

## Master of Science Data Science

Akademischer Grad	Modulnummer	Modulform
Master of Science	07-202-1103A	Wahlpflicht

### Modultitel **Advanced Statistics**

**Modultitel (englisch)** Advanced Statistics

**Empfohlen für:** 1./3. Semester

**Verantwortlich** Professur für Statistik

**Dauer** 1 Semester

**Modulturnus** jedes Wintersemester

**Lehrformen**

- Vorlesung "Advanced Statistics" (2 SWS) = 30 h Präsenzzeit und 70 h Selbststudium = 100 h
- Übung "Advanced Statistics" (1 SWS) = 15 h Präsenzzeit und 35 h Selbststudium = 50 h

**Arbeitsaufwand** 5 LP = 150 Arbeitsstunden (Workload)

**Verwendbarkeit** • M.Sc. Data Science

**Ziele**

Die Studierenden können die theoretischen Grundlagen - d.h. mehrdimensionale Wahrscheinlichkeitsverteilungen, Inferenzkonzepte und deren Anwendung auf Regressionsmodelle - sowie den aktuellen Stand der Forschung im Bereich der Wahrscheinlichkeitsrechnung und induktiven Statistik darstellen und beschreiben. Sie sind in der Lage, selbständig geeignete, dem aktuellen Forschungsstand entsprechende Methoden und Strategien zur Lösung statistischer Problemstellungen auszuwählen, anzuwenden sowie die ermittelten Ergebnisse zu interpretieren und kritisch zu beurteilen.

Die Studierenden können in Übungen die Charakteristika komplexer statistischer Problemstellungen erkennen, passende Modelle zu deren Modellierung auswählen und mittels geeigneter Schätzmethoden anpassen. Hierbei sind sie in der Lage, das Softwarepakete R als Werkzeug für statistische Analysen einzusetzen, die Programmierung einfacher Routinen durchzuführen und ihre erzielten Ergebnisse kritisch auszuwerten.

**Inhalt** Mehrdimensionale Wahrscheinlichkeitsverteilungen, Inferenzkonzepte, Maximum-Likelihood-Schätzung und Likelihood-basierte Testverfahren, Generalisierte Lineare Regressionsmodelle

**Teilnahmevoraussetzungen** keine

**Literaturangabe** Hinweise zu Literaturangaben erfolgen in den Lehrveranstaltungen.

**Vergabe von Leistungspunkten** Leistungspunkte werden mit erfolgreichem Abschluss des Moduls vergeben. Näheres regelt die Prüfungsordnung.

**Prüfungsleistungen und -vorleistungen**

<b>Modulprüfung: Klausur 80 Min., mit Wichtung: 1</b>	
	Vorlesung "Advanced Statistics" (2SWS)
	Übung "Advanced Statistics" (1SWS)

## Master of Science Data Science

Akademischer Grad	Modulnummer	Modulform
Master of Science	09-INF-BI01	Wahlpflicht

<b>Modultitel</b>	<b>Statistisches Lernen</b> Vertiefungsmodul
<b>Modultitel (englisch)</b>	Statistical Learning In-Depth Module
<b>Empfohlen für:</b>	1./3. Semester
<b>Verantwortlich</b>	Institut für Medizinische Informatik, Statistik und Epidemiologie (IMISE)
<b>Dauer</b>	1 Semester
<b>Modulturnus</b>	jedes Wintersemester
<b>Lehrformen</b>	<ul style="list-style-type: none"> <li>• Vorlesung "Grundlagen des statistischen Lernens" (3 SWS) = 45 h Präsenzzeit und 105 h Selbststudium = 150 h</li> <li>• Übung "Grundlagen des statistischen Lernens" (1 SWS) = 15 h Präsenzzeit und 35 h Selbststudium = 50 h</li> <li>• Praktikum "Statistische Analysen mit R" (2 SWS) = 30 h Präsenzzeit und 70 h Selbststudium = 100 h</li> </ul>
<b>Arbeitsaufwand</b>	10 LP = 300 Arbeitsstunden (Workload)
<b>Verwendbarkeit</b>	<ul style="list-style-type: none"> <li>• M.Sc. Bioinformatik</li> <li>• M.Sc. Data Science</li> <li>• M.Sc. Digital Humanities</li> <li>• M.Sc. Informatik</li> </ul>
<b>Ziele</b>	<p>Nach der aktiven Teilnahme am Modul "Statistisches Lernen" sind die Studierenden in der Lage:</p> <ul style="list-style-type: none"> <li>- grundlegende Verfahren der Statistik korrekt anzuwenden,</li> <li>- verschiedene Verfahren des Maschinellen Lernens zu erklären, zu vergleichen, und zu komplexen Workflows zu verbinden und</li> <li>- Workflows der bioinformatischen Datenanalyse in der Statistiksprache R abzubilden und zu implementieren.</li> </ul>
<b>Inhalt</b>	<p>Vorlesung und Übung "Grundlagen des statistisches Lernens"</p> <ul style="list-style-type: none"> <li>- Wahrscheinlichkeitsbegriff</li> <li>- stochastische Modellierung</li> <li>- Entropie und Information</li> <li>- explorative Datenanalyse</li> <li>- Likelihood und Bayesianische Inferenz</li> <li>- Resampling Verfahren (Bootstrap und MCMC)</li> <li>- Modellwahl</li> <li>- multiples Testen</li> <li>- hochdimensionale Statistik (Shrinkage und Regularisierung)</li> <li>- Klassifikation</li> <li>- Regressionsmodelle</li> <li>- Zeitreihenanalyse</li> <li>- räumliche Statistik</li> </ul> <p>Praktikum "Statistische Analysen mit R" (<a href="http://r-project.org">http://r-project.org</a>)</p>

- Einführung in die Datenanalyse in R
- statistisches Programmieren in R
- Anwendung auf Beispieldatensätze
- Erstellung von statistischen Berichten

**Teilnahmevoraussetzungen**

Grundkenntnisse in Statistik oder Biometrie oder gleichwertige Kenntnisse

**Literaturangabe**

Hinweise zu Literaturangaben erfolgen in den Lehrveranstaltungen.

**Vergabe von Leistungspunkten**

Leistungspunkte werden mit erfolgreichem Abschluss des Moduls vergeben. Näheres regelt die Prüfungsordnung.

**Prüfungsleistungen und -vorleistungen**

<b>Modulprüfung: Mündliche Prüfung 20 Min., mit Wichtung: 1</b>	
	Vorlesung "Grundlagen des statistischen Lernens" (3SWS)
	Übung "Grundlagen des statistischen Lernens" (1SWS)
	Praktikum "Statistische Analysen mit R" (2SWS)

## Master of Science Data Science

Akademischer Grad	Modulnummer	Modulform
Master of Science	10-201-2301	Wahlpflicht

<b>Modultitel</b>	<b>Text Mining - Wissensrohstoff Text</b>
<b>Modultitel (englisch)</b>	Text Mining - Text as the Raw Material of Knowledge
<b>Empfohlen für:</b>	1./3. Semester
<b>Verantwortlich</b>	Automatische Sprachverarbeitung
<b>Dauer</b>	1 Semester
<b>Modulturnus</b>	jedes Wintersemester
<b>Lehrformen</b>	<ul style="list-style-type: none"> <li>• Vorlesung "Text Mining" (2 SWS) = 30 h Präsenzzeit und 70 h Selbststudium = 100 h</li> <li>• Übung "Text Mining" (1 SWS) = 15 h Präsenzzeit und 70 h Selbststudium = 85 h</li> <li>• Praktikum "Text Mining" (3 SWS) = 45 h Präsenzzeit und 70 h Selbststudium = 115 h</li> </ul>
<b>Arbeitsaufwand</b>	10 LP = 300 Arbeitsstunden (Workload)
<b>Verwendbarkeit</b>	<ul style="list-style-type: none"> <li>• Vertiefungsmodul im B.Sc. Informatik</li> <li>• B.Sc. Digital Humanities</li> <li>• Lehramt Informatik</li> <li>• M.Sc. Data Science</li> <li>• M.Sc. Journalismus</li> <li>• M.Sc. Wirtschaftspädagogik</li> </ul>
<b>Ziele</b>	<p>Nach der aktiven Teilnahme am Modul „Text Mining“ sind die Studierenden in der Lage:</p> <ul style="list-style-type: none"> <li>- grundlegende Begriffe und Verfahren (wie z.B. Alignments, Clustering, Sentiment Analyse) zu definieren,</li> <li>- algorithmische Lösungsansätze (u.a. string- und musterbasierte Verfahren, Bayes'sche Netzwerke) zu erklären und</li> <li>- diese selbständig auf Problemstellungen anzuwenden.</li> </ul> <p>Die Studierenden sind in der Lage, auch in kleinen Gruppen Fragestellungen zu bearbeiten.</p>
<b>Inhalt</b>	<ul style="list-style-type: none"> <li>• Wissen und Text</li> <li>• Grundlagen der Bedeutungsanalyse</li> <li>• Sprachstatistik (Zipf'sche Gesetze, bedingte Wahrscheinlichkeiten, Kookkurrenzanalyse, small worlds)</li> <li>• Clustering</li> <li>• Musteranalyse</li> <li>• Hybride Verfahren</li> <li>• Beispielanwendungen.</li> </ul>
<b>Teilnahmevoraussetzungen</b>	keine
<b>Literaturangabe</b>	Hinweise zu Literaturangaben erfolgen in den Lehrveranstaltungen.
<b>Vergabe von Leistungspunkten</b>	Leistungspunkte werden mit erfolgreichem Abschluss des Moduls vergeben. Näheres regelt die Prüfungsordnung.

