DHBW Stuttgart
Studiengang Informatik

Anmeldung Thema: Bachelorarbeit

Name: **Oliver Rudzinski** **STG-TINF17A**
Mail: inf17068@lehre.dhbw-stuttgart.de
Zeit der Anmeldung: 09.03.2020, 09:53 Uhr

Thema: Setup and Implementation of an Automated Testing Pipeline for a DataOps Use Case
Kurzbeschreibung: DataOps is a paradigm for building and operating agile data analytics solutions. Examples for data analytics solutions in scope are modern artificial intelligence (AI) and machine learning (ML) microservices as well as traditional data ingestion pipelines, data marts, and business intelligence (BI) reports. In DataOps, learnings from DevOps in software development are adopted and transferred to building data analytics solutions. A crucial discipline within data analytics is data and model testing. As part of the Work-integrated Project III, a provisional DataOps pipeline for building retail business data analytics solutions has been designed and developed within the Analytics department of DXC Technology Company. This bachelor's thesis ties in with this outcome by analyzing and correcting lacking functionality as described in the previous project's documentation. In order to implement a suitable testing environment, established functional as well as non-functional testing paradigms (e.g., unit tests, integration tests, smoke tests, etc.) are surveyed and evaluated. The automation of the testing process is enabled by making use of adequate testing frameworks. Subsequently, these methodologies are implemented in the enhanced DataOps pipeline. The different stages of the pipeline are realized with several Amazon Web Services (AWS) products and coordinated through a superordinate continuous integration and continuous delivery (CI/CD) pipeline as well as Infrastructure-as-Code (IaC) methodologies. The bachelor's thesis' project yields a holistic DataOps pipeline for the agile development and maintenance of data analytics solutions for retail businesses. The performance of the resulting pipeline is exemplarily demonstrated with the help of a market basket analysis (MBA) scenario.
Vertaulichkeit: nicht vertraulich
Ort: Böblingen
Betreuer: Bernd Gloss (Tel: +49 173 7242670, Mail: bgloss@dxc.com)