeweilige Job sowieso nicht gültig ist, da jeweils das andere Attribut fehlt²⁸. Es ist also höchste Zeit, die Tests zu refaktorisieren. Dies geschieht am Besten durch die Verwendung einer Testdaten-Generation, z.B. den eingebauten Fixtures, die Rails uns bei der Codegeneration schon mit generiert hatte. Dabei definieren wir zentralisiert unsere (gültigen) Testdaten, die von Rails vor jedem einzelnen Test in der Datenbank bereitgestellt werden:

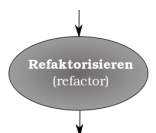
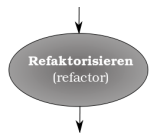
```
test/fixtures/jobs.yml
1 valid_job:
2   title: Rails Entwickler
3   email: info@stefanwienert.net
4   link: "http://www.example.com/jobs"
5   visible: true
6   ...
7 invisible_job:
8   title: Rails Entwickler
9   visible: false
10  ...
```

Listing: Fixtures Testdaten für zwei Jobs

Nun können wir diese Fixtures in unseren Tests verwenden, und das ganze in einer setup-Methode, die vor jedem Testfall aufgerufen wird, laden:

```
test/unit/job_test.rb
1 class JobTest < ActiveSupport::TestCase
2   setup do
3     @job = jobs :valid_job
4     # Dies lädt den Job mit dem Schlüssel "valid_job" und schreibt ihn
5     # in die Instanzvariable @job der Testklasse
6   end
7   test "stelle sicher, dass die Fixtures valide sind" do
8     assert_valid @job
9   end
10  test "ein Job muss einen Titel haben" do
```

²⁸Im ersten Test ist nicht nur der Titel nicht gesetzt, sondern auch die E-Mail entspricht nicht dem Format



```
11  @job.title = nil
12  assert !@job.save
13  end
14  test "ein Job muss eine gültige E-Mail haben" do
15    @job.email = "invalid_email"
16    assert !@job.save
17  end
18  end
```

Listing: Finale Job-Test Klasse nach Refaktorisierung

Am Ende dieser Refaktorisierungen ist es notwendig, die Tests noch einmal auszuführen. Danach würde die Implementierung einer nächsten Teilanforderung sein.

In diesem Abschnitt war zu sehen, dass die Testgetriebene Entwicklung das Arbeiten und Testen in kleinen Schritten favorisiert.



7.2. Implementierung von Controller-Tests (functional tests)

Neben den Unittests stellt Ruby on Rails eine weitere Testart nativ bereit. Technisch gesehen handelt es sich bei diesen Functional Tests aber auch um Unittests, da deren Testobjekt ein Controller ist. Ein Controller hat bei Ruby on Rails die Aufgabe, Anfragen für bestimmte Routen, also Web-Adressen, anzunehmen, die Arbeit an eine Modelklasse auszulagern, und eine View aufzurufen, die letztendlich HTML-Code generiert.

Im ersten Beispiel wollen wir testen, dass ein Gast-Nutzer, also z.B. ein Bewerber, eine sichtbare Stellenanzeige aufrufen darf (`visible = true`). Hierbei verwenden wir wieder unser oben definiertes Fixture für einen gültigen Job.

```
/test/functional/jobs_controller_test.rb
require 'test_helper'

class JobsControllerTest < ActionController::TestCase
  test "Gast Nutzer kann Stellen betrachten" do
    session[:user_id] = nil
    job = jobs(:valid_job)

    get :show, :id => job.id

    assert_response :success
    assert_equal job, assigns(:job)
  end
end
```

```
end
```

Zuerst loggen wir jeglichen Nutzer aus, der eventuell eingeloggt war, dann laden wir das Fixture und führen einen simulierten HTTP Request auf die Detailansicht der Stellenanzeige aus (Die Aktion „show“ mit der ID des Jobs). Nun erwarten wir, dass wir einen HTTP-Status Code 200 (success) erhalten, und dass der Controller eine Variable „@jobs“ bereitstellt, die mit unserem Fixture identisch ist.

Die Implementation dieser Anforderung könnte wie folgt umgesetzt werden:

```
# app/controllers/jobs_controller.rb
class JobsController < ApplicationController
  ...
  def show
    @job = Job.first
  end
  ...
end
```

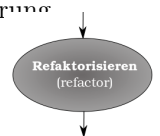
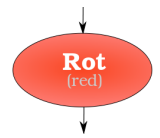
Das Laden des ersten Jobs aus unserer Datenbank genügt zum Erfüllen der Anforderungen, und ist ein schneller Weg, den Test bestehen zu lassen. Allerdings handelt es sich hierbei um eine Fake-Implementierung, da zwar unser Test erfüllt wird, aber die Anwendung nicht das macht, was man sich erhofft hat. Solche Zwischenschritte sind aber ausdrücklich vorgesehen, da das Ziel ist, so schnell wir möglich einen funktionierenden Test zu erhalten mit dem man arbeiten kann.

Wenn wir nun weitere Tests schreiben, so wird es immer schwieriger, die Fake-Implementierung beizubehalten, und früher oder später wird eine korrekte Implementierung folgen. Aber wir können auch die nun folgende Refaktorisierungsphase nutzen, um diesen Makel zu beseitigen:

```
# app/controllers/jobs_controller.rb
def show
  @job = Job.find(params[:id])
end
```

Nun wollen wir testen, ob ein Gast von einer nicht-sichtbaren Stellenanzeige weitergeleitet wird und einen Hinweis erhält.

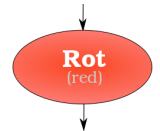
```
test "Gast Nutzer kann nicht-sichtbare Stellen nicht betrachten" do
  session[:user_id] = nil
```



```
job = jobs(:invisible_job)

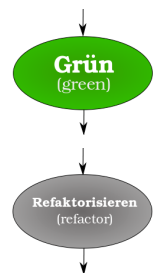
get :show, :id => job.id

assert_response :redirect
assert flash[:notice].present?
end
```



Wir laden unser zweites definiertes Fixture, dass eine unsichtbaren Stellenanzeige. Dieses mal erwarten wir einen HTTP Statuscode 301 (Redirect), und dass unser Controller eine Hinweismnachricht generiert.

```
def show
  @job = Job.find(params[:id])
  if not @job.visible?
    redirect_to root_path, :notice => "Diese Stelle ist zur Zeit nicht
      sichtbar"
  end
end
```



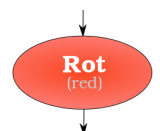
Falls der aktuelle Job nicht sichtbar ist, dann erfolgt eine Weiterleitung auf die Startseite und die Bereitstellung des Hinweistextes. Da auch hier der Quelltext wieder sehr kurz ist, ist ein Refaktorisieren nicht notwendig.

Nun möchten wir, dass ein Kunde dieser Anwendung, also ein Unternehmen seine Stellenanzeige betrachten kann, auch wenn diese unsichtbar ist, sei es aus Gründen der Archivierung als auch der Vorbereitung für eine Veröffentlichung.

```
test "Ein Kunde darf aber seine unsichtbaren Jobs betrachten" do
  job = jobs(:invisible_job)
  session[:user_id] = job.user_id

  get :show, :id => job.id

  assert_response :success
end
```



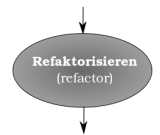
Über die globale Session Variable simulieren wir das Einloggen durch setzen der UserID in dieses Array. Die genaue Implementation hängt natürlich davon ab, wie man die Authentifizierung implementiert hat, oder welche Bibliothek man verwendet. In diesem Beispiel sei darauf hingewiesen, dass die Definition, ob ein Nutzer eingeloggt ist oder nicht, davon abhängt, ob in seiner Session-Variable eine UserID enthalten ist.

```
def show
  @job = Job.find(params[:id])
  if !@job.visible? and @job.user != User.find(session[:user_id])
    redirect_to root_path, :notice => "Diese Stelle ist zur Zeit nicht
      sichtbar"
  end
end
```



Wir lösen diesen Test damit, dass wir in der Weiterleitungsbedingung prüfen, ob der betrachtende Nutzer und der Eigentümer des Jobs gleich sind.

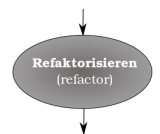
Nun können wir refaktorisieren. Was auffällt, ist z.B. dass unser Controller und die Klasse Job nicht lose gekoppelt sind, da die Bedingung zweimal auf Attribute des Jobs zurückgreift. Eine Lösung wäre die Auslagerung in die Modelklasse von Job:



```
# app/models/job.rb
class Job < ActiveRecord::Base
  ...
  def visible_for_user?(user)
    self.visible and self.user != user
  end
end

# app/controllers/jobs_controller.rb
def show
  @job = Job.find(params[:id])
  unless @job.visible_for_user?(User.find(session[:user_id]))
    redirect_to root_path, :notice => "Diese Stelle ist zur Zeit nicht
      sichtbar"
  end
end
```

Ebenfalls wurde das syntaktische Element „unless“ verwendet, welches ein Alias für „if not“ ist. Weiterhin könnte die Suche nach dem aktuell eingeloggteten Nutzer in eine für alle Controller sichtbare Funktion ausgegliedert werden



```
# app/controllers/application_controller.rb
...
def current_user
  User.find(session[:user_id])
end

# app/controllers/jobs_controller.rb
```



```
def show
  @job = Job.find(params[:id])
  unless @job.visible_for_user? current_user
    redirect_to root_path, :notice => "Diese Stelle ist zur Zeit nicht
      sichtbar"
  end
end
```

Funktionale Tests und deren Controllerimplementierungen sind häufig nicht länger als ein paar Zeilen. Qua Konvention des [†]MVC-Patterns und Rails sollen komplexe Abläufe in den Modellen oder auch in Bibliotheken stattfinden. Die Aspekte, die üblicherweise bei Controllern getestet werden, sind:

- HTTP Statuscodes und Weiterleitungen,
- das Vorhandensein von Statusmeldungen, genannt „Flash“ Messages
- dass ein bestimmtes Template geladen wird
- dass Instanzvariablen gesetzt werden, die die View später darstellen sollen
- falls man Viewtests mit einschließt, dann wird u.U. auch auf das Vorhandensein von bestimmten HTML-Elementen in der am Ende generierten View getestet. Z.B. möchte man wissen, ob das Überschriftenelement „h1“ dem Job-Titel entspricht, wenn die Detailansicht eines Jobs aufgerufen wird.

Skinny Controller,
Fat Model [...] Try
to keep your
controller actions
and views as slim
as possible.

Jamis Buck,
Programmierer
bei 37signals

7.3. Testen von externen Abhängigkeiten

Fast alle Webapplikationen sind auf Kommunikation mit anderen Servern angewiesen. Als Beispiel seien die diversen APIs der sozialen Netzwerke genannt oder Webservices. Für die vorliegende Jobanwendung war gewünscht, ein Feedimport-Feature zu implementieren, sodass bestimmte Kunden ihre Stellenanzeigen automatisiert einlesen lassen könnten.

Die genannten Partner stellen einen XML-Feed nach dem RSS 2.0 Format²⁹ bereit, der ein häufig verwendetes Format zum Austausch von Informationen ist, und durch eine Vielzahl von Werkzeugen und Content-Management-Systemen unterstützt wird. Dabei wird der Inhalt des Haupttextfeldes „description“ um weitere Informationen in einem Subdialekt angereichert.

Im Nachfolgenden sei z.B. eine Stellenanzeige in dem Format beschrieben:

²⁹Spezifikation des RSS 2.0 Formats: <http://cyber.law.harvard.edu/rss/rss.html>

```
beispiel-job.xml
1 <?xml version="1.0" encoding="UTF-8"?>
2 <rss version="2.0">
3   <channel>
4     <title>RSS Feed für Jobangebote </title>
5     <language>de</language>
6     <item>
7       <title>Softwareentwickler Java/JEE (m/w)</title>
8       <description>
9         <![CDATA[
10           <!--
11             <nummer>example_job_01</nummer>
12             <tags>Java,Webentwickler,Softwareentwickler</tags>
13             <ort>Dresden</ort>
14             <kontakt>Max Mustermann bewerbung@example.com</kontakt>
15             <link>http://www.example.com/jobs/512.html</link>
16             -->
17             Zur Verstärkung unseres Teams suchen wir zum nächstmöglichen
18             Zeitpunkt einen Softwareentwickler Java/JEE (m/w) zur Festanstellung.<br />
19             Ihre Aufgaben: ...
20           ]]>
21         </description>
22         <link>http://www.example.com/jobs/512.html</link>
23         <pubDate>Wed, 25 Mar 2011 13:30:00 +0100</pubDate>
24         <guid>example_job_01</guid>
25       </item>
26     </channel>
27 </rss>
```

Listing: Feedimport Beispiel-XML Datei mit einem Job

Der RSS-Feed in dem oben genannten Beispiel enthält eine Stellenanzeige (item). Die description beinhaltet einen HTML-Kommentar, in dem nummer, tags, ort, kontakt und link für die Stellenanzeige definiert werden. Das ganze wurde mit einem Kommentar, und nicht mit einer Erweiterung der Syntax durch eine DDT oder XSD, realisiert, da sich eine Eingliederung der Syntaxelemente mittels DDTs und XSDs in einige der Systeme der Kunden als problematisch herausgestellt hat.