## Prüfungsleistungen und -vorleistungen

Modulprüfung:	
Klausur 60 Min., mit Wichtung: 2	Vorlesung "Text Mining" (2SWS)
	Übung "Text Mining" (1SWS)
Präsentation (45 Min.) mit schriftlicher Ausarbeitung (8 Wochen), mit Wichtung: 1	Praktikum "Text Mining" (3SWS)

## Master of Science Data Science

Akademischer Grad	Modulnummer	Modulform
Master of Science	10-201-2503	Wahlpflicht

### Modultitel Grundlagen der IT-Sicherheit

**Modultitel (englisch)** Introduction to IT-Security

**Empfohlen für:** 1./2./3. Semester

**Verantwortlich** Abteilung Technische Informatik

**Dauer** 1 Semester

**Modulturnus** unregelmäßig

**Lehrformen**

- Vorlesung "Grundlagen der IT-Sicherheit" (2 SWS) = 30 h Präsenzzeit und 45 h Selbststudium = 75 h
- Übung "Grundlagen der IT-Sicherheit" (2 SWS) = 30 h Präsenzzeit und 45 h Selbststudium = 75 h

**Arbeitsaufwand** 5 LP = 150 Arbeitsstunden (Workload)

**Verwendbarkeit**

- M.Sc. Data Science
- M.Sc. Informatik

**Ziele**

Nach der aktiven Teilnahme am Modul sind die Studierenden in der Lage

- grundlegende Begriffe zu definieren, deren Spezifika zu benennen und zwischen ihnen zu unterscheiden
- ausgewählte Bedrohungen der Computersicherheit zu erläutern und auf deren Gegenmaßnahmen einzugehen
- Probleme der Netzwerksicherheit und dem Internet zu benennen
- verbreitete Bewertungskriterien sicherer Software und System zu erläutern
- ausgewählte Sicherheitsmodelle zu benennen und deren Eignung zu beurteilen
- algorithmische Lösungsansätze zu erläutern und anzuwenden
- im Rahmen des Security Engineering verschiedene Analysen zu erarbeiten

**Inhalt**

Vorlesung "Grundlagen der IT-Sicherheit"

- Herangehensweisen, Modell, Konzepte und Verfahren der IT-Sicherheit
- Ausgewählte Bedrohungen und Angriffstechniken der Computersicherheit
  - technisch motivierte Bedrohungen wie Computerviren, Pufferüberläufe, Trojanische Pferde, Rootkits, Schwachstellen in Prozessoren sowie soziotechnische Bedrohungen wie Spam und Phishing
  - Gegenmaßnahmen (Malwarescanner, Sandboxing, DEP, ASLR, etc.)
- Spezifika der Internet- und Netzwerksicherheit
  - Sicherheitsprobleme von Protokollen und Netzwerkdienste
  - Cyberwaffen und Botnetze
  - Gegenmaßnahmen (IDPS, Firewall, DPI etc.)
- Mobile Systeme und deren Sicherheit
- Herangehensweisen beim Security Engineering
  - Entwicklungsprozesse, Struktur, Bedarf, Bedrohungsanalyse
  - Attribution und Penetrationstest
  - BSI-Sicherheitsprozess
- Sicherheitsmodelle und Standards,
- Bewertungskriterien sicher Software und Computersysteme
  - Common Criteria, ITSEC
  - Schutzprofile und Zertifizierung

Übung "Grundlagen der IT-Sicherheit"

- Absicherung von Computersystemen
- Anwendung algorithmischer Verfahren der Kryptografie
  - Erzeugung von Schlüsseln, Digitaler Identitäten
  - Verschlüsselte Mailkommunikation
- Security Engineering
  - Anfertigung einer Struktur-, Schutzbedarfs-, Risiko-, Bedrohungsanalyse
  - Ausfertigung einer Bedrohungs- und Risikomatrix
- Erstellen von CC-Evaluationsdokumenten anhand eines technischen Geräts
  - Definition eines ToE und von PP, ST und SFR
  - Evaluation mittels SAR, EAL (Guidelines)

**Teilnahmevoraussetzungen**

keine

**Literaturangabe**

Hinweise zu Literaturangaben erfolgen in den Lehrveranstaltungen.

**Vergabe von Leistungspunkten**

Leistungspunkte werden mit erfolgreichem Abschluss des Moduls vergeben. Näheres regelt die Prüfungsordnung.

**Prüfungsleistungen und -vorleistungen**

<b>Modulprüfung: Klausur 60 Min., mit Wichtung: 1</b>	
	Vorlesung "Grundlagen der IT-Sicherheit" (2SWS)
	Übung "Grundlagen der IT-Sicherheit" (2SWS)



## Master of Science Data Science

Akademischer Grad	Modulnummer	Modulform
Master of Science	10-202-2128	Wahlpflicht

<b>Modultitel</b>	<b>Künstliche Neuronale Netze und Maschinelles Lernen</b> Kernmodul
<b>Modultitel (englisch)</b>	Artificial Neural Networks and Machine Learning Key Module
<b>Empfohlen für:</b>	1./3. Semester
<b>Verantwortlich</b>	Abteilung Technische Informatik
<b>Dauer</b>	1 Semester
<b>Modulturnus</b>	jedes Wintersemester
<b>Lehrformen</b>	<ul style="list-style-type: none"> <li>• Vorlesung "Künstliche neuronale Netze und Maschinelles Lernen" (2 SWS) = 30 h Präsenzzeit und 45 h Selbststudium = 75 h</li> <li>• Seminar "Künstliche neuronale Netze und Maschinelles Lernen" (2 SWS) = 30 h Präsenzzeit und 45 h Selbststudium = 75 h</li> </ul>
<b>Arbeitsaufwand</b>	5 LP = 150 Arbeitsstunden (Workload)
<b>Verwendbarkeit</b>	<ul style="list-style-type: none"> <li>• Kernmodul Technische Informatik im M.Sc. Informatik</li> <li>• Informatikmodul im M.Sc. Bioinformatik</li> <li>• M.Sc. Data Science</li> </ul>
<b>Ziele</b>	<p>Nach der aktiven Teilnahme am Modul "Künstliche Neuronale Netze und Maschinelles Lernen" sind die Studierenden in der Lage:</p> <ul style="list-style-type: none"> <li>- grundlegende Begriffe der Künstlichen Neuronalen Netze und Maschinellen Lernens zu definieren,</li> <li>- ausgewählte Verfahren und Algorithmen zu beschreiben und zu analysieren,</li> <li>- algorithmische Lösungsansätze zu erklären und diese selbstständig auf Problemstellungen der Daten- und Signalverarbeitung anzuwenden,</li> <li>- einen wissenschaftlichen Vortrag zu halten und</li> <li>- eine wissenschaftliche Veröffentlichung zu erstellen.</li> </ul>
<b>Inhalt</b>	Die Studierenden sollen die grundlegenden überwachten und unüberwachten Lernverfahren und Algorithmen der Künstlichen Neuronalen Netze und des Maschinellen Lernens verstehen und die wesentlichen Lösungsansätze auf Problemstellung der Daten- und Signalverarbeitung anwenden können.
<b>Teilnahmevoraussetzungen</b>	keine
<b>Literaturangabe</b>	Hinweise zu Literaturangaben erfolgen in den Lehrveranstaltungen.
<b>Vergabe von Leistungspunkten</b>	Leistungspunkte werden mit erfolgreichem Abschluss des Moduls vergeben. Näheres regelt die Prüfungsordnung.

