

Modulname	Digitization and Data Analytics
Modulnummer	INF-25-Ma-FTI-DDA
Verantwortliche Dozentin bzw. verantwortlicher Dozent	Prof. Dr. Wolfgang Nagel wolfgang.nagel@tu-dresden.de
Qualifikationsziele	Die Studierenden verstehen verschiedene Konzepte zur Datenspeicherung und Metadatenerhebung und können Datenmanagementkonzepte für existierende Datensätze erarbeiten. Sie kennen und verstehen verschiedene Methoden der Datenvorprozessierung, der explorativen Datenanalyse und des maschinellen Lernens, können deren Stärken und Schwächen in Bezug auf gegebene Daten einschätzen und die Methoden mit Hilfe von aktuellen Software-Frameworks sowohl lokal als auch auf einer HPC-Infrastruktur anwenden.
Inhalte	Die Inhalte des Moduls sind aktuelle Ansätze bei der Analyse großer Datenmengen (Big Data) für spezifische Rechnerarchitekturen sowie die Prozessierung im Umfeld des High-Performance Computings (HPC), verschiedene Data-Analytics-Ansätze sowie deren Einsatz im HPC-Umfeld. Weitere Inhalte sind Data-Analytics-Frameworks, Grundlagen der statistischen Analyse und Verfahren aus dem maschinellen Lernen sowie ausgewählte Anwendungen.
Lehr- und Lernformen	Das Modul umfasst Vorlesungen im Umfang von 2 SWS, Übungen im Umfang von 2 SWS und das Selbststudium. Die Lehrsprache der Vorlesungen und der Übungen ist Englisch.
Voraussetzungen für die Teilnahme	Im Diplomstudiengang Informatik werden die in den Modulen INF-25-Ba-Ma1 Lineare Algebra und Analysis sowie INF-25-Ba-Ma4 Wahrscheinlichkeitstheorie und Statistik zu erwerbenden Kompetenzen vorausgesetzt. Im Masterstudiengang Computer Science werden Kenntnisse in Statistik und linearer Algebra sowie Kenntnisse in der Nutzung von Python auf Bachelorniveau vorausgesetzt. Zur Vorbereitung eignet sich folgende Literatur: 1. Mathematics for Engineers II: Calculus and Linear Algebra, G. Baumann. Oldenbourg Verlag, 2010. http://dx.doi.org/10.1524/9783486598476 , 2. All of Statistics, L. Wasserman. Springer. 2004. https://link.springer.com/book/10.1007/978-0-387-21736-9 , 3. Python for Data Analysis 3rd Ed., W. McKinney. O'Reilly Media, 2022. https://learning.oreilly.com/library/view/-/9781098104023/?ar .
Verwendbarkeit	Das Modul ist im Diplomstudiengang Informatik im Hauptstudium ein Wahlpflichtmodul im Fachgebiet Computer Engineering and High Performance Computing, das nach Maßgabe der Anlage 2 zur Prüfungsordnung zu wählen ist. Das Modul ist im Masterstudiengang Computer Science jeweils ein Wahlpflichtmodul im Open Track im Fachgebiet Computer Engineering and High Performance Computing

	sowie der Ergänzung, das nach Maßgabe der Anlage 2 zur Prüfungsordnung zu wählen ist. Das Modul kann im Masterstudiengang Computer Science nur einmal gewählt werden. Das Modul kann im Masterstudiengang Computer Science nicht gewählt werden, wenn dieses oder ein wesentlich inhaltsgleiches Modul aus einem Studiengang mit dem die Zugangsvoraussetzungen nach § 3 der Studienordnung erfüllt wurden, bereits absolviert wurde. Das Modul schafft die Voraussetzungen für die Module, die es unter Voraussetzungen für die Teilnahme benennen.
Voraussetzungen für die Vergabe von Leistungspunkten	Die Leistungspunkte werden erworben, wenn die Modulprüfung bestanden ist. Die Modulprüfung besteht aus einer Klausurarbeit von 90 Minuten Dauer. Die Prüfungssprache ist Englisch.
Leistungspunkte und Noten	Durch das Modul können 6 Leistungspunkte erworben werden. Die Modulnote entspricht der Note der Prüfungsleistung.
Häufigkeit des Moduls	Das Modul wird jedes Sommersemester angeboten.
Arbeitsaufwand	Der Arbeitsaufwand beträgt insgesamt 180 Stunden.
Dauer des Moduls	Das Modul umfasst ein Semester.