Diese Art des Feedimports ist bereits in den Community-Job-Portalen in Funktion. Allerdings besitzt dieser, in PHP geschriebene Code, keinerlei automatisierte Tests, und war in der Vergangenheit schon oft die Ursache von Fehlern. So ist es notwendig, den Feedimport nun in Ruby als Bibliothek im Rahmen von IT-Jobs neu zu schreiben, und für die bereits laufenden Portale schnellstmöglich einzubauen. Diese Bibliothek soll also unabhängig von Rails funktionieren.

Ziel dieses Abschnittes ist es, zu zeigen, wie das Einlesen eines externen XML-Feeds getestet werden kann.

1. Initialier Test Bevor man anfängt zu implementieren, ist es sinnvoll sich Gedanken darum zu machen, was von den zu implementieren Objekten erwartet wird. Da wir letztendlich eine gewisse Menge von RSS-Feeds einlesen wollen, ist es angebracht, ein entsprechendes Objekt, z.B. „ImportedFeed“ einzuführen.

Auch wenn wir noch nicht genau wissen, wie ein Feed funktioniert, so können wir doch zumindest annehmen, dass ein HTTP-Zugriff auf eine URL erfolgt, um den Feed vom Kunden abzuholen.

Da wir unsere Tests nicht davon abhängig machen wollen, ob ein solcher Feed bereitsteht und sich stets im selben Zustand befindet, müssen wir diesen HTTP-Zugriff simulieren.

```
test/test_imported_feed.rb
1 require "test_helper" # Stuff that we need for convenient tests
2 require "imported_feed" # Object under Test
3
4 class TestImportedFeed < ActiveSupport::TestCase
5
6   test "get an feed through httparty" do
7     HTTParty.expects(:get).with("http://www.example.com/feed.xml")
8     ImportedFeed.new("http://www.example.com/feed.xml")
9   end
end
```

Listing: Feed Test I.

Hier definieren wir einen ersten Test für den ImportedFeed. Für die HTTP-Zugriffe wollen wir die Bibliothek HTTParty³⁰ benutzen. In der ersten Zeile des Tests nutzen wir das Mock-Framework wie folgt: Wir legen eine Erwartung fest, dass innerhalb dieses Tests die Klassenmethode „get“ der Klasse HTTParty aufgerufen wird, mit einem Parameter der die URL angibt.

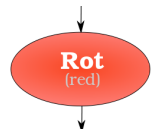
Dann rufen wir unsere (noch nicht existente) ImportedFeed Klasse mit dem einzulesenden Feed.

Nach der Ausführung des Tests erhalten wir einen Fehler.

```
NameError: uninitialized constant TestImportedFeed:ImportedFeed
```

In der reinen TDD-Lehre würde nun als erstes eine Behebung aller Fehler stattfinden, d.h. eine Implementierung der leeren Klasse ImportedFeed. Danach würden wir noch einen Fehler erhalten, da unser Konstruktor noch keinen Parameter entgegennimmt.

³⁰<https://github.com/jnunemaker/httparty>



Erst dann würde man sich den Testfehlschlägen widmen. Aus Platzgründen sind diese Schritte hier nicht explizit ausgeführt. Der Fehlschlag lautet dann:

```
Failure
not all expectations were satisfied
unsatisfied expectations:
- expected exactly once, not yet invoked: HTTParty.get('http://www.
  example.com/feed.xml')
```

Das Mockobjekt hat unseren Test fehlschlagen lassen, ohne dass wir selbst eine Assertion festgelegt hätten. Da wir bisher noch keine Implementation eines Netzwerkzugriffes durch HTTParty implementiert haben, schlägt der Test fehl.

Die Implementierung wäre:

```
lib/imported_feed.rb
1 require "httparty"
2
3 class ImportedFeed
4   def initialize(url)
5     HTTParty.get(url)
6   end
7 end
```

Listing: Feed Implementation I.



Die Funktion initialize stellt innerhalb von Ruby den Konstruktor dar. Dort rufen wir unseren Netzwerkzugriff auf, der allerdings durch unser definiertes Mock-Objekt abgefangen wird. Dies stellt die definierte Erwartung zufrieden, und der Test besteht.

2. Komplexe Objekte durch Mocks zurückgeben Wir haben zwar den Netzwerkzugriff abgefangen, geben aber nun keinerlei Antwort, d.h. ein XML-Dokument zurück. Für den nächsten Test müssen wir unsere Mockanweisung also modifizieren.

```
test/test_imported_feed.rb
1 test "really get content from an feed" do
2   fake_response = OpenStruct.new
3   fake_response.code = 200 # HTTP OK!
4   fake_response.body = "<?xml version='1.0'?><Hallo/>"
5
6   HTTParty.expects(:get).with("http://www.example.com/feed.xml").returns(fake_response)
7
8   import = ImportedFeed.new(@url)
9   assert_match "Hallo", import.body
10 end
```

Listing: Feed Test II

Wir bilden das Antwortobjekt, das HTTParty normalerweise generieren würde, beschränken uns hierbei aber nur auf die für uns notwendigen Methoden von „body“ und „code“ (Dem HTTP-Status Code). Wir nutzen dazu die Klasse OpenStruct, die Getter und Setter für das Objekt beim Benutzen erstellt. Unserem Mock können wir dann anweisen, diese Antwort zurückzugeben.

Bei Ausführung des Tests stellen wir fest, dass zwar die Erwartung erfüllt wurde, aber unser ImportedFeed noch kein Attribut „body“ besitzt (Fehler), und dass dieser keine String beinhaltet.

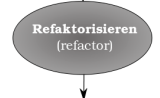
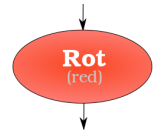
```
app/models/job.rb
1 class ImportedFeed
2   attr_reader :body
3   def initialize(url)
4     response = HTTParty.get(url)
5     @body = response.body
6   end
7 end
```

Listing: Feed Implementation II

Mithilfe des Makros „attr_reader“ generieren wir ein Attribut body und gleichzeitig einen Getter für den Zugriff von außen. Innerhalb unseres Konstruktors speichern wir den Body der HTTP-Antwort in diesem Attribut.

Da unser Testfall nun ziemlich lang geworden ist, und beide Testfälle ein Mock initialisieren, ist dies eine gute Gelegenheit, den Mock zentral zu definieren. Dazu nutzen wir z.B. eine Datei „test_helper.rb“, in der wir Anweisungen schreiben, die alle Testfälle nutzen können:

```
test/test_helper.rb
1 class ActiveSupport::TestCase
2   def mock_feed(opts={})
3     options = {
4       :url => "http://example.com/feed.xml",
5       :code => 200,
6       :body => '<?xml version="1.0" encoding="UTF-8"?><Hallo>Hallo</Hallo>'
7     }.merge(opts)
8     response = OpenStruct.new
9     response.code = options[:code]
10    response.body = options[:body]
11    HTTParty.expects(:get).with(options[:url]).returns(response)
12  end
end
```



```
13 end
```

Listing: Zentrale Implementierung des Mocks in der Test Helper

Da Ruby offene Klassen unterstützt öffnen wir die Basisklasse der Testfälle und definieren eine neue Methode. Diese erhält einen Hash als Parameter, den wir mit unseren Standardwerten zusammenmergen. Diese Art der Parameterübergabe ist ein sehr gebräuchliches Idiom innerhalb der Ruby-Community.

Der Aufruf unsere neuen Hilfsfunktion erfolgt dann mittels:

```
test/test_imported_feed.rb  
1 def setup  
2   @url = "http://example.com/feed.xml"  
3 end  
4  
5 test "get an feed through httparty" do  
6   should "perform a get request when initializing" do  
7     mock_feed :url => @url  
8     ImportedFeed.new(@url)  
9   end  
10  
11 test "really get content from an feed" do  
12   mock_feed :url => @url, @body => "Hallo"  
13   import = ImportedFeed.new(@url)  
14   assert_match "Hallo", import.body  
15 end
```

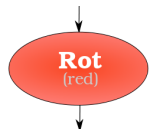
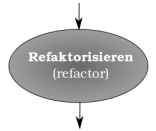
Listing: Feed Test IIb nach Refaktorisierung

Als zusätzliche Maßnahmen haben wir die Definition der URL in eine gemeinsame Initialisierungsmethode gesetzt.

3. Validität Unser Feed soll später feststellen können, ob er Fehler beinhaltet oder nicht, um dann ggf. eine E-Mail an den Verantwortlichen zu schreiben.

```
test/test_imported_job.rb  
1 test "have a valid method" do  
2   mock_feed  
3   import = ImportedFeed.new(@url)  
4   assert import.respond_to?(:valid?)  
5 end
```

Listing: Feed Test III



Nun testen wir lediglich darauf, ob das ImportedFeed Objekt eine Methode oder ein Attribut mit dem Namen „valid?“ besitzt.

Um den Test zu bestehen, reicht es, eine leere Methode zu definieren:

```
lib/imported_job.rb
1 ...
2 def valid?
3 end
```

Listing: Feed Implementation III - Fake Implementierung

Nun werden wir etwas konkreter, und erwarten, dass falls der kontaktierte Server einen Status-Code 404 (Dokument nicht gefunden – Ein wahrscheinlicher Fehlerfall, falls das XML-Dokument verschoben wurde) erhalten.

```
test/test_imported_job.rb
1 test "not validate if the user server reports a problem" do
2   mock_feed :code => 404
3   feed = ImportedFeed.new(@url)
4   assert !feed.valid?
5 end
```

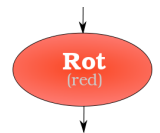
Listing: Feed Test IV – Test auf nicht-erfolgreichen HTTP-Status Code

Implementieren können wir das so:

```
ib/imported_job.rb
1 class ImportedFeed
2   attr_reader :body, :status_code
3   def initialize(url)
4     response = HTTParty.get(url)
5     @status_code = response.code
6     @body = response.body
7   end
8
9   def valid?
10    @status_code == 200,
11  end
12 end
```

Listing: Feed Implementation IV - Implementation der Validierung

Zu bemerken ist, dass wir den Funktionsrückgabewert nicht explizit mit „return“ kennzeichnen müssen. Bei Ruby hat jeder Ausdruck einen Rückgabewert. Innerhalb einer Funktion ist dies das letzte Statement, falls nicht mit return spezifiziert.



4. Testen auf Exceptions Zum Abschluss dieses Kapitels möchten wir noch sicher gehen, dass unser ImportedFeed robust gegenüber Exceptions von Fremdbibliotheken ist. Auch dies können wir in einer Erwartung durch unser Mock-Objekt spezifizieren.

```
test/test_imported_job.rb
1 test "should be resistant to any thrown errors from library" do
2   HTTParty.expects(:get).raises( ArgumentError)
3
4   feed = nil
5   assert_nothing_raised(ArgumentError) do
6     feed = ImportedFeed.new(@url)
7   end
8   assert !feed.valid?
9
10 end
```

Listing: Feed Test V - Testen der Robustheit gegen Exceptions

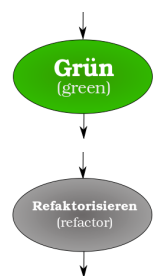
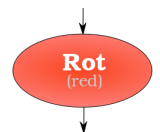
Durch die Zusicherung „assert_nothing_raised(exception, message, block)“ testen wir, dass innerhalb des übergebenen Blocks keine Exception vom Typ exception (Hier: „ArgumentError“) geworfen wird. Wir möchten in diesem Fall auch sichergehen, dass unser Feed als nicht-valide markiert wird.

```
lib/imported_job.rb
1 class ImportedFeed
2   attr_reader :body, :status_code
3   def initialize(url)
4     response = HTTParty.get(url)
5     @status_code = response.code
6     @body = response.body
7     @error_thrown = false
8     rescue Exception => e
9       @error_thrown = true
10    end
11    def valid?
12      !@error_thrown and @status_code == 200
13    end
14  end
```

Listing: Feed Implementation V

Alle Exceptions die innerhalb des Konstruktors geworfen werden, werden abgefangen, und das neue Attribut @error_thrown auf true gesetzt. Dies kann dann unsere valid Funktion verwenden.

Unser Konstruktor ist nun schon relativ lang geworden, und hat inzwischen schon mehrere Aufgaben: Abruf eines Feeds und setzten der HTTP-Antwort. Jede unserer



Methoden sollte eine klar umrissene Aufgabe haben. Dies erreichen wir nun durch das Auslagern in eine neue (private) Funktion.

```
lib/imported_job.rb
1 class ImportedFeed
2   attr_reader :body, :status_code
3   def initialize(url)
4     get_feed(url)
5     @error_thrown = false
6   rescue Exception => e
7     @error_thrown = true
8   end
9   def valid?
10    !@error_thrown and @status_code == 200
11  end
12
13  private
14  def get_feed(url)
15    response = HTTParty.get(url)
16    @status_code = response.code
17    @body = response.body
18  end
19 end
```

Listing: Feed Implementation Vb - nach Refaktorisierung

Diese private Methode hat nun keine eigenen Tests, ist aber aus einer Refaktorisierung hervorgegangen, und damit innerhalb von TDD ein erlaubter Schritt (Das scheinbare Dilemma des Testens privater Methoden wurde bereit auf S. 19 erläutert).