**Prüfungsleistungen und -vorleistungen**

<b>Modulprüfung:</b>	
Mündliche Prüfung 25 Min., mit Wichtung: 1	Vorlesung "Künstliche neuronale Netze und Maschinelles Lernen" (2SWS)
Referat (20 Min.) mit schriftlicher Ausarbeitung (8 Wochen), mit Wichtung: 1	Seminar "Künstliche neuronale Netze und Maschinelles Lernen" (2SWS)

## Master of Science Data Science

Akademischer Grad	Modulnummer	Modulform
Master of Science	10-202-2201	Wahlpflicht

<b>Modultitel</b>	<b>Wissenschaftliche Visualisierung</b> Vertiefungsmodul
<b>Modultitel (englisch)</b>	Scientific Visualization In-Depth Module
<b>Empfohlen für:</b>	1./3. Semester
<b>Verantwortlich</b>	Abteilung Bild- und Signalverarbeitung
<b>Dauer</b>	1 Semester
<b>Modulturnus</b>	jedes Wintersemester
<b>Lehrformen</b>	<ul style="list-style-type: none"> <li>• Vorlesung "Wissenschaftliche Visualisierung" (4 SWS) = 60 h Präsenzzeit und 120 h Selbststudium = 180 h</li> <li>• Praktikum "Wissenschaftliche Visualisierung" (4 SWS) = 60 h Präsenzzeit und 60 h Selbststudium = 120 h</li> </ul>
<b>Arbeitsaufwand</b>	10 LP = 300 Arbeitsstunden (Workload)
<b>Verwendbarkeit</b>	<ul style="list-style-type: none"> <li>• M.Sc. Biologie</li> <li>• M.Sc. Data Science</li> <li>• M.Sc. Informatik</li> <li>• Staatsexamen Lehramt Informatik Gymnasium</li> </ul>
<b>Ziele</b>	Nach der aktiven Teilnahme am Vertiefungsmodul Wissenschaftliche Visualisierung können die Studierenden alle Grundkonzepte der wissenschaftlichen Visualisierung skizzieren. Die Studierenden können beurteilen, welches wissenschaftliche Visualisierungsverfahren der für eine bestimmte Aufgabe am besten geeignete Ansatz ist. Die Studierenden können grundlegende Verfahren der wissenschaftlichen Visualisierung in Programmen selbständig implementieren.
<b>Inhalt</b>	<p>Das Modul umfasst eine Vorlesung und ein Praktikum, die beide zu belegen sind. Die wissenschaftliche Visualisierung beschäftigt sich mit der Nutzung der Computergrafik zur Generierung von Bildern und Animationen, die einer verbesserten Auswertung von Experimenten und Simulationen durch den Menschen dienen. Sie gehört in vielen Disziplinen zu den grundlegenden Techniken der Datenauswertung. Die Vorlesung behandelt vor allem Prinzipien, Methoden und erfolgreiche Beispiele zur Visualisierung von Felddaten, wie sie bei Simulationen und Messungen in Physik, Chemie, Meteorologie und den Ingenieurwissenschaften, aber auch der Medizin auftreten. Ferner werden Aspekte des Entwurfs von Visualisierungssystemen behandelt. Themen sind u. a. Datenpräsentation, Grundlagen aus Theorie und Anwendungsdomänen, direkte Visualisierung, geometrische Visualisierung, Direct Volume Rendering, topologische Visualisierung, struktur- und merkmalsorientierte Visualisierung, Visualisierungssysteme.</p> <p>Das Praktikum dient der eigenständigen Umsetzung von Verfahren aus der Vorlesung und aktuellen Publikationen, wobei auch Erfahrungen zur Entwicklung ganzer Visualisierungssysteme gewonnen werden.</p>

**Teilnahmevoraussetzungen** keine

**Literaturangabe** Hinweise zu Literaturangaben erfolgen in den Lehrveranstaltungen.

**Vergabe von Leistungspunkten** Leistungspunkte werden mit erfolgreichem Abschluss des Moduls vergeben. Näheres regelt die Prüfungsordnung.

### **Prüfungsleistungen und -vorleistungen**

<b>Modulprüfung: Mündliche Prüfung 30 Min., mit Wichtung: 1</b>	
<i>Prüfungsvorleistung: Praktikumsleistung (Präsentation (30 Min) mit schriftlicher Ausarbeitung) im Praktikum, Bearbeitungszeit (8 Wochen)</i>	
	Vorlesung "Wissenschaftliche Visualisierung" (4SWS)
	Praktikum "Wissenschaftliche Visualisierung" (4SWS)

## Master of Science Data Science

Akademischer Grad	Modulnummer	Modulform
Master of Science	10-202-2205	Wahlpflicht

<b>Modultitel</b>	<b>Graphen und biologische Netze</b>
<b>Modultitel (englisch)</b>	Graphs and Biological Nets
<b>Empfohlen für:</b>	1./3. Semester
<b>Verantwortlich</b>	Lehrstuhl Bioinformatik
<b>Dauer</b>	1 Semester
<b>Modulturnus</b>	jedes Wintersemester
<b>Lehrformen</b>	<ul style="list-style-type: none"> <li>• Vorlesung "Einführungsvorlesung Graphentheorie" (2 SWS) = 30 h Präsenzzeit und 56 h Selbststudium = 86 h</li> <li>• Vorlesung "Aktuelle Forschungsthemen aus dem Bereich Graphen und biologische Netze" (1 SWS) = 15 h Präsenzzeit und 28 h Selbststudium = 43 h</li> <li>• Seminar "Seminar zur Spezialvorlesung" (1 SWS) = 15 h Präsenzzeit und 28 h Selbststudium = 43 h</li> <li>• Praktikum "Praktikum" (3 SWS) = 45 h Präsenzzeit und 83 h Selbststudium = 128 h</li> </ul>
<b>Arbeitsaufwand</b>	10 LP = 300 Arbeitsstunden (Workload)
<b>Verwendbarkeit</b>	<ul style="list-style-type: none"> <li>• Vertiefungsmodul im M.Sc. Informatik, insbesondere im Schwerpunkt Bioinformatik</li> <li>• M.Sc. Data Science</li> </ul>
<b>Ziele</b>	<p>Nach der aktiven Teilnahme am Modul "Graphen und Biologische Netze" sind die Studierenden in der Lage:</p> <ul style="list-style-type: none"> <li>- grundlegende Begriffe und Konzepte der Graphentheorie zu formulieren und zu erklären,</li> <li>- biologische Fragestellungen als graphentheoretische Probleme zu modellieren und mithilfe geeigneter algorithmischer Ansätze zu lösen und</li> <li>- die Ergebnisse im Kontext der biologischen Fragestellung zu interpretieren und kritisch zu diskutieren.</li> </ul>
<b>Inhalt</b>	<p>Grundvorlesung:</p> <ul style="list-style-type: none"> <li>- Grundlegende Eigenschaften von Graphen: Zusammenhang, Planarität, Kreise, Färbungen</li> <li>- Zufallsgraphen</li> </ul> <p>Spezialvorlesung/ Seminar: aktuelle Forschungsthemen, z.B.</p> <ul style="list-style-type: none"> <li>- Metabolische Netzwerke: Flussanalyse, Organisationen, Netzwerk-Evolution</li> <li>- Genregulationsnetzwerke: Dynamik, Stabilität,</li> <li>- Modelle komplexer biologischer Netzwerke: Wachsende Netzwerke, Skalenfreiheit, Selbstähnlichkeit</li> </ul>
<b>Teilnahmevoraussetzungen</b>	keine
<b>Literaturangabe</b>	Hinweise zu Literaturangaben erfolgen in den Lehrveranstaltungen.

**Vergabe von Leistungspunkten**

Für die Vergabe von Leistungspunkten müssen alle vorgesehenen Studienleistungen erbracht sowie die Prüfungsleistung bestanden sein.

**Prüfungsleistungen und -vorleistungen**

<b>Modulprüfung: Mündliche Prüfung 30 Min., mit Wichtung: 1</b>	
<i>Prüfungsvorleistung: • Referat (30 Min.) im Seminar, • Praktikumsleistung als schriftliche Ausarbeitung im Praktikum, Bearbeitungszeit 8 Wochen</i>	
	Vorlesung "Einführungsvorlesung Graphentheorie" (2SWS)
	Vorlesung "Aktuelle Forschungsthemen aus dem Bereich Graphen und biologische Netze" (1SWS)
	Seminar "Seminar zur Spezialvorlesung" (1SWS)
	Praktikum "Praktikum" (3SWS)

\* Diese Prüfungsleistungen müssen bestanden sein.

