Popularität von Java-Annotationen im zeitlichen Verlauf

Fortgeschrittene Softwaretechnik Wintersemester 2017/2018

Michael Wolz

12. März 2018

1 Reflexion

Die in der Vorlesung behandelte zeitliche Analyse der Entwicklung von JUnit-Testmethoden mithilfe von BOA weckte bereits großes Interesse bei mir. Gerade die Entwicklung von verschiedenen Veränderungen an Quellcodes über die Zeit finde ich sehr interessant und somit bot es sich an, dass ich mich im Rahmen des Abschlussprojektes für das Modul Fortgeschrittene Softwaretechnik mit der Popularität von Java-Annotationen im zeitlichen Verlauf beschäftige. Hierbei werden zunächst die relativen Häufigkeiten von Java-Annotationen betrachtet und anschließend die meist verwendeten Annotationen im zeitlichen Verlauf analysiert. Die für die Analyse herangezogenen Daten stammen dabei aus dem GitHub September 2015 Datensatz von BOA, welcher insgesamt 380125 Repositories beinhaltet [1]. Die Daten werden im Anschluss mit R verarbeitet und ausgewertet.

2 Datenbeschaffung mit BOA

Im Ordner boa befindet sich das Skript, welches zur Datenbeschaffung auf dem BOA-Datensatz verwendet wurde. Für jedes Repository wird die erste (startYear), sowie die letzte Revision (endYear) als zeitliche Referenz verwendet. Zwischen diesen beiden Werten werden im Anschluss quartalsweise Snapshots betrachtet und die verwendeten Annotationen in einem mehrdimensionalem Array mit dem Aufbau Annotation: Jahr: Quartal: Count gespeichert. Zusätzlich wird gezählt, wie viele Repositories eine bestimmte Annoation beinhalten, sodass beispielsweise firmen- oder projektinterne Annoationen in der späteren Analyse nicht mit einbezogen werden müssen. Des Weiteren werden noch die insgesamt betrachteten Revisionen pro Quartal gezählt, wobei diese nur gezählt werden, falls sie Annotationen beinhalten.

3 Datenauswertung

Betrachteter Zeitraum: 2005/Q1 - 2013/Q4

4 Fazit

Literatur

[1] Iowa State University. BOA - Dataset Statistics. http://boa.cs.iastate.edu/stats/index.php. Zugriff: 12.03.2018.