In diesem Abschnitt wurden einige Spezialfälle beim Umsetzen von TDD praktisch erläutert. Dies waren:

- Einsatz von Mocks zur Entkoppelung von externen Datenquellen und Spezifizierung von Erwartungen (Expectations)
- Nutzung von privaten Methoden bei der Refaktorisierung
- Tests mit Exceptions

7.4. System/Akzeptanztests

Nun wollen wir am Beispiel einer Suche nach Stellenanzeigen die Entwicklung von Akzeptanztests und damit die Akzeptanztestgetriebene Softwareentwicklung (vgl. Abschnitt 3.6.1 betrachten.

Zum Einsatz kommt dabei die für Systemtests entwickelte domainspezifische Sprache Cucumber, die Bereits in Abschnitt 4.3.2 vorgestellt wurde.

Ein Cucumber-Verzeichnis ist qua Konvention immer gleich aufgebaut. In Abbildung 9 ist der Verzeichnisbaum, wie er in den Beispielen dieses Kapitels benutzt wird, zu sehen. Innerhalb unseres Projektverzeichnisses existiert ein Unterverzeichnis „features“ in dem alle Feature-Dateien mit der Endung „.feature“ lagern. Ebenfalls existiert ein Unterverzeichnis „step_definitions“, das die Implementation der Testschritte beinhaltet, sowie ein Verzeichnis „support“, für z.B. Standardmethoden, Aktionen die vor jedem Feature-Durchlauf ausgeführt werden, oder die Definition der Browser-Engine. Cucumber lädt automatisch alle Ruby-Dateien in den beiden Verzeichnissen.

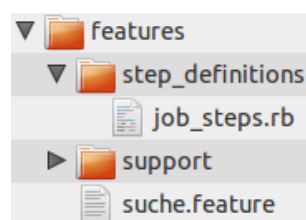


Abbildung 9: Aufbau eines Cucumber- Verzeichnisses

1. Definition des gesamten Akzeptanztest Zusammen mit dem Kunden, oder basierend auf den Anforderungen entwickeln wir zuerst eine Spezifikation für ein Feature, und darauf aufbauend, Da die Akzeptanztests in erster Linie dazu dienen, die Software gegenüber den Anforderungen des Kunden zu validieren, und auch als Kommunikationsmittel genutzt wird, orientiert sich das Vokabular an gebräuchlichen Begriffen. So ist ein einzelner Testfall ein „Szenario“ und eine Test-Suite ein „Feature“. Statt „Assertions“ (Zusicherungen) gibt es Vor- und Nachbedingungen.

Nachfolgend sei ein erstes Szenario für eine Suche nach Stellenanzeigen gezeigt.

```
features/job.feature
1 # language: de
2 Funktionalität: Job-Suche
3   Um Jobs zu finden
4   Als ein Gast
5   Soll es möglich sein mittels einer Suche Jobs zu finden
6   Szenario: Auffinden durch Titel
7     Angenommen wir befinden uns auf der Startseite
8     Und die folgenden Jobs sind vorhanden:
9       | title | visible |
10      | Ruby on Rails Entwickler | true |
```

```
11 | Java Programmierer | true |
12 Wenn wir "Rails" für "search" eintippen
13 Und wir auf den Button "Suchen" klicken
14 Dann sehen wir "Ruby on Rails Entwickler"
```

Listing: Definition eines Szenarios in einem Cucumber-Feature

Die ersten drei Zeilen des Features („Funktionalität“) beinhalten hier einen Kommentar, der lediglich die Testziele und Rahmenbedingungen definiert. Danach folgen die Testschritte. Eine Besonderheit ist der zweite: „Und die folgenden Jobs sind vorhanden“. Dort ermöglicht es ein Syntaxelement von Cucumber über eine Tabelle mehrere Datensätze zu definieren. Alle anderen Elemente sollten aus dem Abschnitt 4.3.2 bereits bekannt sein.

Wenn wir dieses Feature nun ausführen, erhalten wir folgende Ausgabe:

```
1 scenario (1 undefined)
5 steps (5 undefined)

You can implement step definitions for undefined steps with these
snippets:

Angenommen /^wir befinden uns auf der Startseite$/ do
  pending # express the regexp above with the code you wish you had
end
Angenommen /^die folgenden Jobs sind vorhanden:$/ do |table|
  # table is a Cucumber::Ast::Table
  pending # express the regexp above with the code you wish you had
end
Wenn /^wir "([^"]*)" für "([^"]*)" eintippen$/ do |arg1, arg2|
  pending # express the regexp above with the code you wish you had
end
Wenn /^wir auf den Button "([^"]*)" klicken$/ do |arg1|
  pending # express the regexp above with the code you wish you had
end
Dann /^sehen wir "([^"]*)"$/ do |arg1|
  pending # express the regexp above with the code you wish you had
end
```

Bevor wir also beginnen können das Feature zu entwickeln, müssen wir erst die Testschritte implementieren. Cucumber hat uns schon ein paar Codeschnipsel generiert, mit denen wir unsere eigenen Testschritte implementieren können. Deren Implementierung ist i.d.R davon abhängig, welche zugrunde liegende Methode man verwendet,

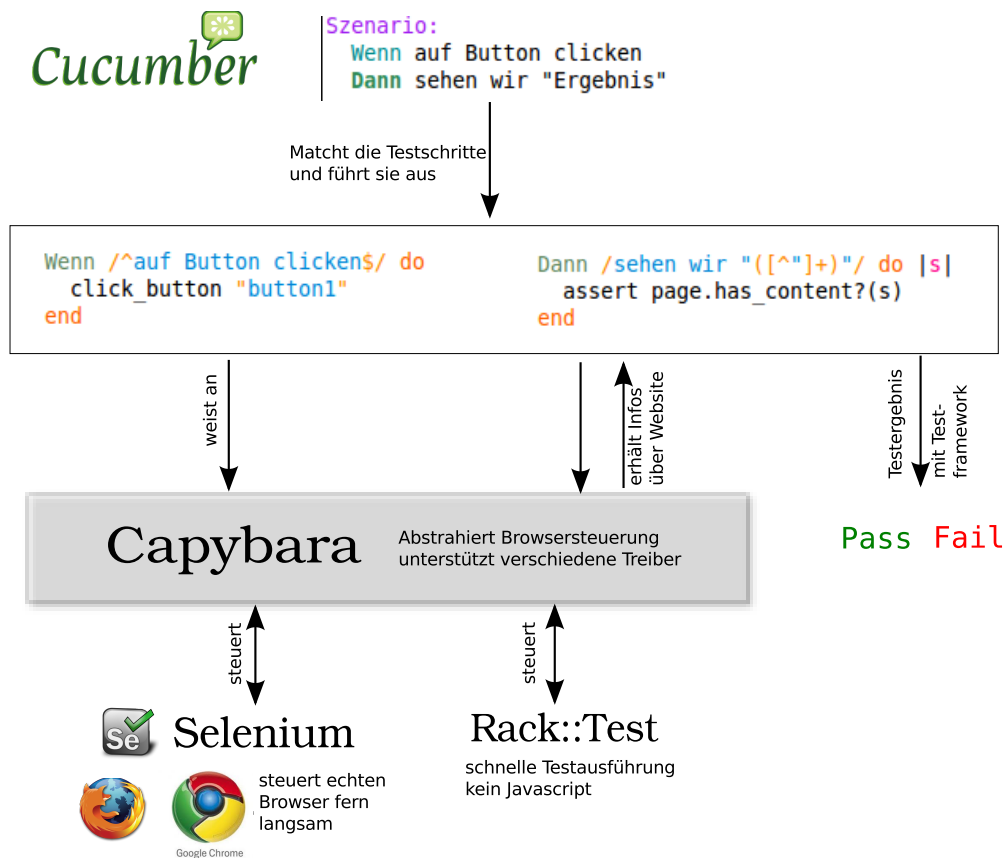


Abbildung 10: Ablauf beim Akzeptanztest mit Cucumber und Capybara

einen Browser zu simulieren, und welches Test-Framework für Zusicherungen benutzt wird. Wir haben uns für Capybara entschieden, welches sehr flexibel bei der Wahl der Browser-Engine ist, und für die Definition der Zusicherungen für das schon bekannte Test::Unit. Der Ablauf des Akzeptanztest ist in Abbildung 10 abgebildet. Die Anweisungen des Akzeptanztests werden auf die definierten Testschritte gemappt, die Cucumber dann sequentiell ausführt.

Hier zuerst die Basis-Testschritte, die für fast jedes Feature benötigt werden. Falls wir die englische Testschrittdefinition nutzen, entfällt dieser Schritt, da Capybara bereits häufig gebrauchte Testschritte mitliefert.

```
features/step_definitions/job_steps.rb
1 Angenommen /^wir befinden uns auf der Startseite$/ do
2   visit "/"
3 end
4 Wenn /^wir "([^"]*)" für "([^"]*)" eintippen$/ do |text, input_name|
5   fill_in input_name, :with => text
6 end
7 Und /^wir auf den Button "([^"]*)" klicken$/ do |text|
8   click_button text
9 end
10
11 Dann /^sehen wir "([^"]*)"$/ do |string|
12   assert page.has_content?(string)
13 end
```

Listing: Implementierung der Testschritte zur Interaktion mit einem Browser

Der erste Testschritt gibt unserem simulierten Browser die Anweisung, die Startseite zu besuchen. Der Zweite Testschritt spezifiziert einen Testschritt mit 2 variablen Texten, die an unseren Block mit übergeben werden. Wir möchten, dass der erste Begriff in „...“ als Text in ein Formularelement mit dem Bezeichner des zweiten Begriffes eingetragen wird. Der Dritte implementiert den Klick auf einen HTML-Button mit dem angegebenen Inhalt. Während die ersten 3 Schritte Aktionen ausführen, hat der 4. eine andere Funktion: Er spezifiziert eine Zusicherung, dass auch unserer aktuellen Webseite irgendwo ein gewisser Inhalt steht.

```
features/step_definitions/job_steps.rb
1 # table.hashes ->
2 # [ {:title => "Ruby on Rails Entwickler", :visible => true},
3 #   {:title => "Java Programmierer", :visible => true}]
4 Angenommen /^die folgenden Jobs sind vorhanden:$/ do |table|
5   valid_job = jobs(:visible_job).attributes
6   table.hashes.each do |hash|
```

```
7   attributes = valid_job.merge(hash)
8   Job.create(attributes)
9   end
10 end
```

Listing: Testschritt zum Anlegen von Testdaten auf Basis der Cucumber Definition

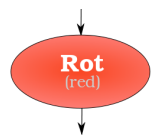
Der Testschritt für die Implementation des Testschrittes mit der Tabelle, ist etwas umfangreicher. Wir bekommen von Cucumber ein Array von Hashes übergeben, die die gepasste Tabelle beinhaltet. Über diese können wir nun mit dem Iterator „each“ iterieren, und die definierten Attribute mit denen unserer Job-Fixture³¹ vereinigt wird, um somit einen neuen Job in der Testdatenbank anzulegen. Wir haben den Testschritt so allgemein gehalten, dass wir ihn in späteren Szenarien gut verwenden können, um weitere Jobs mit speziellen Attributen zu generieren.

Nun können wir Cucumber erneut ausführen, und das Ergebnis ist ein anderes:

```
Wenn wir "Rails" für "search" eintippen      # features/step_definitions/
  job_steps.rb:5
  cannot fill in, no text field, text area or password field with id,
    name, or label 'search' found (Capybara::ElementNotFound)
    (eval):2:in 'fill_in'

Failing Scenarios:
cucumber features/suche.feature:6 # Scenario: Auffinden durch Titel

1 scenario (1 failed)
5 steps (1 failed, 3 skipped, 1 passed)
```



Der Test schlug also fehl, da noch kein Suchfeld eingebaut wurde. Dies lösen wir, indem wir auf der View der Startseite eines einbauen:

```
app/views/layouts/application.html.erb
1 ...
2 <div id='search-field'>
3   <%= form_tag "/jobs" do %>
4     <%= text_field_tag(:search) %>
5     <%= submit_tag("Suchen") %>
6   <% end %>
7 </div>
```

Listing: Implementation eines Suchfeld im Layout der Webseite

³¹Anmerkung: Für diese Aufgabe eignen sich Factories deutlich besser. Da mit Fixtures schon eine Testdatengenerierung eingeführt wurde, bleiben wir aber aus Konsistenzgründen dabei.

Wir implementieren im Layout der Applikation ein Suchfeld mit den Form-Hilfsfunktionen, die Rails anbietet. Das Layout ist eine Basisview, die für alle Webseiten generiert wird. Diese beinhaltet Elemente der Website, die überall benutzt werden, z.B. eine Navigation, Einbindung von Stylesheets oder wie hier, ein Suchfeld, das auf jeder Seite erscheinen soll. Dabei verwenden wir ERB³², eine Template-Sprache, die normalerweise HTML-Code erzeugt, und es ermöglicht innerhalb der „<%= ... \>“ Ruby-Code auszuführen.