## Master of Science Data Science

Akademischer Grad	Modulnummer	Modulform
Master of Science	10-202-2207	Wahlpflicht

### Modultitel **Sequenzanalyse und Genomik**

**Modultitel (englisch)** Sequence Analysis and Genomics

**Empfohlen für:** 1./3. Semester

**Verantwortlich** Lehrstuhl für Bioinformatik

**Dauer** 1 Semester

**Modulturnus** jedes Wintersemester

**Lehrformen**

- Vorlesung "Einführungsvorlesung Sequenzanalyse und Genomik" (2 SWS) = 30 h Präsenzzeit und 56 h Selbststudium = 86 h
- Vorlesung "Spezialvorlesung Sequenzanalyse und Genomik" (1 SWS) = 15 h Präsenzzeit und 28 h Selbststudium = 43 h
- Seminar "Sequenzanalyse und Genomik" (1 SWS) = 15 h Präsenzzeit und 28 h Selbststudium = 43 h
- Praktikum "Sequenzanalyse und Genomik" (3 SWS) = 45 h Präsenzzeit und 83 h Selbststudium = 128 h

**Arbeitsaufwand** 10 LP = 300 Arbeitsstunden (Workload)

**Verwendbarkeit**

- Vertiefungsmodul im M.Sc. Informatik
- Pflichtmodul im Schwerpunktfach Bioinformatik
- Wahlpflichtmodul im M.Sc. Biologie
- Wahlpflichtmodul im M.Sc. Biochemie
- Lehramt Informatik
- M.Sc. Data Science

**Ziele**

Nach der aktiven Teilnahme am Modul "Bioinformatik von RNA und Proteinstrukturen" sind die Studierenden in der Lage

- Sequenzdaten im biologischen Kontext zu interpretieren,
- die grundlegenden Algorithmen zum Sequenzvergleich in hinreichender Tiefe zu verstehen, um die geeigneten Werkzeuge für konkrete Anwendungen auszuwählen,
- die grundlegenden Algorithmen zum Sequenzvergleich anzuwenden und in einfacher Weise zu modifizieren,
- einfache Aufgabenstellungen aus der vergleichenden Genomik eigenständig zu bearbeiten und
- die Ergebnisse der praktischen Arbeit zu präsentieren und kritisch zu diskutieren.

**Inhalt**

Vorlesung "Sequenzanalyse und Genomik":

- Exakte und approximative Suche in Sequenzdaten
- lokale und globale Alignierung von Sequenzen
- Phylogenetische Rekonstruktion in Theorie und Praxis
- Einführendes zur Vorhersage von RNA- und Proteinstrukturen.

Eine Spezialvorlesung wird auf einem der folgenden Themengebiete angeboten:

- Evolutionäre Algorithmen: Kombinatorische Optimierungs-Probleme; Simulated Annealing; Werte-Landschaften; Genetische Algorithmen; Genetic Programming.
- Hidden-Markov-Modelle in der Bioinformatik: Grundlagen von HMMs: Baum-Welch- und Viterbi-Algorithmus; Parameterschätzung; paarweise Alignments mit

HMMs; Profile-HMMs für Sequenzfamilien; multiple Alignments mit Lernen von Profile-HMMs.

- Präbiotische Evolution“: Astrophysikalische Grundlagen; Präbiotische Chemie; Chemische Reaktionsnetzwerke; Die RNA Welt und alternative Szenarien; Mathematische Modelle: Quasispecies, Hyperzyklus, und Co.; Der Genetische Code.

Praktikum "Nukleinsäuren" oder Praktikum "Phylogenetische Rekonstruktion":

- Nukleinsäuren“: Praxisnaher Umgang mit Standard-Programmen (u.a. “blast“, “clustalW“ und “dialign“) zur genomweiten Suche und zum Sequenzvergleich.

- Nukleinsäuren“: Suche nach strukturierter Information, wie z.B. Protein-kodierenden Regionen, nicht-kodierenden RNAs oder regulatorischen Elementen in Genomen unter Zuhilfenahme aktueller Werkzeuge und Methoden (z.B. “tracker“, “RNAz“ oder “infernai“)

- Phylogenie“: Rekonstruktion von Phylogenien mit Standard-Werkzeugen wie “phylip“, “MEGA“ oder “NeighborNet“

- Phylogenie“: Problemgerechte Auswahl einer Methode (Maximum Parsimony, Maximum Likelihood oder distanzbasiert); kritische Bewertung von Ergebnissen.

- Nukleinsäuren und Phylogenie“: Umgang mit Datenquellen wie dem “UCSC Genome Browser“.

**Teilnahmevoraussetzungen**

keine

**Literaturangabe**

Hinweise zu Literaturangaben erfolgen in den Lehrveranstaltungen.

**Vergabe von Leistungspunkten**

Leistungspunkte werden mit erfolgreichem Abschluss des Moduls vergeben. Näheres regelt die Prüfungsordnung.

**Prüfungsleistungen und -vorleistungen**

<b>Modulprüfung: Mündliche Prüfung 30 Min., mit Wichtung: 1</b>	
<i>Prüfungsvorleistung: • Referat (30 Min.) im Seminar, • Praktikumsbericht im Praktikum, Bearbeitungszeit 8 Wochen</i>	
	Vorlesung "Einführungsvorlesung Sequenzanalyse und Genomik" (2SWS)
	Vorlesung "Spezialvorlesung Sequenzanalyse und Genomik" (1SWS)
	Seminar "Sequenzanalyse und Genomik" (1SWS)
	Praktikum "Sequenzanalyse und Genomik" (3SWS)



## Master of Science Data Science

Akademischer Grad	Modulnummer	Modulform
Master of Science	10-202-2314	Wahlpflicht

<b>Modultitel</b>	<b>Fortgeschrittene Methoden des Information Retrieval</b>
<b>Modultitel (englisch)</b>	Advanced Methods in Information Retrieval
<b>Empfohlen für:</b>	1./3. Semester
<b>Verantwortlich</b>	Automatische Sprachverarbeitung
<b>Dauer</b>	1 Semester
<b>Modulturnus</b>	jedes Wintersemester
<b>Lehrformen</b>	<ul style="list-style-type: none"> <li>• Vorlesung "Fortgeschrittene Methoden des Information Retrieval" (2 SWS) = 30 h Präsenzzeit und 70 h Selbststudium = 100 h</li> <li>• Praktikum "Fortgeschrittene Methoden des Information Retrieval" (3 SWS) = 45 h Präsenzzeit und 155 h Selbststudium = 200 h</li> </ul>
<b>Arbeitsaufwand</b>	10 LP = 300 Arbeitsstunden (Workload)
<b>Verwendbarkeit</b>	<ul style="list-style-type: none"> <li>• M.Sc. Data Science</li> <li>• M.Sc. Digital Humanities</li> <li>• Vertiefungsmodul im M.Sc. Informatik</li> </ul>
<b>Ziele</b>	<p>Nach der aktiven Teilnahme am Modul "Fortgeschrittene Methoden des Information Retrieval" sind die Studierenden in der Lage:</p> <ul style="list-style-type: none"> <li>- grundlegende Begriffe (wie z.B. Small World, semantische Ähnlichkeit) zu definieren,</li> <li>- algorithmische Lösungsansätze (u.a. für Abstracting, Eigennamenerkennung, Crawling) zu erklären,</li> <li>- ausgewählte Verfahren auf konkrete Fragestellungen anzuwenden und</li> <li>- mit sehr großen Mengen von Textdaten umzugehen.</li> </ul>
<b>Inhalt</b>	<ul style="list-style-type: none"> <li>- Sprachermittlung</li> <li>- Cross-Language Information Retrieval</li> <li>- Ähnlichkeit von Wörtern, Sätzen und Dokumenten</li> <li>- Latent Semantic Analysis</li> <li>- Erkennung von Eigennamen</li> <li>- Sachgebietsermittlung</li> <li>- Automatisches Abstracting</li> <li>- Linkstruktur des Internet</li> <li>- Crawling im Internet.</li> </ul>
<b>Teilnahmevoraussetzungen</b>	keine
<b>Literaturangabe</b>	Hinweise zu Literaturangaben erfolgen in den Lehrveranstaltungen.
<b>Vergabe von Leistungspunkten</b>	Leistungspunkte werden mit erfolgreichem Abschluss des Moduls vergeben. Näheres regelt die Prüfungsordnung.

