

# Bachelor-Praxisphase im Projekt „PARASUITE“ bei Cognidata GmbH

Michael Eckel

9. Juni 2009

## Allgemeines

### Zeitraum der Praxisphase

Vom 16.02.2009 bis zum 15.05.2009 (Sommersemester 2009).

### Betreuende Personen

- Herr Prof. Dr. Thomas Letschert
- Herr M. Sc. Sebastian Süß

### Über das Projekt „PARASUITE“

„PARASUITE“ bedeutet **P**roduct **A**nalysis and **R**eporting **A**pplication **S**UITE und ist ein Entscheidungsunterstützendes System im Bereich Produktlebenszyklus. Kunden sind u. a. Bombardier Transportation und Thyssen Elevators. Bombardier wird im Bereich *Beginning of Life* des Produktlebenszyklus durch PARASUITE unterstützt mit dem Ziel bessere Produkte zu bauen. Thyssen wird im Bereich *Middle of Life* unterstützt und benötigt vorausschauende Produktwartung (engl. *Predictive Maintenance*).

PARASUITE wendet zur Analyse der vom Kunden gesammelten Daten *Data Mining*-Methoden an, um Erkenntnisse aus den Daten zu ziehen. Somit können im Falle von Bombardier Fehlerquellen ausfindig gemacht werden und im Falle von Thyssen anstehende Wartungsarbeiten „vorhergesagt“ werden.

## Eingesetzte Technologie(n)

Die Programmiersprache im Projekt ist Java. Benutzt wird die Java Enterprise Edition in Version 5. Als Application Server kommt JBoss zum Einsatz und als Datenbanksystem MySQL. Als Entwicklungsumgebung wird die Eclipse IDE eingesetzt und als GUI-Framework Eclipse RCP.

## Meine Tätigkeiten

Zur Analyse der vom Kunden gesammelten riesigen Datenmengen ist es nötig Vorberechnungen zu machen, um einen wesentlich kleineren Satz von Daten zu erhalten, welcher dann vom Reporting-Plugin „BIRT“ für das Eclipse RCP-Framework weiter ausgewertet und grafisch dargestellt werden kann.

Eine frühere Lösung sah so aus, dass die Berechnungen in SQL hart-kodiert in den Java-Quelltext geschrieben wurden. Das Ganze wurde dann umgestellt auf *Flow-Based Programming* (FBP), welches ein Programmierparadigma ist, in dem man miteinander verbundene Rechen-Komponenten als ein Berechnungsnetzwerk ansieht. Ein Berechnungsnetz besteht aus Komponenten (Knoten), Verbindungen und Sub-Netzen.

Meine Aufgabe während der Praxisphase war Die Entwicklung eines Datenbank-Leseknotens für das neue Datenbankschema. Nach der Einarbeitung in FBP und der Einarbeitung ins FBP-Framework für Java entwickelten sich diverse andere „Probleme“. Die meiste Zeit verwendete ich auf die Konfiguration der Komponente. Dort galt es den „Filter“-Konfigurationsparameter, welcher Client-seitig ein Objektbaum war, textuell auf die Server-Seite zu bekommen in menschenlesbarer Form. Eine Möglichkeit, diesen Objektbaum in SQL zu wandeln bestand bereits, sodass mein erster Versuch war das generierte SQL zu parsen und somit den ursprünglichen Objektbaum wieder herzustellen. Es stellte sich aber heraus, dass im generierten SQL-Code einfach nicht genügend Informationen waren. Auch das „künstliche“ hinzufügen dieser fehlenden Informationen durch Kommentare endete darin, dass der Code sehr unleserlich wurde.

Der zweite Versuch war eine eigene Sprache einzuführen, die nicht mit zu viel Overhead belastet ist und somit der Lesbarkeit entgegenkommt. Ich nannte die Sprache PFEL (**PARASUITE Filter Expression Language**). Dieser Teil der Praxisphase war der zeitintensivste, denn ich musste mich erst in die Techniken des Compilerbaus einarbeiten. Am Ende hatte ich dann einen funktionierenden Parser für diese Sprache und einen Printer, der den Objektbaum in PFEL wandeln kann.

Weitere, vom Umfang her kleinere Aufgaben, wie z. B. das Bereitstellen eines Datenzugriffsobjekts (DAO) für die Datenbank, gehörten auch zu meinem Aufgabenbereich. JUnit-Testing wurde auch stets betrieben und sollte für (fast) jede erstellte Klasse implementiert werden.