Jetzt folgt die Implementierung der Suche im Jobs-Controller:

```
app/controllers/jobs_controller
1 class JobsController
2   def index
3     @jobs = Job.where("title like ?", "%#{params[:search]}%")
4   end
5   ...
```

Listing: JobsController am Ende der Testphase

Im Jobs-Controller implementieren wir die Bereitstellung der Jobs mittels einer Datenbankabfrage. Dabei nutzen wir die Datenbankfunktion von ActiveRecord und führen eine partielle Match-Suche der Suchparameter über die Tabellenspalte „title“ aus. ActiveRecord übernimmt das sichere Escapen des übergeben Strings selbstständig in den dafür vorgesehenen Platzhalter „?“³². Das Ergebnis speichern wir in der Instanzvariable @jobs, und stellen es so der View zur Verfügung.

Damit nun der Test besteht, muss als letztes noch eine View implementiert werden, die (zumindest) die Titel aller in der Suche gefundenen Jobs ausgibt, hier z.B. als Überschrift.

```
app/views/jobs/index.html.erb
1
2 <% @jobs.each do |job| %>
3   <div class='job'>
4     <h3><%= job.title %></h3>
5   </div>
6 <% end %>
```

Listing: View-Implementation für eine Liste von Jobs

Mittels des Iterators „each“, geben wir für jeden gefundenen Job ein DIV mit einer darin enthaltenen H3-Überschrift des Titels aus.

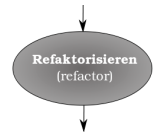
Cucumber bestätigt uns nun die erfolgreiche Implementation des Features:

³²Embedded Ruby, eine Syntax für Rails Views



```
1 scenario (1 passed)
5 steps (0 failed, 0 skipped, 5 passed)
```

Auch innerhalb der Akzeptanztestgetriebenen Entwicklung ist Refaktorisieren ein fester Bestandteil innerhalb des Entwicklungszyklus. Ein guter Ansatz wäre z.B. die Suchlogik aus dem Controller in die Job-Modelklasse auszulagern.



```
app/models/job.rb
1 class Job
2   ...
3   def self.perform_search(params)
4     where("title like ?", "#{params[:search]}")
5   end
```

Listing: Auslagerung der Suchlogik in die Job-Klasse

Wir lagern die Suchmethode in eine neue Klassenmethode der Job-Klasse aus. Klassenmethoden werden definiert, indem „self.“ vor den Methodennamen geschrieben wird.

```
app/controllers/jobs_controller.rb
1 class JobsController
2   def index
3     @jobs = Job.perform_search(params)
4   end
5 end
```

Listing: Finaler Jobcontroller nach Refaktorisierung

Nun können wir die neu implementierte Funktion in unserem Controller verwenden.

Ausblick Nun haben wir ein erstes Szenario mit Cucumber implementiert. Da wir schon ein paar Testschritte implementiert haben, ist die Entwicklung von weiteren Szenarien leichter. Die Tests wurden bisher durch einen simulierten Browser RackTest ausgeführt. Da wir Capybara als Abstraktion des Browsers genutzt haben, wäre eine Umstellung auf eine reale Browserumgebung, z.B: Firefox oder Chrome, kein Problem. Deren Ausführung ist zwar deutlich langsamer als bei RackTest, dafür findet hier der Test unter realen Bedingungen, nämlich in einem Browser, wie ihn auch später Nutzer der Web-Anwendung besitzen.

7.5. Testen von Javascript Ereignissen

Mit dem ständig zunehmenden Grad der Entwicklung von Javascript-Frameworks und den in den letzten Jahren zunehmenden Ausführungsgeschwindigkeiten von Javascript in den einzelnen Browsern, ist Javascript bei der Programmierung von modernen Webanwendung ein immer wichtigerer Teil geworden. Das Aufkommen von Cloud-Computing propagiert die Ablösung traditioneller Desktop-Software durch Browser-basierte Anwendungen. So beinhaltet fast jede moderne Webanwendung Javascript-Funktionalität.

Meist geschieht dies durch Ajax³³, also durch asynchrone Kommunikation mit einem Server und Aktualisieren der Seite ohne dass diese neugeladen werden muss. Durch sehr gute Frameworks, wie z.B. jQuery, dass ab Rails 3 standardmäßig Teil der Rails-Distribution ist, sind solche Aktualisierungen meist in sehr wenigen Zeilen zu implementieren. Aber auch diese Zeilen sind Teil unseres Applikationscodes, und es entsteht ein Bedürfnis diesen Teil ebenfalls zu testen.

Test im Rahmen von Systemtests Die für uns jetzt leichteste Methode, ist das Testen von Javascript im Rahmen unserer System/Akzeptanztest, wie bereits im vorherigen Abschnitt beschrieben wurde. Dafür müssen wir allerdings eine Browser-Engine wählen, die Javascript ausführen kann. Für die Browsersimulation nahmen wir Capybara, haben also nun leicht die Möglichkeit die Engine zu wechseln. So kann z.B. mittels der Middleware Selenium Firefox oder Chrome simuliert werden. Dies verzögert zwar unsere Testausführung, stellt aber sicher, dass auch Javascript in einer Umgebung ausgeführt wird, die der letztendlichen Realen des Benutzers sehr nahe kommt.

Dedizierte Unittests für Javascript Falls eine Anwendung nicht nur einfache Ajax-Funktionen der Art „Klicke Link – Sende Anfrage an Server – setze fertiggenerierten HTML Code der Antwort direkt in den DOM ein“, sondern Funktionen, die klassische Desktopanwendungen nachbilden, gefordert sind, ist der Einsatz von eigenen Unittests für Javascript angebracht. Diese Art der Anwendungsgestaltung gewinnt zunehmend an Popularität, z.B. das aktuelle Twitter-Layout oder Google Mail machen nach dem initialen Laden der Webseite keine weitere komplette Seitenaktualisierung mehr. Auch hier bieten sich einige Javascript Test-Frameworks an, z.B. JsUnit³⁴, einem Vertreter

³³Asynchrones Javascript und XML (oder JSON)

³⁴<http://www.jsunit.net/>

der xUnit-Reihe, und Jasmine, einem Vertreter aus dem BDD-Kontext.

Bei IT-Jobs beschränkt sich der Einsatz von Javascript allerdings auf die erwähnten einfachen Funktionen.

Diskussion Das explizite Testen von Javascript ist zwar nicht Thema dieser Diplomarbeit, aber im Kontext der Entwicklung einer Webanwendung sollte Javascript nicht unerwähnt bleiben. Für die vorliegende Anwendung IT-jobs hat die Methode, Javascript im Rahmen von Systemtests zu testen, ausgereicht. Für komplexe Javascriptanwendungen sind aber unbedingt dedizierte Unittests erforderlich.

8. Auswertung

In diesem Abschnitt wollen wir nun verschiedene Resultate des Projekts IT-Jobs im Bezug auf Code-Qualität untersuchen und vergleichen.

Zuerst werden verschiedene Metriken, welche im Verlauf der Entwicklung regelmäßig festgehalten wurden, näher betrachten. Im weiteren Teil vergleichen wir die erreichten Maßzahlen mit denen früherer Projekte der pludoni GmbH und einiger bekannter OpenSource Ruby-Projekte, um die Ergebnisse besser einordnen zu können.

8.1. Entwicklungsstand ITjobs

Zum Zeitpunkt der Beendigung dieser Arbeit sind folgende Teilmodule abgeschlossen:

- Modul: Statischer Content. Blog und statische Seiten, sowie die Möglichkeit variablen Inhalt auf beliebigen Seiten innerhalb der Sidebar darzustellen
- Modul: Job, Basis. Implementation des Datenbankschemas für Jobs, und die Möglichkeit Jobs über das Interface anzulegen
- Modul: Feedimport. Möglichkeit, Jobs mittels einer XML-Schnittstelle in das System einzuspielen. E-Mail Benachrichtigung im Fehlerfalle an den Inhaber der XML-Feeds

Noch offen für eine zukünftige Entwicklung sind:

- Modul: Bewerbung. Bewerber können sich über die Webseite bewerben. Ebenfalls können sie sich mit Facebook und LinkedIn verbinden, um automatisch ihre Stammdaten eintragen zu lassen
- Modul: Bezahlungssystem: Beim Anlegen eines Jobs wird dieser über einen externen Dienstleister abgewickelt.

Während der Arbeiten an IT-Jobs wurden andere Aufgaben priorisiert, und deswegen der Abgabetermin für das System nach 2012 verlegt. Nichtsdestotrotz soll IT-Jobs als Basistechnologie und Referenzprojekt für weitere Rails-Anwendungen dienen, und die Entwicklung kann insgesamt als Erfolg bewertet werden, auch wenn sie zum gegenwärtigen Zeitpunkt noch nicht abgeschlossen ist.

An der Entwicklung beteiligt waren 2 Programmierer, einschließlich des Autors. Beiden Programmierern gelang es, sich mit der Testgetriebenen Entwicklung und mit Rails vertraut zu machen und wertvolle Erfahrungen zu sammeln.

Code-Qualität und Testabdeckung

Während der Entwicklung wurden in regelmäßigen Abständen Code-Metriken festgehalten, und können nun so über den Verlauf der Entwicklung dargestellt werden.

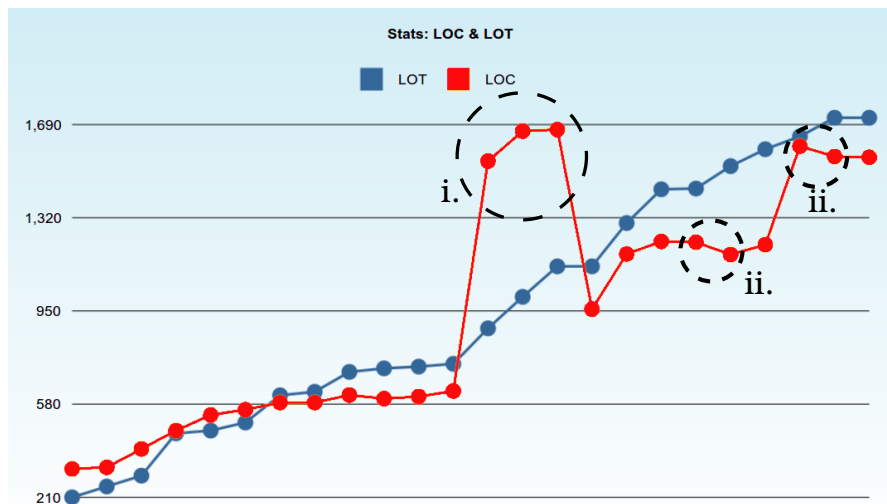


Abbildung 11: Entwicklung des Umfangs des Test- und Programmcodes

In Abbildung 11 ist der Verlauf der Codezeilen und der Testzeilen abgebildet. Während in den ersten Tagen der Entwicklung, wegen dem Einsatzes von Codegeneratoren und einer langsamen Gewöhnung an den TDD Ablauf, noch weniger Testcode als Programmcode geschrieben wurde, verlaufen die Linien ab dann nahezu proportional. Einen Ausreißer stellen die mit i. markierte Codezeilen dar: Hier ist eine Fehlkonfiguration der Statistikberechnung die Ursache dafür, dass Fremdcode mitgezählt wurde, der nicht Teil der Anwendung war. Eine andere interessante Erkenntnis ist, dass zwar die Lines Of Test monoton steigend ist, die Lines Of Code dagegen durch Refaktorisierungen abgefallen sind (z.B. mit ii. markierte Stellen). Die hohe Testdichte hat eine Refaktorisierung ermöglicht, und in diesen Phasen konnte effektiv Design zu dem System hinzugefügt werden.

In Abbildung 12 ist der Verlauf der Testausführungsabdeckung über den Verlauf der Entwicklung abgebildet. Insbesondere in den ersten Tagen gab es einige Probleme, die davon verursacht wurden, dass das Tool mit dem die Abdeckung gemessen wurde, „rcov“ nicht kompatibel mit der aktuellen Ruby Version 1.9.2 ist. RCov lieferte deshalb falsche Werte liefert. Später erfolgte eine Umstellung auf SimpleCov (bereits

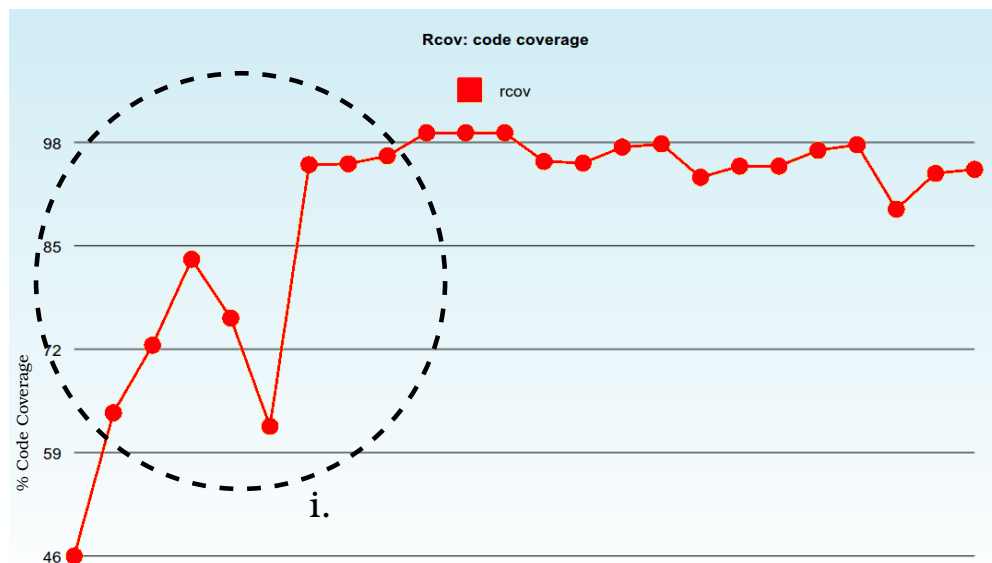


Abbildung 12: C0 Testabdeckung über den Verlauf der Entwicklung

in Abschnitt 6.2 vorgestellt). Ab dann lag die Testabdeckung innerhalb von 90% bis 100%.