**Prüfungsleistungen und -vorleistungen**

<b>Modulprüfung: Mündliche Prüfung 30 Min., mit Wichtung: 1</b>	
<i>Prüfungsvorleistung: • Präsentation (45 Min.) im Praktikum</i>	
	Vorlesung "Fortgeschrittene Methoden des Information Retrieval" (2SWS)
	Praktikum "Fortgeschrittene Methoden des Information Retrieval" (3SWS)

## Master of Science Data Science

Akademischer Grad	Modulnummer	Modulform
Master of Science	10-202-2323	Wahlpflicht

<b>Modultitel</b>	<b>Wissens- und Content Management</b> Vertiefungsmodul
<b>Modultitel (englisch)</b>	Information and Content Management In-Depth Module
<b>Empfohlen für:</b>	1./3. Semester
<b>Verantwortlich</b>	Automatische Sprachverarbeitung
<b>Dauer</b>	1 Semester
<b>Modulturnus</b>	jedes Wintersemester
<b>Lehrformen</b>	<ul style="list-style-type: none"> <li>• Vorlesung "Wissens- und Content Management" (2 SWS) = 30 h Präsenzzeit und 70 h Selbststudium = 100 h</li> <li>• Praktikum "Wissens- und Content Management" (3 SWS) = 45 h Präsenzzeit und 155 h Selbststudium = 200 h</li> </ul>
<b>Arbeitsaufwand</b>	10 LP = 300 Arbeitsstunden (Workload)
<b>Verwendbarkeit</b>	<ul style="list-style-type: none"> <li>• M.Sc. Data Science</li> <li>• M.Sc. Digital Humanities</li> <li>• M.Sc. Journalismus</li> <li>• Vertiefungsmodul im M.Sc. Informatik</li> </ul>
<b>Ziele</b>	<p>Nach der aktiven Teilnahme am Modul "Wissens- und Contentmanagement" sind die Studierenden in der Lage,</p> <ul style="list-style-type: none"> <li>- grundlegende Begriffe und Verfahren (wie z.B. Content Objekte, Metadaten, Annotation) zu definieren,</li> <li>- algorithmische Lösungsansätze (u.a. Terminologie-Extraktion, Metadaten basierte Suche) zu erklären,</li> <li>- algorithmische Lösungsansätze selbständig auf Problemstellungen anzuwenden und</li> <li>- in kleinen Gruppen Fragestellungen zu bearbeiten und zu diskutieren.</li> </ul>
<b>Inhalt</b>	<ul style="list-style-type: none"> <li>- Grundlagen des Wissens- und Content Management</li> <li>- Content Management, Content Management Systeme</li> <li>- Strukturieren und Finden von Informationen</li> <li>- Wissensmanagement</li> <li>- Aktuelles Fallbeispiel.</li> </ul>
<b>Teilnahmevoraussetzungen</b>	Teilnahme an den Modulen "Linguistische Informatik" (10-201-2317), "Vertiefung Text Mining/ Wissensrohstoff Text" (10-201-2301) oder gleichwertige Kenntnisse
<b>Literaturangabe</b>	Hinweise zu Literaturangaben erfolgen in den Lehrveranstaltungen.
<b>Vergabe von Leistungspunkten</b>	Leistungspunkte werden mit erfolgreichem Abschluss des Moduls vergeben. Näheres regelt die Prüfungsordnung.

**Prüfungsleistungen und -vorleistungen**

<b>Modulprüfung: Mündliche Prüfung 30 Min., mit Wichtung: 1</b>	
<i>Prüfungsvorleistung: • Präsentation (45 Min.) im Praktikum</i>	
	Vorlesung "Wissens- und Content Management" (2SWS)
	Praktikum "Wissens- und Content Management" (3SWS)

## Master of Science Data Science

Akademischer Grad	Modulnummer	Modulform
Master of Science	10-202-2342	Wahlpflicht

<b>Modultitel</b>	<b>Linguistische Annotationen und Datenextraktion mit XQuery</b>
<b>Modultitel (englisch)</b>	Linguistic Annotation and Data Extraction with XQuery
<b>Empfohlen für:</b>	1./3. Semester
<b>Verantwortlich</b>	Lehrstuhl für Digital Humanities
<b>Dauer</b>	1 Semester
<b>Modulturnus</b>	jedes Wintersemester
<b>Lehrformen</b>	<ul style="list-style-type: none"> <li>• Vorlesung "Linguistic annotation and data extraction with XQuery" (2 SWS) = 30 h Präsenzzeit und 120 h Selbststudium = 150 h</li> <li>• Praktikum "Linguistic annotation and data extraction with XQuery" (0 SWS) = 0 h Präsenzzeit und 150 h Selbststudium = 150 h</li> </ul>
<b>Arbeitsaufwand</b>	10 LP = 300 Arbeitsstunden (Workload)
<b>Verwendbarkeit</b>	<ul style="list-style-type: none"> <li>• M.Sc. Data Science</li> <li>• M.Sc. Digital Humanities</li> <li>• Ergänzungsfach im M.Sc. Informatik</li> </ul>
<b>Ziele</b>	<ol style="list-style-type: none"> <li>1) Die Studierenden werden mit den Grundlagen der Linguistik vertraut gemacht, die notwendig sind, um linguistische Annotationen im informatischen Kontext umfassend verstehen und anwenden zu können.</li> <li>2) Vermittlung von XQuery zur Anwendung und Abfrage von Dokumenten und Annotationen</li> </ol>
<b>Inhalt</b>	<p>Vorlesung:</p> <ol style="list-style-type: none"> <li>1) Die verschiedenen Layers der Linguistischen Annotation: Morphologie, Syntax und Semantik</li> <li>2) XQuery: Navigieren eines XML Dokuments</li> <li>3) XQuery: FLWOR expression</li> <li>4) XQuery: Funktionen</li> </ol> <p>Praktikum:</p> <ol style="list-style-type: none"> <li>1) Entwicklung von Ressourcen und Annotieren eines Pilotkorpus` für eine bestimmte Sprache</li> </ol> <p>Die Lehrveranstaltungen werden in englischer Sprache gehalten. Studien- und Prüfungsleistungen sind in englischer Sprache zu erbringen.</p>
<b>Teilnahmevoraussetzungen</b>	keine
<b>Literaturangabe</b>	Hinweise zu Literaturangaben erfolgen in den Lehrveranstaltungen.
<b>Vergabe von Leistungspunkten</b>	Leistungspunkte werden mit erfolgreichem Abschluss des Moduls vergeben. Näheres regelt die Prüfungsordnung.

**Prüfungsleistungen und -vorleistungen**

<b>Modulprüfung: Praktikumsleistung (Programmierung eines Skripts) 90 Min., mit Wichtung: 1</b>	
	Vorlesung "Linguistic annotation and data extraction with XQuery" (2SWS)
	Praktikum "Linguistic annotation and data extraction with XQuery" (0SWS)

## Master of Science Data Science

Akademischer Grad	Modulnummer	Modulform
Master of Science	10-DIH-0001	Wahlpflicht

