CMD-EL3- Big Data Computing

Person in charge and Representative	Höppper
Contact person	Prof. Dr. Frank Höppner
	email: f.hoeppner@ostfalia.de
	Ostfalia University of Applied Sciences
	Institute of Information Engineering
Semester	Semester 3 (?)
Topic cluster	Computational Method Development
Duration/Credit	12 weeks, 3h per week
	which includes "lab times" where prepared assignments are discussed
Time	- course is held roughly once every three semesters
	- lecture 1x/week
	(course is already part of another curriculum)
Place	lecture takes place at Ostfalia University, Dept of Computer Science, Wolfenbüttel (The lecture is not broadcasted in general, only during the Covid pandemics.)
Prerequisite for the lecture/course	Programming (Java), algorithms and data structures
	programming in Java
	 algorithms and data structures
	(- this course is part of another curriculum, where it is usually held in german)
Aim of the lecture/course	The course aims at developing algorithmic solutions for big data that are scalable due to distributed, parallel execution in a cluster. Objectives:
	 understand problems and challenges of large data sets understand parallel architecture of Hadoop (e.g. HDFS) understand the programming model of MapReduce and Spark transfer algorithmic solutions for small data to solutions for big data techniques to achieve efficiency and scalability of solutions for big data use best practices / design patterns for own scalable algorithmic solutions coping with skew probabilistic algorithms (e.g. LSH) selected topics (Spark SQL, Mllib, GraphX, search engine) practical exercises in Java (part of exams)