Ein Indikator für die Code-Qualität sind die Anzahl und Arten von Code-Smells. Diese wurden ebenfalls über den Entwicklungszeitraum gemessen und sind in Abbildung 13 dargestellt. Insgesamt betrachtet, stiegen diese mit zunehmender Codemenge langsam an. Dank der Messung durch Code-Metriken konnte einige suboptimale Stellen nach einer gewissen Zeit refaktoriert werden. So ist das wiederkehrende Muster, dass die Anzahl der Vorkommen eines Codesmells erst anstieg, und nach einigen Tagen wieder abfiel. Das Ergebnis gilt als positiv zu bewerten, da die Menge an Code-Smells langsamer anstieg, als die dazugehörige Menge an Code (LOC, vgl. Abbildung 11). Die Code-Smells wurden mit dem Werkzeug Reek gemessen, und die Definition der Code-Smells sowie Informationen über das jeweilige Messverfahren entnehmen sie bitte [Rutherford, 2010].

Diskussion

Falls das Projekt in einem reinen TDD-Vorgehen entwickelt worden wäre, dann hätte die Testabdeckung 100% betragen müssen. Ist war aber nicht immer der Fall. Gründe hierfür sind:

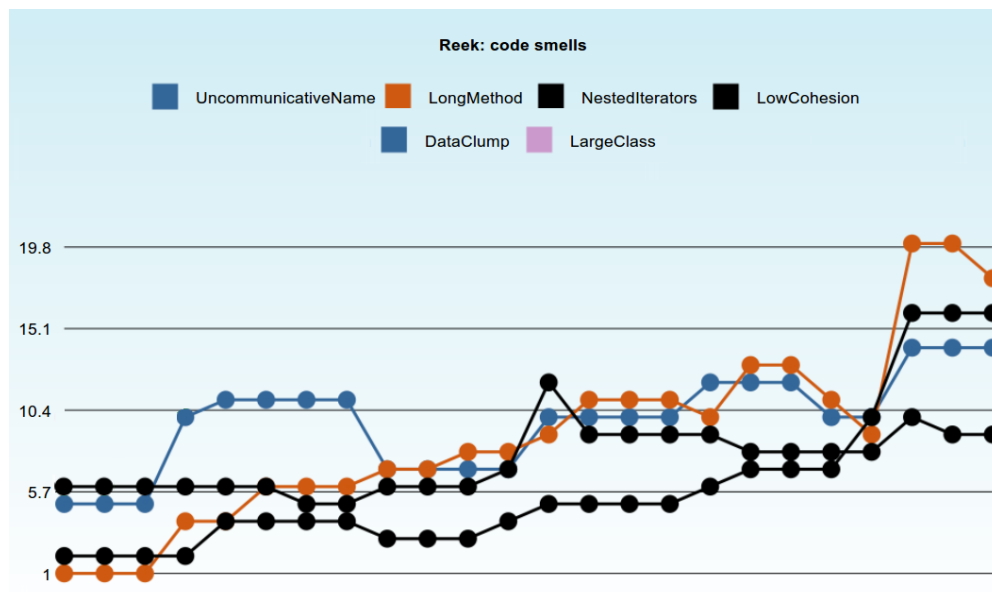


Abbildung 13: Verlauf der verschiedenen Klassen an Code-Smells über den Entwicklungszeitraum

- Nutzung von Codegeneratoren. Bei der Erstellung der Gerüste und der Programmierung einfacher Administrationsinterfaces waren nicht immer Testfälle mitgeneriert worden. Stellenweise wurde dieser generierte Code später neu in TDD geschrieben, allerdings nicht in allen Fällen
- Vorher waren schon einige Erfahrung mit Rails vorhanden. Allerdings sind in der neusten Version von Rails einige Features dazu gekommen, und wurden nun einige Bibliotheken verwendet, die vorher noch nicht bekannt waren. Aus der Unkenntnis entstand so viel Code, der eigentlich zum Experimentieren mit diesen neuen Features gedacht war. Diese Spikes (siehe Abschnitt 3.3) dienen dem Erkunden neuer Funktionalitäten und dem Prototypisieren. Diese sollten normalerweise nicht in den Hauptzweig der Entwicklung mit eingecheckt werden. In zukünftigen Projekten, die nach dem Prinzip der kontinuierlichen Integration stattfinden werden (siehe Abschnitt 6.1.2), wird ein Einchecken dieser Spikephasen in den Hauptzweig der Entwicklung nicht erfolgen.
- Trotz bisheriger Erfahrungen mit Ruby und Rails waren nur wenig Erfahrungen zum Thema Testen und Testgetriebene Entwicklung vorhanden. Insbesondere

die Nutzung von [†]Test-Doubles musste erst gelernt werden, da ohne eine solche bestimmte Testszenarien schwierig bis überhaupt nicht zu testen sind.

Trotzdem ermöglichte die vorhandene Testdichte eine sichere Ausgangsbasis für ständige Refaktorisierungen. Ein Ergebnis dessen ist, dass die Anzahl der Code-Smells im akzeptablen Rahmen geblieben ist. Eine gewisse Menge an Komplexität lässt sich nicht vermeiden, und Code-Smells können hier nur ein Indikator sein. In vielen Fällen muss direkt am Code entschieden werden, ob ein gewisser Code-Smell für eine gewisse Stelle relevant ist oder nicht.

8.2. Code-Quality-Benchmark mit anderen Ruby-Projekten

Um die Ergebnisse besser einordnen zu können, vergleichen wir die Ergebnisse aus den Code-Quality-Benchmark mit vergangenen Ruby-Projekten in der pludoni GmbH und mit bekannten Ruby/Rails-OpenSource Applikationen und Bibliotheken.

Eine vollständige Metrik wie oben an IT-Jobs, war in vielen Fällen nicht möglich, da die Testwerkzeuge in vielen Fällen Inkompatibilitäten mit neueren Rubyversionen haben. Wir beschränken uns deshalb auf eine exemplarische Überblicksmetrik, bestehend aus:

LOC Anzahl der Quellcodezeilen, gemessen durch das Rails-Kommando „rake stats“.

LOT Anzahl der Quellcodezeilen aller Tests, gemessen durch das Rails-Kommando „rake stats“.

LOT/TOC Verhältnis aus Testzeilen und Quellcodezeilen

AVGComplexity Durchschnittliche Komplexität aller Klassen, gemessen durch **Flog**³⁵. Flog besitzt ein eigenes Maß für Code-Komplexität. Es vergibt für jede Zuweisungen, Verzweigungen und Funktionsaufrufe unterschiedlich viele Punkte, und bildet so eine Summe per Funktion oder per Klasse. Dabei vergibt Flog besonders viele Punkte für schwer nachzuvollziehende Funktionsaufrufe, wie z.B. „eval(string)“, welches einen String als Ruby-Code auswertet.

H5Complexity Komplexität der 5% komplexesten Klassen, nach Flog

DSmell Anzahl der Code-Smells nach Roodi. Beinhaltet u.a.: hohe Cyclomatische Komplexität in einer Methode (min. 4), lange Methoden, lange Parameterlisten, u.v.m.³⁶

³⁵<https://github.com/seattlerb/flog>

³⁶http://roodi.rubyforge.org/files/README_txt.html

DSmell/KLOC Anzahl der Code-Smells nach Roodi pro tausend Codezeilen

CSmell Anzahl der Code-Smells nach Reek. Diese beinhalten: Geringe Kohäsion, Duplikation, Control-Couple, Unkommunikativer Name von Methoden/Variablen/Parameter, verschachtelte Iteratoren, u.v.m.³⁷

CSmell/KLOC Anzahl der Code-Smells nach Reek pro tausend Codezeilen

8.2.1. Vergleich von IT-Jobs-und-Stellen mit eigenen Projekten

Bisherige Projekte der pludoni GmbH und des Autors basierend auf Ruby/ Ruby on Rails

feedimport Ist die fertiggestellte Bibliothek, die im Rahmen der Entwicklung von IT-Jobs ausgelagert wurde, um den Community-Portalen bereits jetzt zur Verfügung zu stehen. Die Grundzüge wurden in Abschnitt 7.3 vorgestellt.

feedmerger Eine Rails-Anwendung zur Verwalten, Cachen, Filtern und Zusammenfügen von RSS und Atom-Feeds. Als OpenSource unter <https://github.com/zealot128/WenShanWenHai> zu finden.

pludonidb Ist eine auf ActiveRecord basierende Bibliothek zur Anbindung der Datenbanken der Communityportale an Ruby-Skripte. Beinhaltet außerdem weitere Hilfsfunktionen für Berechnung von Tag-Wolken, Häufigkeiten u.ä.

backlinks Backlink und SEO-Success-Control ist eine Rails-2.3 Anwendung, zum Messen der sogenannten Backlinks³⁸ der Communitymitglieder und der Platzierung für relevante Sucheingaben in den großen Suchmaschinen (konkret: Welchen Platz hatte ein Communityportal an einem bestimmten Tag für „it jobs dresden“, usw.). Details dazu sind im Praktikumsbericht des Autors zu finden.

SiteAnalyzer (SAnalyzer) ist ein Werkzeuge zur Analyse von kompletten Websites/Webdomains, mit dem Ziel tote Links zu finden, alle Seiten auf HTML-Gültigkeit zu prüfen, und insgesamt die Link-Architektur zu analysieren. Auch dieses Projekt ist als OpenSource verfügbar: <https://github.com/zealot128/Site-Structure-Analyzer>.

³⁷<https://github.com/kevinrutherford/reek/wiki/Code-Smells>

³⁸Backlinks sind Links anderer Webseiten auf die eigenen. In den Communityportalen sind die Mitglieder vertraglich verpflichtet, einen Backlink auf die Communityportale anzulegen.

8.2.2. Vergleich von IT-Jobs-und-Stellen mit anderen Rails Projekten

Weiterhin wird das Projekt mit folgenden beliebten Webprojekten, welche ebenfalls auf Ruby on Rails basieren, verglichen:

lpp Linkpartnerprogramm ist eine Studenteninitiative der TU Dresden, um ausländische Studenten mit deutschen Sprach- und Lernpartnern zusammenzubringen. Dazu wurde 2008 von einem Studententeam im Rahmen einer Semesterarbeit eine Rails-Webanwendung geschrieben. Der Autor dieser Arbeit war zwar nicht an der Entwicklung beteiligt, übernahm aber die Pflege und Weiterentwicklung selbiger.

diaspora Ist ein verteiltes soziales Netzwerk.

Code: <https://github.com/diaspora/diaspora>

bucketwise (bucket) ist ein web-basierter persönlicher Finanzmanager mit Budgetierung nach dem Briefumschlagsystem

Code: <https://github.com/jamis/bucketwise>

chiliproject (chilli) Ist ein Fork³⁹ des sehr beliebten Bugtrackers Redmine⁴⁰. Statt des originalen Redmines wurde dieser Fork genommen, da er eine neuere Codebasis hat

Code: <https://github.com/chiliproject/chiliproject>

railscast (rCasts) Ist der Code der Website <http://www.railscasts.com>, in welche allwöchentlich ein Screencast zum Thema Ruby und Rails veröffentlicht wird (bis dato 283 Episoden).

Code: <https://github.com/ryanb/railscasts>

ActiveSupport [aSupport] Beinhaltet Hilfsklassen und Erweiterungen der Ruby-Standardbibliothek und ist Kernbestandteil von Ruby on Rails

Code: <https://github.com/rails/rails.git>

ActionPack (aPack) ist ein Framework um Anfragen und Antworten eines Webservers zu verarbeiten und ist ein weiterer Kernbestandteil von Rails.

Code: <https://github.com/rails/rails.git>

Die Auswahl erfolgt z.T. sehr willkürlich. Als Ansatzpunkt diente die Liste der beliebtesten Rails-Projekte, die auf Github gehostet sind. Weiterhin wurde ActiveSupport

³⁹In der (Open-Source) Software-Entwicklung ist ein Fork eine legale Kopie eines bestehenden Software-Produktes, um von nun an eine unabhängige Entwicklung zu betreiben

⁴⁰<https://github.com/edavis10/redmine>

und ActiveSupport ausgewählt, um die Code-Qualität von Rails selbst mit zu beurteilen.