<b>Modultitel</b>	<b>Verfahren und Anwendungen in den Digital Humanities</b>
<b>Modultitel (englisch)</b>	Methods and Applications in the Digital Humanities
<b>Empfohlen für:</b>	1./3. Semester
<b>Verantwortlich</b>	Lehrstuhl für Computational Humanities
<b>Dauer</b>	1 Semester
<b>Modulturnus</b>	jedes Wintersemester
<b>Lehrformen</b>	<ul style="list-style-type: none"> <li>• Vorlesung "Einführung in grundlegende Verfahren und Anwendungen der Digital Humanities" (2 SWS) = 30 h Präsenzzeit und 60 h Selbststudium = 90 h</li> <li>• Übung "Praktischer Einsatz von Verfahren und Anwendungen in den Digital Humanities" (2 SWS) = 30 h Präsenzzeit und 60 h Selbststudium = 90 h</li> <li>• Praktikum "Projektarbeit" (2 SWS) = 30 h Präsenzzeit und 90 h Selbststudium = 120 h</li> </ul>
<b>Arbeitsaufwand</b>	10 LP = 300 Arbeitsstunden (Workload)
<b>Verwendbarkeit</b>	<ul style="list-style-type: none"> <li>• M.Sc. Data Science</li> <li>• M.Sc. Digital Humanities</li> </ul>
<b>Ziele</b>	<p>Nach der aktiven Teilnahme am Modul "Verfahren und Anwendungen in den Digital Humanities" sind die Studierenden in der Lage:</p> <ol style="list-style-type: none"> <li>1) grundlegende Forschungsfelder und Forschungsansätze sowie auch aktuelle Trends und Herausforderungen in den Digital Humanities zu skizzieren,</li> <li>2) konkrete Methoden und Tools zur Durchführung von Digital Humanities-Forschungsprojekten korrekt anzuwenden,</li> <li>3) zu entscheiden, welches Verfahren der für eine Aufgabe am besten geeignete Ansatz ist.</li> </ol>
<b>Inhalt</b>	<p>Das Modul umfasst eine Vorlesung zur Einführung in grundlegende Verfahren und Anwendungen in den Digital Humanities. Diese Verfahren und Anwendungen werden in einer begleitenden Übung praktisch erprobt und in einem Praktikum von den Studierenden in Projektarbeiten eigenständig angewandt.</p> <p>Vorlesung:</p> <ul style="list-style-type: none"> <li>- Theoretische Grundlagen der Digital Humanities</li> <li>- Kurzer geschichtlicher Überblick zu Methoden und Anwendungen in den Digital Humanities</li> <li>- Aktuelle Trends und Herausforderungen in den Digital Humanities</li> <li>- Typische Tätigkeiten und Arbeitsabläufe in einem Digital Humanities-Projekt</li> <li>- Systematisierung bestehender Methoden und Anwendungen</li> <li>- Vorstellen konkreter Methoden aus unterschiedlichen Anwendungsbereichen (Text, Bild, Musik, etc.) der Digital Humanities, bspw. Crowdsourcing, Stilometrie, Topic Modeling, Social Network Analysis, Farbanalyse, Melodic Similarity, etc.</li> </ul> <p>Übung:</p> <ul style="list-style-type: none"> <li>- Aufbau einer grundlegenden "Digital Humanities"-Toolbox und praktische Anwendung der Verfahren und Anwendungen aus der Vorlesung</li> </ul>

Praktikum:

- Übersicht zu aktuellen Fragestellungen in den Digital Humanities und Umsetzung eigener Forschungsprojekte

Die Lehrveranstaltungen werden in englischer Sprache gehalten. Studien- und Prüfungsleistungen sind in englischer Sprache zu erbringen.

**Teilnahmevoraussetzungen**

keine

**Literaturangabe**

Hinweise zu Literaturangaben erfolgen in den Lehrveranstaltungen.

**Vergabe von Leistungspunkten**

Leistungspunkte werden mit erfolgreichem Abschluss des Moduls vergeben. Näheres regelt die Prüfungsordnung.

**Prüfungsleistungen und -vorleistungen**

<b>Modulprüfung: Mündliche Prüfung 20 Min., mit Wichtung: 1</b>	
<i>Prüfungsvorleistung: Präsentation (30 Minuten) mit schriftlicher Ausarbeitung (Bearbeitungszeit 8 Wochen)</i>	
	Vorlesung "Einführung in grundlegende Verfahren und Anwendungen der Digital Humanities" (2SWS)
	Übung "Praktischer Einsatz von Verfahren und Anwendungen in den Digital Humanities" (2SWS)
	Praktikum "Projektarbeit" (2SWS)



## Master of Science Data Science

Akademischer Grad	Modulnummer	Modulform
Master of Science	10-INF-DS01	Pflicht

### Modultitel **Skalierbare Datenbanktechnologien 1**

**Modultitel (englisch)** Scalable Database Technologies 1

**Empfohlen für:** 1. Semester

**Verantwortlich** Abteilung Datenbanken

**Dauer** 1 Semester

**Modulturnus** jedes Wintersemester

**Lehrformen**

- Vorlesung "Skalierbare Datenbanktechnologien 1" (2 SWS) = 30 h Präsenzzeit und 45 h Selbststudium = 75 h
- Vorlesung "Skalierbare Datenbanktechnologien 2" (2 SWS) = 30 h Präsenzzeit und 45 h Selbststudium = 75 h
- Seminar "Skalierbare Datenbanktechnologien" (2 SWS) = 30 h Präsenzzeit und 120 h Selbststudium = 150 h

**Arbeitsaufwand** 10 LP = 300 Arbeitsstunden (Workload)

**Verwendbarkeit** • M.Sc. Data Science

**Ziele** Nach der aktiven Teilnahme am Modul "Skalierbare Datenbanktechnologien 1" sind die Studierenden in der Lage verschiedene verteilte und parallele Datenmanagementsysteme anzugeben.

Sie können Eigenschaften und Architekturen von Integrationssystemen sowie Techniken zur Anfragebearbeitung und -optimierung in verteilten und parallelen Datenbanksystemen erklären.

Die Studierenden sind imstande wissenschaftliche Publikationen aus dem Bereich moderne Datenbanktechnologien, Cloud und Big Data zu erläutern und zu präsentieren.

Sie vermögen aktuelle Datenbanktechnologien selbständig in einer schriftlichen Ausarbeitung zu beurteilen.

Die Studierenden sind in der Lage, verschiedene Ansätze in der Gruppe zu diskutieren.

**Inhalt** Es werden mindestens zwei der folgenden Vorlesungen angeboten. Zwei Vorlesungen und das Seminar werden ausgewählt.

Vorlesung Mehrrechner-Datenbanksysteme

- Klassifikation von Mehrrechner-DBS
- Architektur von Verteilten DBS
- Datenverteilung
- Verteilte und parallele Anfrageoptimierung
- Transaktionsverwaltung in Verteilten DBS
- Replizierte DBS
- Cluster-DBS (Shared Disk).

#### Vorlesung Cloud Data Management

- Cloud Computing, Infrastrukturen und Dienste
- Verteilte Dateisysteme
- MapReduce Konzept, MapReduce im Kontext von Datenbanken
- Anwendungsimplementierung in verteilten Umgebungen und
- Optimierungstechniken
- Large-scale Datenanalyse, Analyse-Frameworks

#### Vorlesung Data Mining

- Skalierbare Algorithmen zur Analyse von großen Datenmengen
- Arbeit mit hochdimensionalen Daten
- Analyse von Datenströmen
- Netzwerkanalyse
- Assoziationsregeln, Clusteranalyse, Empfehlungsdienste
- Large-Scale Machine Learning

#### Vorlesung Implementierung von Datenbanksystemen I

- Aufbau von DBS (Schichtenmodell)
- Externspeicherverwaltung: Dateiverwaltung, Einsatz von Speicherhierarchien,
- Disk-Arrays, nicht-flüchtige Halbleiterspeicher
- Pufferverwaltung: Lokalität, Speicherallokation, Seitenersetzung
- Satzverwaltung: Satzadressierung, lange Felder, Column stores
- Indexstrukturen für DBS: B-Bäume, Hash-Verfahren, Grid-File, Text-Indizes, etc.
- Implementierung relationaler Operationen: Selektion, Joins, Sortierung
- Anfragebearbeitung: Übersetzung/Interpretation, Query-Optimierung.

#### Seminar Forschungsseminar Datenbanken

- Präsentieren und Diskussion von Arbeiten aus dem Gebiet der Datenbanktechnologie oder verwandten Gebieten
- Die Themenstellung richtet sich nach den aktuellen Entwicklungen auf dem Gebiet der Datenbanktechnologie bzw. verwandten Gebieten.
- Im Rahmen des Seminars ist eine Ausarbeitung zu einem Teilthema anzufertigen und über ihren Inhalt vorzutragen.

#### **Teilnahmevoraussetzungen**

keine

#### **Literaturangabe**

Hinweise zu Literaturangaben erfolgen in den Lehrveranstaltungen.

#### **Vergabe von Leistungspunkten**

Leistungspunkte werden mit erfolgreichem Abschluss des Moduls vergeben. Näheres regelt die Prüfungsordnung.

