

# Abstract

## 1. Deutsch

Die vorliegende Masterarbeit beschreibt die Softwarestruktur-Analyse - Reverse Engineering - der *SnoWMA*n-Web-Plattform des *Instituts für Genomik und Bioinformatik* der *TU Graz* zur Untersuchung von Mikrobiom-Daten aus Hochdurchsatz Sequenzierverfahren. Die extrahierten Architektur-Muster dienten als Wissensbasis für die Erweiterungen im zweiten Teil der Arbeit. Sie wurden analysiert und führten zum Refactoring von Software-Modulen, die für die gewünschten Erweiterungen notwendig waren. Resultierende Verbesserungsvorschläge wurden herausgearbeitet und festgehalten.

Nach Abschluss des Reverse Engineerings wurde die Plattform um die *Chimerafiltering*-Funktion erweitert, bei der geplanten *Denoising*-Erweiterung erwies sich die verfügbare Rechenleistung des Rechenclusters als Hürde.

Die positiven Auswirkungen des eingeführten *Chimerafilterings* wurden durch eigene Analysen gezeigt. Bei als eindeutig erkannten Sequenzen zeigte sich eine Reduzierung von zumindest -3,10%, die höchste Reduzierung der OTU-Anzahl trat in der *UCLUST*-Pipeline beim Ähnlichkeitsmaß 95% mit -15,15% auf.

Abschließend werden Verbesserungsvorschläge bei zukünftigen Arbeiten an *SnoWMA*n diskutiert.

**Stichwörter:** *SnoWMA*n, Reverse Engineering, Refactoring, *Denoising*-Erweiterung, *Chimerafiltering*-Erweiterung

## 2. English

The present master thesis describes the reverse engineering of the *SnoWMAAn*-Web-application of the *Institute of Genomics and Bioinformatics* at the *University of Technology* in Graz which is for analysis of large microbiome sequencing data. The extracted software-patterns provided the necessary knowledge to extend the platform in the second part of the work. They were analyzed and software modules which had impeded the planned extensions were refactored. Resulting improvement suggestions have been worked out and have been collected.

After the reverse engineering had been done the platform was extended with *Chimerafiltering*, for the intended *Denoising*-extension the provided computing power of the used computer cluster turned out to be an obstacle. The positive effects of the introduced *Chimerafiltering* are shown by means of own analysis. The reduction of unique sequences was -3.10% or higher, the highest reduction of OTUs occurred in the *UCLUST*-pipeline which -15.15% at 95% similarity.

In conclusion some improvement suggestions for future works on *SnoWMAAn* are discussed.

**Keywords:** *SnoWMAAn*, reverse engineering, refactoring, denoising-extension, chimerafiltering-extension