8.2.3. Auswertung und Visualisierung

In den Abbildungen 14, 15 und 16 werden die gesammelten Ergebnisse für alle genannten Projekte visualisiert. Der obere Teil jeder Abbildung enthält die externen Projekte, der untere alle Projekte der pludoni GmbH.

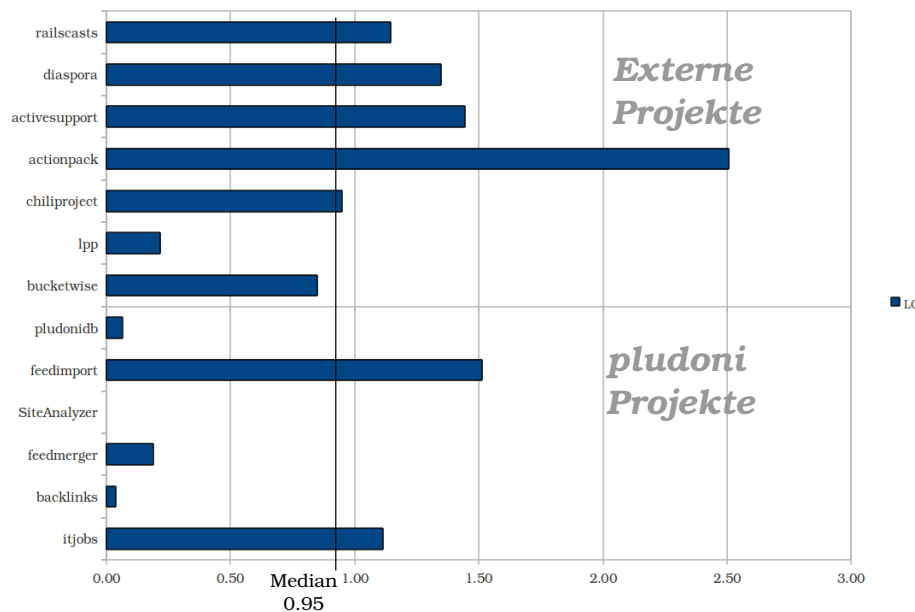


Abbildung 14: Vergleich des Verhältnisses aus Anzahl Testcodezeilen / Anzahl Codezeilen

Abbildung 14 zeigt das Verhältnis aus Testcode zu Programmcode. Der Median über alle Projekte ist 0.95. Ausreißer sind hierbei vier der internen Projekte pludonidb, SiteAnalyzer, feedmerger und backlinks, die wenig bis gar keine automatisierten Tests besitzen. Als Gegenbeispiel sei ActionPack zu nennen, dass über 2.5x soviel Testcodezeilen wie Programmcode verfügt. Gegenüber den bisherigen Projekten in der pludoni GmbH kann man anhand der Projekte Feedimport und IT-Jobs, die im Rahmen dieser Diplomarbeiten entstanden, sehen, dass die bloße Anzahl an Tests stark zugenommen hat. Beide Projekte wurden mittels TDD entwickelt. Gegenüber dem Feedimport verfügt IT-Jobs aber über mehr Code, der durch das Framework bereits zur Verfügung

Tabelle 2: Vergleich von IT-jobs mit anderen Ruby Projekten des Autors/der pludoni GmbH

Projekt	ItJobs	Backlinks	Feedmerger	SAnalyzer	Feedimport	PludoniDb
Technologie	Rails3	Rails2	Rails3	Rails3	Ruby	Ruby
Testverfahren	TDD	manuell	manuell	manuell	TDD	manuell
LOC	1570	1884	724	214	571	1933
LOT	1750	75	137	0	865	126
LOT/LOC	1.11	0.04	0.19	0.00	1.51	0.07
AVGCmplx	8.00	19.90	10.50	14.80	16.80	13.70
H5Cmplx	43.80	70.60	26.90	39.70	25.10	113.20
DSemll	11	69	11	11	8	25
DSmell/KLOC	7.00	36.60	15.20	51.40	14.00	12.90
CSmell	53	152	52	13	28	113
CSmell/KLOC	33.80	80.70	71.80	60.70	49.00	58.50

Tabelle 3: Vergleich von IT-jobs mit Ruby/Rails Projekten aus der Community

Projekt	itjobs	bucket	lpp	chili	aPack	aSupport	diaspora	rCasts
Technologie	Rails3	Rails2	Rails1	Rails2	Rails3	Rails3	Rails3	Rails3
LOC	1570	1979	7116	21201	12995	9407	7466	653
LOT	1750	1684	1557	20127	32570	13590	10072	748
LOT/LOC	1.11	0.85	0.22	0.95	2.51	1.44	1.35	1.15
AVGCmplx	8.00	12.80	17.10	19.10	11.10	10.90	13.10	11.00
H5Cmplx	43.80	32.20	181.80	84.40	64.00	47.90	54.00	34.60
DSmell	11	37	250	651	370	166	99	2
DSmell /KLOC	7.00	18.70	35.10	30.70	28.50	17.60	13.30	3.10
CSmell	53	129	386	1535	982	291	324	42
CSmell /KLOC	33.80	65.20	54.20	72.40	75.60	30.90	43.40	64.30

gestellt wird, während der Feedimport eine weitestgehend eigene Objekthierarchie besitzt, und damit auch mehr Tests bedurfte.

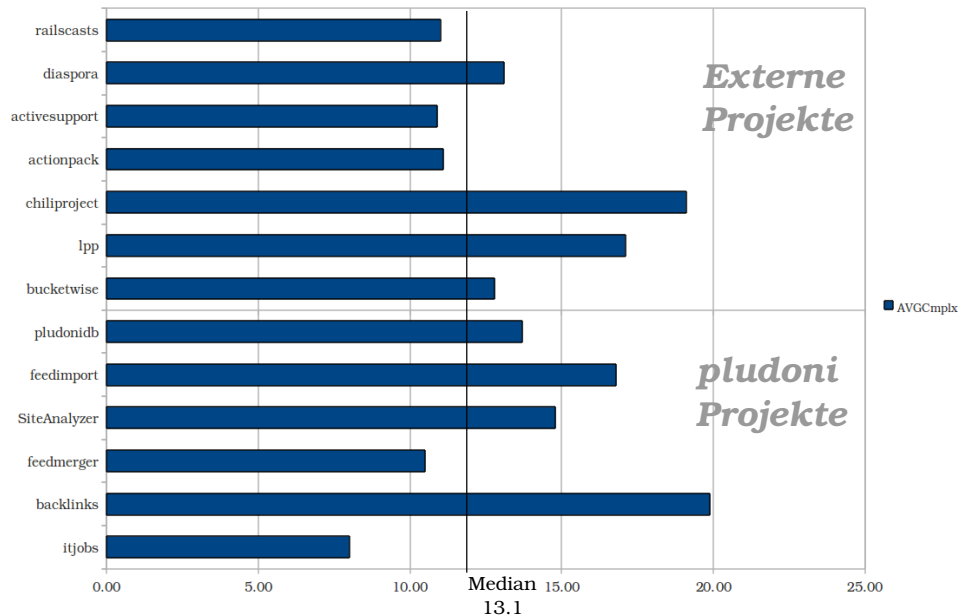


Abbildung 15: Vergleich der durchschnittlichen Komplexität

In Abbildung 15 wurde die durchschnittliche Komplexität nach dem Flog-Verfahren gegenübergestellt. Der Median war hier 13,1. Unter den eigenen Projekten sind hier Backlinks und der Feedimport die komplexesten Projekte. IT-Jobs und Feedmerger, die beide Rails-Projekte sind, haben dagegen eine geringere Code-Komplexität. Die Ursache ist hierbei, dass während der Entwicklung von IT-Jobs ständig die Code-Qualität durch die Metriken gemessen wurde, und im Falle von Ausreißern gegengesteuert wurde. Bei dem Feedimport, der zwar durch TDD entwickelt wurde, wurde eine solche Analyse erst am Ende der Entwicklung durchgeführt. Für uns ist das ein Indiz, dass das Nutzen einer Testgetriebenen Entwicklung kein alleiniger Garant für eine gute Code-Qualität ist.

Bei den externen Projekten war der Bugtracker-Anwendung Chiliproject/Redmine die komplexeste Anwendung, dicht gefolgt von der Sprachpartnervermittlung Linkpartnerprogramm. Gerade wenn man das Verhältnis aus Testcode zu Programmcode mit einbezieht, kann man eine Tendenz ableiten: Beide Projekte haben gegenüber den anderen externen eine relativ geringe Anzahl an Testcode und beide haben eine relativ

hohe Komplexität.

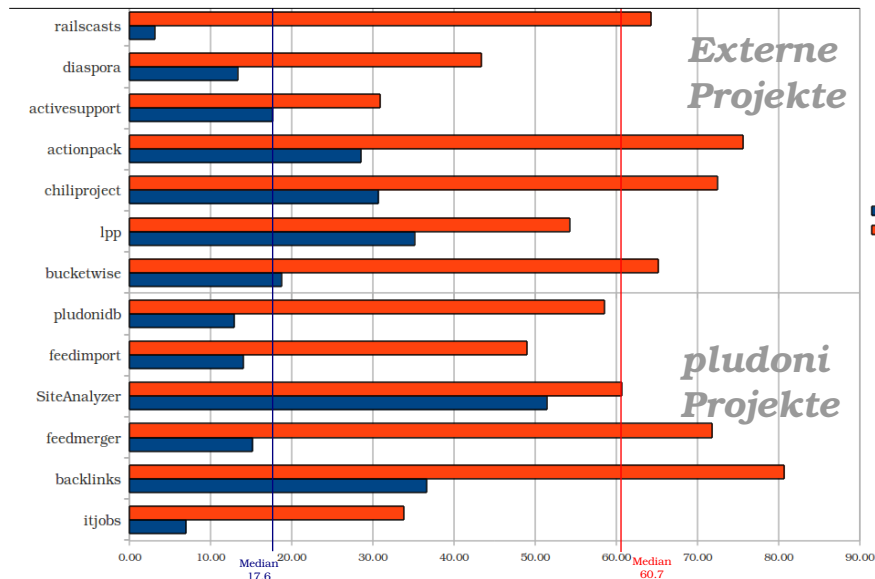


Abbildung 16: Vergleich der Anzahl Smells pro KLOC

Die Abbildung 16 zeigt zwei verschiedene Arten von Smells: Einerseits durch Reek gemessen (Rot) und andererseits durch Roodi gemessen (Blau). Zu bemerken sei, dass IT-Jobs eine deutlich geringere Code-Smell-Dichte ($1/2$ Median) besitzt. Auch der Feedimport ist noch im Rahmen und leicht unter dem Schnitt. Dagegen hat das interne Projekt Backlinks die höchste Dichte an Code-Smells. Interessanterweise hat dieses Projekt auch so gut wie gar keine automatisierten Tests, was eine nachträgliche Refaktorisierung sehr erschwert. Für das Projekt Feedimport existiert allerdings eine gute Testabdeckung, und so ließen sich leicht weitere Refaktorisierungen durchführen, sollte das gewünscht sein.

Actionpack verfügt unter den externen Projekten die höchste Smell-Dichte, trotz der großen Menge an Tests und der relativ geringen Komplexität. Da Actionpack auf das umfangreichste Projekt unter den untersuchten ist, wäre hier ein guter Ansatzpunkt für weitere Untersuchungen in der Sinnhaftigkeit gewissen Code-Smells. Die mit Abstand größte Code-Smell hierbei war Geringe Kohäsion bzw. starke Kopplung unter den Klassen. Eventuell lassen sich einige der Smells bei großen Modulen nicht vermeiden, da diese zwangsweise untereinander gekoppelt sein müssen.

Für statistisch signifikante Ergebnisse reichen die hier präsentierten Daten nicht aus, stattdessen sollte nur eine Tendenz untersucht werden. Insbesondere der Vergleich mit den internen Projekten hinsichtlich Code-Qualität. Das Gefühl der Programmierer, das z.B. das Projekt Backlinks eine relativ schlechte Codebasis verfügt, wurde durch die hier gezeigten Metriken bestätigt. Der aktuelle Stand von IT-Jobs dagegen lässt eine gute Ausgangsbasis für eine Weiterentwicklung vermuten.

9. Fazit

Die im Rahmen dieser Diplomarbeit entwickelte Web-Anwendung ist zum gegenwärtigen Zeitpunkt noch nicht fertiggestellt. Aufgrund einer Repriorisierung innerhalb der Firma, wird diese voraussichtlich 2012 weiter entwickelt werden. Gleichwohl kann die Entwicklung als Erfolg gewertet werden, da alle Qualitäts-Kennzahlen einen wesentlichen Unterschied gegenüber den bisherigen Projekten aufzeigt. Im Großen ist dies der Testgetriebenen Entwicklung und der ständigen Beobachtung der Code-Metriken zu verdanken. Dank der Testgetriebenen Entwicklung entstand relativ lose gekoppelter Code und ein feinmaschiges Testnetz, dank dessen Empfehlungen aus den Code-Metriken leicht umgesetzt werden konnten.