#### **Prüfungsleistungen und -vorleistungen**

<b>Modulprüfung: Klausur 120 Min., mit Wichtung: 2</b>	
	Vorlesung "Skalierbare Datenbanktechnologien 1" (2SWS)
	Vorlesung "Skalierbare Datenbanktechnologien 2" (2SWS)
Referat (60 Min.) mit schriftlicher Ausarbeitung (4 Wochen), mit Wichtung: 1	Seminar "Skalierbare Datenbanktechnologien" (2SWS)

## Master of Science Data Science

Akademischer Grad	Modulnummer	Modulform
Master of Science	07-202-2302	Wahlpflicht

### Modultitel **Multivariate Statistik und Data Mining**

**Modultitel (englisch)** Multivariate Statistics and Data Mining

**Empfohlen für:** 2. Semester

**Verantwortlich** Professur für Statistik

**Dauer** 1 Semester

**Modulturnus** Sommersemester (im ungeradzahligen Jahr)

**Lehrformen**

- Vorlesung "Multivariate Statistik und Data Mining" (4 SWS) = 60 h Präsenzzeit und 140 h Selbststudium = 200 h
- Übung "Multivariate Statistik und Data Mining" (2 SWS) = 30 h Präsenzzeit und 70 h Selbststudium = 100 h

**Arbeitsaufwand** 10 LP = 300 Arbeitsstunden (Workload)

**Verwendbarkeit**

- M.Sc. Betriebswirtschaftslehre (Management Science)
- M.Sc. Data Science
- M.Sc. Volkswirtschaftslehre (Economics)
- M.Sc. Wirtschaftsinformatik (Business Information Systems)
- M.Sc. Wirtschaftspädagogik (Business Education and Management Training)
- M.Sc. Wirtschaftswissenschaften/ Sciences Économiques

**Ziele**

Die Studierenden können die theoretischen Grundlagen, d.h. Verfahren zur Dimensionsreduktion, Clusteranalyse, Diskriminanzanalyse, Klassifikations- und Assoziationsverfahren und künstliche neuronale Netze sowie den aktuellen Stand der Forschung im Bereich der Multivariaten Statistik sowie des Data Mining darstellen und beschreiben.

Sie sind in der Lage, selbständig geeignete Methoden zur Lösung volks- und betriebswirtschaftlicher Problemstellungen auszuwählen, anzuwenden sowie die ermittelten Ergebnisse zu interpretieren und kritisch zu bewerten. Hierbei erweitern sie ihren Wissensstand neben statistischen Ansätzen um aus der Informatik entlehnte Methoden des Maschinenlernens.

Die Studierenden können in Übungen die Charakteristika umfangreicher multivariater Datensätze erkennen, mittels geeigneter Methoden die Anzahl der zu analysierenden Dimensionen reduzieren sowie mittels Klassifikations- und Assoziationsverfahren Ähnlichkeiten und komplexe Zusammenhänge in den Daten offenlegen. Sie sind in der Lage, die Statistiksoftwarepakete R und SPSS als Werkzeuge statistischer Analysen einzusetzen und ihre erzielten Ergebnisse kritisch auszuwerten.

**Inhalt**

- Aufgaben, Ziele und Konzepte multivariater Analysen
- Methoden der multivariaten Statistik: Hauptkomponenten- und Faktorenanalyse, Clusteranalyse, Diskriminanzanalyse
- Methoden des Maschinenlernens: Klassifikationsbäume, Assoziationsregeln, Künstliche Neuronale Netze

**Teilnahmevoraussetzungen** keine

**Literaturangabe** Hinweise zu Literaturangaben erfolgen in den Lehrveranstaltungen.

**Vergabe von Leistungspunkten** Leistungspunkte werden mit erfolgreichem Abschluss des Moduls vergeben. Näheres regelt die Prüfungsordnung.

**Prüfungsleistungen und -vorleistungen**

<b>Modulprüfung: Klausur 90 Min., mit Wichtung: 1</b>	
	Vorlesung "Multivariate Statistik und Data Mining" (4SWS)
	Übung "Multivariate Statistik und Data Mining" (2SWS)

## Master of Science Data Science

Akademischer Grad	Modulnummer	Modulform
Master of Science	10-201-2502	Wahlpflicht

### Modultitel **Holistisches Praktikum der IT-Sicherheit**

**Modultitel (englisch)** Holistic IT-Security Lab

**Empfohlen für:** 2./3. Semester

**Verantwortlich** Abteilung Technische Informatik

**Dauer** 1 Semester

**Modulturnus** unregelmäßig

**Lehrformen** • Praktikum "IT-Sicherheit" (2 SWS) = 30 h Präsenzzeit und 120 h Selbststudium = 150 h

**Arbeitsaufwand** 5 LP = 150 Arbeitsstunden (Workload)

**Verwendbarkeit** • M.Sc. Data Science  
• M.Sc. Informatik

**Ziele** Nach der aktiven Teilnahme am Modul sind die Studierenden in der Lage, Angriffe auf IT-Systeme aktiv zu verstehen, Gegenmaßnahmen zu erarbeiten und durchzusetzen.

Die Studierenden sind anschließend imstande, Sicherheitslücken selbstständig zu identifizieren, Angriffe zu implementieren und Sicherheitstechniken zu entwickeln.

Sie lernen arbeitsteilig und effizient im Team zu arbeiten, Probleme zu identifizieren und zielgerichtet zu lösen.

### Inhalt

Das Praktikum ist zweigeteilt. Der erste Teil besteht aus einem schrittweisen Heranführen an grundlegende Techniken und Werkzeuge. Die Lerninhalte sind dabei

- Erläuterung eines umsichtigen, sicherheitsbewussten und analytischen Vorgehens, um Sicherheitsprobleme aufzufinden (White Hat Hacker)
- Grundlagen
- Implementierung und Ausnutzung grundlegender Schwachstellen (Nachstellen)
- Umgang mit dem Debugger und Disassembler
- Lösen mehrerer Hackme zum Finden von Schwachstellen
- Umgang mit mehreren Analysewerkzeugen
- Erkennen und Bewerten von Schwachstellen

Im zweiten Teil werden die Werkzeuge und Techniken auf universitätsinterne Sicherheitsobjekte angewandt. Ziel ist es Schwachstellen zu entdecken.

- Die Studierenden formieren sich zu kleineren Gruppen (2-3) und erarbeiten Zielstellung, Analyse- und Ausführungsplan.

- Die Studierenden dokumentieren ihr Vorgehen und präsentieren die Ergebnisse am Ende des Praktikums.

Hinweise: Die gefundenen Schwachstellen werden vom Betreuer bewertet und Empfehlungen für das weitere Vorgehen gegeben. Handelt es sich um eine sicherheitskritische, datenschutzrelevante Schwachstelle mit hohem Bedrohungspotential wird von einer weiteren Untersuchung abgesehen und das

entsprechende Institut / Verwaltungseinheit sofort informiert.

Modulprüfungen finden (1) über Hackmes, (2) dem Implementieren und Disassemblieren von Code sowie (3) der Bewertung des Vortrags (Herangehensweise, Dokumentation, Ergebnisse) statt.

**Teilnahmevoraussetzungen**

keine

**Literaturangabe**

Hinweise zu Literaturangaben erfolgen in den Lehrveranstaltungen.

**Vergabe von Leistungspunkten**

Leistungspunkte werden mit erfolgreichem Abschluss des Moduls vergeben. Näheres regelt die Prüfungsordnung.

**Prüfungsleistungen und -vorleistungen**

<b>Modulprüfung: Praktikumsleistung (2 Testate (a 10 Min.) und 1 Vortrag (30 Min.)), mit Wichtung: 1</b>	
	Praktikum "IT-Sicherheit" (2SWS)

## Master of Science Data Science

Akademischer Grad	Modulnummer	Modulform
Master of Science	10-202-2218S	Wahlpflicht

### Modultitel **Grundlagen Komplexer Systeme (S)**

**Modultitel (englisch)** Foundations of Complex Systems (S)

**Empfohlen für:** 2. Semester

**Verantwortlich** Professur für Schwarmintelligenz und Komplexe Systeme

**Dauer** 1 Semester

**Modulturnus** jedes Sommersemester

**Lehrformen**

- Vorlesung "Grundlagen Komplexer Systeme" (2 SWS) = 30 h Präsenzzeit und 45 h Selbststudium = 75 h
- Seminar "Grundlagen Komplexer Systeme" (2 SWS) = 30 h Präsenzzeit und 45 h Selbststudium = 75 h

**Arbeitsaufwand** 5 LP = 150 Arbeitsstunden (Workload)

**Verwendbarkeit** • M.Sc. Data Science

**Ziele** Nach der aktiven Teilnahme am Modul "Grundlagen Komplexer Systeme" sind die Studierenden in Lage:

- grundlegende Begriffe beispielhafter komplexe Systeme zu formulieren und zu erklären,
- für mindestens ein ausgewähltes komplexe System (z.B. Schwarmsysteme, diskrete Simulationssysteme oder Zellularautomaten) grundlegende Funktionsprinzipien zu analysieren und
- diese Funktionsprinzipien selbständig so einzusetzen, dass das System zur Lösung von Problemstellungen eingesetzt werden kann.

