Software Engineering for Data Intensive Sciences

Course Description

Building a successful data-based product requires a significant amount of high-quality code, which needs to run in a professional production environment. This course starts by introducing the agile approaches Scrum and Kanban and then discusses the shift from more traditional software development approaches to the DevOps culture. Special focus is given to the discussion and understanding of techniques and approaches for producing high-quality code such as unit and integration testing, test-driven development, pair programing, and continuous delivery and integration. Since many software artifacts are accessed via APIs, this course introduces the concepts of API design and paradigms. Finally, the course addresses the challenges of bringing code into a production environment, building a scalable environment, and using cloud-based approaches.

Contents

1. Agile Project Management
   1. Introduction to SCRUM
   2. Introduction to Kanban
2. DevOps
   1. Traditional Lifecycle Management
   2. Bringing Development and Operations Together
   3. Impact of Team Structure
   4. Building a DevOps Infrastructure
3. Software Development
   1. Unit and Integration Testing, and Performance Monitoring
   2. Test-Driven Development and Pair Programming
   3. Continuous Delivery and Integration
   4. Overview of Relevant Tools
4. API
   1. API Design
   2. API Paradigms
5. From Model to Production
   1. Building a Scalable Environment
6. Model Versioning and Persistence
7. Cloud-Based Approaches