Allerdings hat sich gezeigt, dass die Umstellung auf eine Testgetriebene Entwicklung nicht von heute auf morgen stattfinden kann. Das konstante Testen-Vor-Implementieren bedarfs insbesondere in der Anfangsphase einer hohen Disziplin, da es zu leicht ist, in das traditionelle Schema zurückzufallen. Auch das Schreiben der Tests sollte gelernt werden. In dieser Arbeit wurden Tests lediglich in ihrer Quantität untersucht. Dies ist aber absolut nicht ausreichend für eine Beurteilung der Testvollständigkeit oder ein Beweis für eine Korrektheit des untersuchten Programmes.

Die Erfahrung und bisherige Studien haben aber gezeigt, dass durch den konsequenten Einsatz der Testgetriebenen Entwicklung im Schnitt deutlich besserer Code hinsichtlich der Komplexität/Umfangs produziert wird, und auch die Testung selbst stark vereinfacht wird, da Code und Test gemeinsam entstehen, und nicht im Nachhinein.

Die Testgetriebene Entwicklung ist allerdings kein Allheilmittel („Silver Bullet“) für alle Probleme der Software-Entwickler. Der Einsatz von TDD erfordert vom Programmierer ein hohes Maß an Eigenverantwortung und ein hohes Wissen über Software-Design, um Refaktorisierungen sinnvoll durchführen zu können. TDD kann auch nicht davor schützen, ein falsches Produkt zu entwickeln, da die Anforderungen falsch verstanden wurden. Auch benötigt TDD auf Kurze Sicht betrachtet mehr Zeit bei der Entwicklung, dafür kann es auf lange Sicht Zeit beim Debuggen sparen, da die Tests schnell Anhaltspunkte über Ursachen von Problemen geben. Die Effektivität einer Testgetriebenen Entwicklung ist auch maßgeblich von den vorhandenen Werkzeugen innerhalb der Programmiersprache und Framework abhängig. So verfügen die weit-verbreiteten Sprachen über ein Vielzahl von Werkzeugen, die den Prozess unterstützen, z.B. Code-Metriken, parallelisierbare Test-Runner oder hochentwickelte Test-Frameworks.

Nichtsdestotrotz ist TDD eine Technik, die es nach Ansicht des Autors wert ist, von jedem Programmierer gelernt zu werden. Bisherige Studien und Erfahrungen zeigen fast überwiegend positive Auswirkungen von TDD. Die Technik ist zwar schnell beschrieben („Red“ „Green“ „Refactor“), allerdings bedarf es viel Übung diesen Zyklus beizubehalten, und scheinbar schwierig zu testende Probleme anzugehen.

9.1. Ausblick

Wie erwähnt ist die Entwicklung der Anwendung noch nicht abgeschlossen, und wird zu gegebener Zeit vom Autor fortgesetzt. In diesem Zusammenhang werden dann öfter Akzeptanztest als treibende Kraft eingesetzt werden, und deren Auswirkungen untersucht werden.

Der hier gezeigte Ansatz eines Benchmarks der Code-Qualität innerhalb der Ruby/Rails-Community kann so ausgebaut werden, dass weitere Parameter untersucht und deutlich mehr Projekte einfließen, um so letztendlich repräsentative Kennzahlen zu erhalten, was eine gute Code-Qualität für Railsanwendungen ausmacht. Dazu wird in Zukunft ein Benchmarking-Tool entwickelt, welcher die Messungen halbautomatisch vornehmen wird. Leider ist die Situation bei den Metrik-Werkzeugen nicht optimal; viele Werkzeuge funktionieren nur in der älteren Ruby Version 1.8.7. Allerdings nutzen viele der aktuellen Anwendungen bereits Features, die nur die Ruby Version 1.9.x unterstützt. Hier wäre eine teilweise Neuentwicklung einiger Messwerkzeuge sinnvoll.

Interessant wäre die Effektivität von TDD in Abhängigkeit von der eigenen Erfahrung mit dieser Technik zu untersuchen, um daraus Schlußfolgerungen zu ziehen, wie lange ein Programmierer im Durchschnitt benötigt, um effektiv testgetrieben zu entwickeln. Ebenfalls wären Studien über die psychologische Auswirkungen von TDD interessant, da erfahrene TDD-Entwickler zwar gefühlsmäßig wissen, dass TDD einen Fluss erzeugt und richtig angewendet Spaß macht, aber dies nicht anhand von empirischen Zahlen belegen können.

Innerhalb der pludoni GmbH werden außerdem zukünftige Ruby Projekte mittels TDD durchgeführt um weitere Erfahrungen zu sammeln. Falls diese als Fallstudie für die Wissenschaft interessant sein sollten, so werden diese veröffentlicht.

A. Nutzung von Cucumber in Verbindung mit Selenium für Firefox und Guard ohne X-Server

Für ein effektives Test-Setup wurden folgende Testtools benutzt

Guard Automatische Testausführung bei Änderungen der Dateien

Xvfb Ist ein X-Server, der ohne Grafikausgabe arbeitet. Er eignet sich also für die Ausführung von GUI-Programmen, wie Mozilla Firefox auf Remote-Servern.

Selenium Interface zur Fernsteuerung von Browsern

Spork Erhöht die Testausführungsgeschwindigkeit, da Teile von Rails im Hintergrund gehalten werden, und zwischen den Tests nicht neu geladen werden müssen

Installation

Es muss bereits installiert sein: firefox, xvfb, Rails > 3.0

1. In das Gemfile folgendes eintragen:

Listing 43: RAILS_ROOT/Gemfile

```
group :test, :cucumber do
  gem "capybara"
  gem 'cucumber'
  gem "cucumber-rails"
  gem 'database_cleaner'
  gem "launchy"
  gem "guard"
  gem 'guard-cucumber'
  gem 'guard-spork'
  gem "rb-inotify" # Für Linux
  gem "spork", :git => "git://github.com/timcharper/spork.git"
end
```

Danach folgendes ausführen:

```
bundle install
rails g cucumber:install
```

2. Eine Datei Guardfile anlegen. Hier kommen alle Anweisungen zur Steuerung von guard hinein

Listing 44: RAILS_ROOT/Guardfile

```
group "cucumber" do
  guard 'spork' do
    watch('config/application.rb')
    watch('config/environment.rb')
    watch('features/support/env.rb')
    watch(%r{^config/environments/.+\.rb$})
    watch(%r{^config/initializers/.+\.rb$})
    watch('spec/spec_helper.rb')
  end
  guard 'cucumber', :cli => '--no-profile --color --format pretty --
    strict RAILS_ENV=test' do
    watch(%r{^features/.+\.feature$})
    watch(%r{^features/support/.+$}) { |m| Dir[
      features' }
    watch(%r{^features/step_definitions/(.+)_steps\.rb$}) { |m| Dir[
      File.join("**/#{m[1]}.feature")][0] || 'features' }
  end
end
```

3. In der Datei config/cucumber.yml die Option --drb für „default“ und „wip“ setzen.
4. Die Datei features/support/env.rb anpassen, um sie mit Spork kompatibel zu machen

Listing 45: features/support/env.rb

```
require 'cucumber/rails'
require "spork"
require 'test/unit/testresult'
Test::Unit.run = true

Spork.prefork do
  ENV["RAILS_ENV"] ||= "test"
  require File.expand_path(File.dirname(__FILE__) + '/../../config/
    environment')
  require 'cucumber/formatter/unicode'
  require 'cucumber/rails'
```

```
require 'test/unit/testresult'
ActionController::Base.allow_rescue = false
begin
  DatabaseCleaner.strategy = :truncation
rescue NameError
  raise "You need to add database_cleaner to your Gemfile (in the
    :test group) if you wish to use it."
end
end

Spork.each_run do
  require 'cucumber/rails/world'
end
```

Nutzung

1. Aktivierung von Xvfb auf Displayport 99

```
Xvfb :99 -ac
```

2. Aktuelle Shell auf diesen Displayport einstellen

```
export DISPLAY=:99
```

3. Guard starten

```
guard -g cucumber
```

Nun werden automatisch alle Cucumber-Features getestet. Falls Selenium verwendet wird, dann wird der Firefox im Hintergrund gestartet, ohne dass dies sichtbar wird.

B. Eigenschaften erfolgreicher Tests

Das Vorhandensein von zahlreichen Tests reicht nicht aus, um die Qualität der Tests zu beurteilen. Stattdessen existieren einige Qualitätskriterien, um Tests auf ihre Sinnhaftigkeit und Wartbarkeit zu untersuchen:

Unabhängigkeit (Independence) und Isolation Ein Test ist unabhängig, falls er nicht

durch andere Tests beeinflusst wird. Auch die Reihenfolge, in der die Tests ausgeführt werden, darf auf das Ergebnis keinen Einfluss üben. Siehe auch [Beck, 2002], [Palermo, 2006] sowie [Langr and Ottinger, 2011, Karte 45].

Wiederholbarkeit (Repeatability) Ein Test wird als wiederholbar bezeichnet, wenn er mehrmals hintereinander ausgeführt werden kann, und dabei jedes mal dasselbe Ergebnis liefert [Langr and Ottinger, 2011, Karte 45] [Rappin, 2011]. Problematisch sind dabei z.B. Datum und Zeit, sowie Zufallsfunktionen

Klarheit (Clarity) Ein Test ist klar, wenn sein Zweck sofort verständlich wird [Rappin, 2011], [Palermo, 2006]. Damit wird einerseits die Lesbarkeit gemeint. Andererseits schließt dies auch ein, ob der Test genau eine Eigenschaft testet und nicht redundant zu anderen Tests ist. Dies hat zur Folge, dass die Tests wartbarer werden und als Code Dokumentation dienen können.

Präzise (Conciseness) Ein Test ist präzise, wenn er so wenig Code und so wenige Objekte wie möglich benötigt, um sein Ziel zu erreichen [Palermo, 2006] [Rappin, 2011]. Eine Auswirkung ist, dass der Test schneller wird.

Robustheit (Robustness) Ein Test ist robust, wenn es eine direkte Korrelation zum zu testenden Code gibt: Ist der Code korrekt, so ist der Test erfolgreich. Ist der Code falsch, so schlägt der Test fehl. Nicht-robuste Tests werden auch „zerbrechlich“ (brittle) genannt. Dazu zählen auch sogenannte tautologische Tests, die immer erfolgreich Verlaufen, und keine Aussage über den zugrunde liegenden Programmcode geben

Schnelligkeit Gute Tests sollten schnell sein, da die Test-Suite ansonsten deutlich seltener ausgeführt wird [Langr and Ottinger, 2011] [Palermo, 2006] [Nagappan u. a., 2008]. Siehe auch Präzision.

Zeitliche Nähe Gute Tests entstehen in zeitlicher Nähe zu dem zugrunde liegenden Programmstück Langr and Ottinger [2011]. Innerhalb von TDD ist dieses Kriterium immer erfüllt.

Abbildungsverzeichnis

1. Anwendungsfälle	5
2. Einteilung der Tests	10
3. Red-Green-Refactor: Der TDD Entwicklungszyklus	17
4. Flussdiagramm für TDD	18
5. Entwicklungsablauf	19
6. MVC Modell von Rails	36
7. Attribute des Modells „Job“	55
8. Funktionsweise von save bei ActiveRecord Objekten	58
9. Aufbau eines Cucumber- Verzeichnisses	75
10. Ablauf beim Akzeptanztest mit Cucumber und Capybara	77
11. Entwicklung des Umfangs des Test- und Programmcodes	85
12. C0 Testabdeckung über den Verlauf der Entwicklung	86
13. Verlauf der verschiedenen Klassen an Code-Smells über den Entwick- lungszeitraum	87
14. Vergleich des Verhältnisses aus Anzahl Testcodezeilen / Anzahl Codezeilen	91
15. Vergleich der durchschnittlichen Komplexität	93
16. Vergleich der Anzahl Smells pro KLOC	94

Listings

1. Ruby Beispiele	26
2. Ruby Beispiel: Blöcke	27
3. Klassenhierarchien	27
4. Ruby Beispiel offene Klassen	28
5. Demonstration von generischen	28
6. Nutzung von Metaprogrammierung zur Erstellung von Objektbeziehungen	29
7. Testen mit Test::Unit in Ruby	31
8. Cucumber: Additionsfeature in Deutsch	33

9.	Cucumber: Implementierung der Additionstestschritte in Ruby	33
10.	Nutzung der Codegeneratoren von der Kommandozeile	38
11.	Abstraktion von SQL-Befehlen	40
12.	Nutzung von Validierungskriterien und DSL für Modellbeziehungen . . .	40
13.	Listing Test	57
14.	Test auf Vorhandensein eines Titels	57
15.	Implementierung der Validierung in die Klasse Job	59
16.	refaktorisierter Test	59
17.	Auslagerung des Regulären Ausdrucks in einen Initialisierer	60
18.	Alle bisherigen Testmethoden in der Klasse JobTest	61
19.	Fixtures Testdaten für zwei Jobs	61
20.	Finale Job-Test Klasse nach Refaktorisierung	62
21.	Feedimport Beispiel-XML Datei mit einem Job	67
22.	Feed Test I.	68
23.	Feed Implementation I.	69
24.	Feed Test II	70
25.	Feed Implementation II	70
26.	Zentrale Implementierung des Mocks in der Test Helper	71
27.	Feed Test IIb nach Refaktorisierung	71
28.	Feed Test III	71
29.	Feed Implementation III - Fake Implementierung	72
30.	Feed Test IV – Test auf nicht-erfolgreichen HTTP-Status Code	72
31.	Feed Implementation IV - Implementation der Validierung	72
32.	Feed Test V - Testen der Robustheit gegen Exceptions	73
33.	Feed Implementation V	73
34.	Feed Implementation Vb - nach Refaktorisierung	74
35.	Definition eines Szenarios in einem Cucumber-Feature	76
36.	Implementierung der Testschritte zur Interaktion mit einem Browser . .	78
37.	Testschritt zum Anlegen von Testdaten auf Basis der Cucumber Definition	79
38.	Implementation eines Suchfeld im Layout der Webseite	79
39.	JobsController am Ende der Testphase	80
40.	View-Implementation für eine Liste von Jobs	80
41.	Auslagerung der Suchlogik in die Job-Klasse	81
42.	Finaler Jobcontroller nach Refaktorisierung	81
43.	RAILS_ROOT/Gemfile	98
44.	RAILS_ROOT/Guardfile	99

