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 artefacts are accessed via APIs, this course introduces concepts of API design and paradigms.Finally, this course addresses the challenges of bringing code into a production environment, building a scalable environment, and using cloud-cased 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 & integration test, performance monitoring
   2. Test-driven development & pair programing
   3. Continuous delivery & integration
   4. Overview of relevant tools
4. API
   1. API design
   2. API paradigms
5. From Model to Production

5.1 Building a scalable environment

1. Model versioning and persistence
2. Cloud-based approaches