45. features/support/env.rb	99
---------------------------------------	----

Literaturverzeichnis

- [Ambler 2002] AMBLER, Scott: *Introduction to Test Driven Design (TDD)*. Letzter Zugriff: 2011-08-28 08:07:09. 2002. – URL <http://www.agiledata.org/essays/tdd.html>
- [Andrew Hunt 1999] ANDREW HUNT, David T.: *The Pragmatic Programmer: From Journeyman to Master*. 1. Addison-Wesley Professional, Oktober 1999. – ISBN 020161622X
- [Baggen u. a. 2011] BAGGEN, Robert ; CORREIA, José P. ; SCHILL, Katrin ; VISSER, Joost: Standardized code quality benchmarking for improving software maintainability. In: *Software Quality Journal* (2011), Mai. – URL <http://www.springerlink.com/content/vm77272104085276/>. – ISSN 0963-9314
- [Beck 2002] BECK, Kent: *Test Driven Development. By Example*. Addison-Wesley Longman, Amsterdam, November 2002. – ISBN 0321146530
- [Brown 2008] BROWN, Roger: *Test Driven Development and Flow*. Web 2011-08-07 18:31:57. April 2008. – URL <http://www.agilecoachjournal.com/post/Test-Driven-Development.aspx>
- [Cangiano 2010] CANGIANO, Antonio: *The Great Ruby Shootout (July 2010)*. Letzter Zugriff: 09.09.2011. Juli 2010. – URL <http://programmingzen.com/2010/07/19/the-great-ruby-shootout-july-2010/>
- [Caroli 2008] CAROLI, Paulo: *Agile Tips: Testing private methods, TDD and Test-Driven Refactoring*. Letzter Zugriff: 09.09.2011. November 2008. – URL <http://agiletips.blogspot.com/2008/11/testing-private-methods-tdd-and-test.html>
- [Cornett 1996] CORNETT, Steve: *Code Coverage Analysis*. Letzter Zugriff: 2011-08-07 20:53:34. 1996. – URL <http://www.bullseye.com/coverage.html>
- [Dave Thomas 2009a] DAVE THOMAS, Chad Fowler Andrew H.: *Programming Ruby 1.9: The Pragmatic Programmers' Guide*. 3rd. Pragmatic Bookshelf, April 2009. – ISBN 9781934356081
- [Dave Thomas 2009b] DAVE THOMAS, David Heinemeier H.: *Agile Web Development with Rails, Third Edition*. Third Edition. Pragmatic Bookshelf, April 2009. – ISBN 1934356166

- [DeVries and Naberezny 2008] DEVRIES, Derek ; NABEREZNY, Mike: *Rails for PHP Developers*. 1. Pragmatic Bookshelf, Februar 2008. – ISBN 1934356042
- [Elssamadisy 2007] ELSSAMADISY, Amr: *InfoQ: 100% Test Coverage?* Letzter Zugriff: 2011-06-19 10:08:32. Mai 2007. – URL http://www.infoq.com/news/2007/05/100_test_coverage
- [Ford 2010a] FORD, Neal: *Evolutionary architecture and emergent design: Leveraging reusable code, Part 1*. Letzter Zugriff: 09.09.2011. März 2010. – URL <http://www.ibm.com/developerworks/java/library/j-eaed11/index.html>
- [Ford 2010b] FORD, Neal: *Vortrag: Emergent Design And Evolutionary Architecture*. Letzter Zugriff: 09.09.2011. März 2010. – URL <http://www.thoughtworks.com/emergent-design>
- [Game 2011] GAME, Computer Language B.: *Ruby 1.9 Speed - C/GNU GCC speed*. <http://shootout.alioth.debian.org/u32q/benchmark.php?p=all&lang=yarv&lang2=gcc>. Juni 2011. – URL <http://shootout.alioth.debian.org/u32q/benchmark.php?test=all&lang=yarv&lang2=gcc>
- [Goodliffe 2006] GOODLIFFE, Pete: *Code Craft: The Practice of Writing Excellent Code*. 1st ed. No Starch Press, August 2006. – ISBN 1593271190
- [Hansson 2011] HANSSON, David H.: *RailsConf 2011 Keynote*. Mai 2011. – URL <http://www.rubyinside.com/dhh-keynote-streaming-live-from-railsconf-2011-right-here-right-now-4769.html>
- [Herrmann 2005] HERRMANN, Dr. C.: *Metaprogrammierung*. In: *Funktionale Programmierung WS2005/2006* (2005). – URL <http://www.infosun.fim.uni-passau.de/cl/lehre/sips1-ss06/Uebung/Metaprogrammierung.pdf>
- [HQ 2010] HQ, Selenium: *Selenium Contributors*. Letzter Zugriff: 09.09.2011. 2010. – URL <http://seleniumhq.org/about/contributors.html>
- [Hulkko and Abrahamsson 2005] HULKKO, Hanna ; ABRAHAMSSON, Pekka: A multiple case study on the impact of pair programming on product quality. In: *Proceedings of the 27th international conference on Software engineering - ICSE '05*. St. Louis, MO, USA, 2005, S. 495. – URL <http://dl.acm.org/citation.cfm?id=1062545>

- [IEEE 1998] IEEE: IEEE 1061-1998 Standard for a Software Quality Metrics Methodology. In: *IEEE Computer Society* (1998), Nr. 31 Dec. 1998
- [Inc. 2007] INC., Stelligent: *Stelligent Survey Reveals Majority of Organizations Do Not Practice Test-Driven Development*. Letzter Zugriff: 09.09.2011. November 2007. – URL http://www.businesswire.com/portal/site/google/index.jsp?ndmViewId=news_view&newsId=20071127006009&newsLang=en
- [Janzen and Saiedian 2008] JANZEN, David ; SAIEDIAN, Hossein: Does Test-Driven Development Really Improve Software Design Quality? In: *IEEE Software* 25 (2008), Nr. 2, S. 77–84. – ISSN 0740-7459
- [Klukas 2011a] KLUKAS, Jörg: *Referenzen | pludoni GmbH - the community experts*. <http://www.pludoni.de/referenzen>. Juni 2011. – URL <http://www.pludoni.de/referenzen>
- [Klukas 2011b] KLUKAS, Jörg: *Startseite - Aktuelles*. <http://www.itmitte.de/>. Juni 2011. – URL <http://www.itmitte.de/>
- [Langr and Ottinger 2011] LANGR, Jeff ; OTTINGER, Tim: *Agile in a Flash: Speed-Learning Agile Software Development*. Crds. Pragmatic Bookshelf, Februar 2011. – ISBN 1934356719
- [Liggesmeyer 1990] LIGGESMEYER, Peter: *Modultest und Modulverifikation. State of the Art*. BI-Wiss.-Verl., 1990 (Angewandte Informatik 4). – ISBN 3411143614
- [Madeyski 2009] MADEYSKI, Lech: *Test-Driven Development: An Empirical Evaluation of Agile Practice*. 1st Edition. Springer, Dezember 2009. – ISBN 9783642042874
- [McCabe 1976] MCCABE: A Complexity Measure. In: *IEEE Transactions on Software Engineering* (1976), Dezember, S. 308–320. – URL <http://www.literateprogramming.com/mccabe.pdf>
- [Meijer and Drayton 2005] MEIJER, E. ; DRAYTON, P.: Static typing where possible, dynamic typing when needed: The end of the cold war between programming languages, 2005
- [Müller 2006] MÜLLER, Matthias M.: The Effect of Test-Driven Development on Program Code. In: *PROC. INT'L CONF. EXTREME PROGRAMMING AND AGILE PROCESSES IN SOFTWARE ENG. (XP 06)* (2006), S. 94–103. – URL <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.65.8244>

- [Nagappan u. a. 2008] NAGAPPAN, Nachiappan ; MAXIMILIEN, E ; BHAT, Thirumalesh ; WILLIAMS, Laurie: Realizing quality improvement through test driven development: results and experiences of four industrial teams. In: *Empirical Software Engineering* 13 (2008), Juni, Nr. 3, S. 289–302. – URL <http://dx.doi.org/10.1007/s10664-008-9062-z>. – ISSN 1382-3256
- [Orsini 2007] ORSINI, Rob: *Rails Cookbook (Cookbooks)*. 1st Ed. O'Reilly Media, Januar 2007. – ISBN 0596527314
- [Ousterhout 1998] OUSTERHOUT, J. K.: Scripting: Higher level programming for the 21st century. In: *Computer* 31 (1998), Nr. 3, S. 23–30
- [Palermo 2006] PALERMO, Jeffrey: *Guidelines for Test-Driven Development*. Letzter Zugriff: 22.08.2011. Mai 2006. – URL [http://msdn.microsoft.com/en-us/library/aa730844\(v=vs.80\).aspx](http://msdn.microsoft.com/en-us/library/aa730844(v=vs.80).aspx)
- [Powell 2008] POWELL, Catherine: *Abakas: Code Coverage Complexity*. Letzter Zugriff: 2011-08-07 20:59:33. April 2008. – URL <http://blog.abakas.com/2008/04/code-coverage-complexity.html>
- [on Rails.org 2011] RAILS.ORG, Ruby on: *Ruby on Rails: Applications*. Letzter Zugriff: 09.09.2011. 2011. – URL <http://rubyonrails.org/applications>
- [Rappin 2011] RAPPIN, Noel: *Rails Test Prescriptions*. 1. Pragmatic Bookshelf, März 2011. – ISBN 9781934356647
- [Reeves 1992] REEVES, Jack W.: Three Essays by Jack W. Reeves - I. What Is Software Design? In: *C++ Journal* (1992), Nr. Herbst 1992. – URL http://www.developerdotstar.com/mag/articles/reeves_design.html
- [Rutherford 2010] RUTHERFORD, Kevin: *Code Smells - Reek*. Letzter Zugriff: 09.08.2011. 2010. – URL <https://github.com/kevinrutherford/reek/wiki/Code-Smells>
- [Sadists 2010] SADISTS, Ruby: *Confessions of a Ruby Sadist sudo gem install heckle*. Letzter Zugriff: 2011-08-07 21:15:44. 2010. – URL <http://ruby.sadi.st/Heckle.html>
- [Savoia 2007] SAVOIA, Alberto: *The Code C.R.A.P. Metric Hits the Fan - Introducing the crap4j Plug-in*. Letzter Zugriff: 2011-08-07 21:56:27. Oktober 2007. – URL <http://www.artima.com/weblogs/viewpost.jsp?thread=215899>

- [Shore 2007] SHORE, James: *The Art of Agile Development*. 1. O'Reilly Media, November 2007. – ISBN 0596527675
- [Spiewak and Harrop 2010] SPIEWAK, Daniel ; HARROP, Jon: *Dynamic type languages versus static type languages - Stack Overflow*. <http://stackoverflow.com/questions/125367/dynamic-type-languages-versus-static-type-languages>. 2010. – URL <http://stackoverflow.com/questions/125367/dynamic-type-languages-versus-static-type-languages>
- [Stackoverflow] STACKOVERFLOW, Jon Skeet Ilja P.: *testing - Is there a difference between TDD and Test First Development (or Test First Programming)? - Stack Overflow*. Letzter Zugriff: 09.09.2011. – URL <http://stackoverflow.com/questions/334779/is-there-a-difference-between-tdd-and-test-first-development-or-test-first-progr>
- [Stephens and Rosenberg 2010] STEPHENS, Matt ; ROSENBERG, Doug: *Design Driven Testing: Test Smarter, Not Harder*. Apress, November 2010. – ISBN 9781430229438
- [Team 2011] TEAM, Ruby Visual I.: *About Ruby*. <http://www.ruby-lang.org/en/about/>. Juni 2011. – URL <http://www.ruby-lang.org/en/about/>
- [Vanderburg 2010] VANDERBURG, Glenn: *Vortrag: Real Software Engineering*. Letzter Zugriff: 09.09.2011. August 2010. – URL <http://confreaks.net/videos/282-lsrc2010-real-software-engineering>
- [Wasmus and Gross 2007] WASMUS, Hans ; GROSS, Hans: Evaluation of Test Driven Development. In: *Technical Report Series* (2007), Nr. TUD-SERG-2007-014. – URL <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.100.6964>