**Inhalt** Es muss eine der folgenden Vorlesungen und das Seminar gewählt werden.

Vorlesung "Diskrete Simulation": Simulationsparadigmen, Grundlagen von Warteschlangen/Bediensystemen, Formale Modelle für Diskrete Ereignissysteme und Systemspezifikation, Ein- und Ausgabemodellierung, Simulationssprachen, Parallele/Verteilte Simulation.

Vorlesung "Zellularautomaten": Berechnungsmächtigkeit, Selbstreproduktion, Schnelles Sortieren, Synchronisations- und Markierungsprobleme, Diskretisierung kontinuierlicher Systeme, Modellierung realer Phänomene.

Vorlesung "Verfahren der Schwarm Intelligenz": Ameisenalgorithmen, Schwarmalgorithmen, Prinzipien der Selbstorganisation in biologischen Systemen und ihre Nutzung in der Informatik.

Seminar "Grundlagen Komplexer Systeme": Das Seminar behandelt ergänzende Arbeiten zu den Themen der Vorlesungen und befasst sich mit aktuellen wissenschaftlichen Abhandlungen aus den komplexen Systemen.

**Teilnahmevoraussetzungen** Die Module "Grundlagen Komplexer Systeme (S)" (10-202-2218S) und "Grundlagen Komplexer Systeme (V)" (10-202-2218V) schließen sich gegenseitig aus.

**Literaturangabe** Hinweise zu Literaturangaben erfolgen in den Lehrveranstaltungen.

**Vergabe von Leistungspunkten** Leistungspunkte werden mit erfolgreichem Abschluss des Moduls vergeben. Näheres regelt die Prüfungsordnung.

**Prüfungsleistungen und -vorleistungen**

<b>Modulprüfung: Mündliche Prüfung 20 Min., mit Wichtung: 1</b>	
	Vorlesung "Grundlagen Komplexer Systeme" (2SWS)
Referat (45 Min.) mit schriftlicher Ausarbeitung (2 Wochen), mit Wichtung: 1	Seminar "Grundlagen Komplexer Systeme" (2SWS)



## Master of Science Data Science

Akademischer Grad	Modulnummer	Modulform
Master of Science	10-202-2218V	Wahlpflicht

<b>Modultitel</b>	<b>Grundlagen Komplexer Systeme (V)</b>
<b>Modultitel (englisch)</b>	Foundations of Complex Systems (V)
<b>Empfohlen für:</b>	2. Semester
<b>Verantwortlich</b>	Professur für Schwarmintelligenz und Komplexe Systeme
<b>Dauer</b>	1 Semester
<b>Modulturnus</b>	jedes Sommersemester
<b>Lehrformen</b>	<ul style="list-style-type: none"> <li>• Vorlesung "Grundlagen Komplexer Systeme" (2 SWS) = 30 h Präsenzzeit und 45 h Selbststudium = 75 h</li> <li>• Vorlesung mit seminaristischem Anteil "Grundlagen Komplexer Systeme 2" (2 SWS) = 30 h Präsenzzeit und 45 h Selbststudium = 75 h</li> </ul>
<b>Arbeitsaufwand</b>	5 LP = 150 Arbeitsstunden (Workload)
<b>Verwendbarkeit</b>	• M.Sc. Data Science
<b>Ziele</b>	<p>Nach der aktiven Teilnahme am Modul "Grundlagen Komplexer Systeme" sind die Studierenden in Lage:</p> <ul style="list-style-type: none"> <li>- grundlegende Begriffe beispielhafter komplexe Systeme zu formulieren und zu erklären,</li> <li>- für mindestens ein ausgewähltes komplexe System (z.B. Schwarmssysteme, diskrete Simulationssysteme oder Zellularautomaten) grundlegende Funktionsprinzipien zu analysieren und</li> <li>- diese Funktionsprinzipien selbständig so einzusetzen, dass das System zur Lösung von Problemstellungen eingesetzt werden kann.</li> </ul>
<b>Inhalt</b>	<p>Studierende wählen die Vorlesung "Grundlagen Komplexer Systeme" und eine Vorlesung mit seminaristischem Anteil.</p> <p>Diskrete Simulation: Simulationsparadigmen, Grundlagen von Warteschlangen/Bediensystemen, Formale Modelle für Diskrete Ereignissysteme und Systemspezifikation, Ein- und Ausgabemodellierung, Simulationssprachen, Parallele/Verteilte Simulation.</p> <p>Vorlesungen mit seminaristischem Anteil  Zellularautomaten: Berechnungsmächtigkeit, Selbstreproduktion, Schnelles Sortieren, Synchronisations- und Markierungsprobleme, Diskretisierung kontinuierlicher Systeme, Modellierung realer Phänomene.</p> <p>Verfahren der Schwarm Intelligenz: Ameisenalgorithmen, Schwarmalgorithmen, Prinzipien der Selbstorganisation in biologischen Systemen und ihre Nutzung in der Informatik.</p>
<b>Teilnahmevoraussetzungen</b>	Die Module "Grundlagen Komplexer Systeme (S)" (10-202-2218S) und "Grundlagen Komplexer Systeme (V)" (10-202-2218V) schließen sich gegenseitig aus.

**Literaturangabe** Hinweise zu Literaturangaben erfolgen in den Lehrveranstaltungen.

**Vergabe von Leistungspunkten** Leistungspunkte werden mit erfolgreichem Abschluss des Moduls vergeben. Näheres regelt die Prüfungsordnung.

**Prüfungsleistungen und -vorleistungen**

<b>Modulprüfung: Mündliche Prüfung 20 Min., mit Wichtung: 1</b>	
	Vorlesung "Grundlagen Komplexer Systeme" (2SWS)
	Vorlesung mit seminaristischem Anteil "Grundlagen Komplexer Systeme 2" (2SWS)

## Master of Science Data Science

Akademischer Grad	Modulnummer	Modulform
Master of Science	10-202-2322	Wahlpflicht

<b>Modultitel</b>	<b>Textdatenbanken</b>
	Kernmodul
<b>Modultitel (englisch)</b>	Text Databases
	Key Module
<b>Empfohlen für:</b>	2. Semester
<b>Verantwortlich</b>	Automatische Sprachverarbeitung
<b>Dauer</b>	1 Semester
<b>Modulturnus</b>	jedes Sommersemester
<b>Lehrformen</b>	<ul style="list-style-type: none"> <li>• Vorlesung "Textdatenbanken" (2 SWS) = 30 h Präsenzzeit und 70 h Selbststudium = 100 h</li> <li>• Übung "Textdatenbanken" (1 SWS) = 15 h Präsenzzeit und 35 h Selbststudium = 50 h</li> </ul>
<b>Arbeitsaufwand</b>	5 LP = 150 Arbeitsstunden (Workload)
<b>Verwendbarkeit</b>	<ul style="list-style-type: none"> <li>• M.Sc. Data Science</li> <li>• Kernmodul der Angewandten Informatik im M.Sc. Informatik</li> <li>• M.Sc. Journalismus</li> <li>• M.Sc. Wirtschaftsinformatik</li> </ul>
<b>Ziele</b>	<p>Nach der aktiven Teilnahme am Modul "Textdatenbanken" sind die Studierenden in der Lage:</p> <ul style="list-style-type: none"> <li>- grundlegende Begriffe (wie z.B. Kookkurrenz, Wortnetz, Desambiguierung) zu definieren,</li> <li>- algorithmische Lösungsansätze (u.a. POS-Tagging, Finden bedeutungsähnlicher Wörter) zu erklären und</li> <li>- algorithmische Lösungsansätze selbständig auf Problemstellungen anzuwenden.</li> </ul>
<b>Inhalt</b>	<ul style="list-style-type: none"> <li>- Aufbau von Textdatenbanken: Sammeln, Aufbereiten, Indexieren</li> <li>- Statistische Analysemethoden</li> <li>- Kookkurrenzen</li> <li>- Part-Of-Speech-Tagging</li> <li>- Semantische Wortnetze</li> <li>- Grundformreduktion, Kompositazerlegung, Desambiguierung</li> <li>- Finden bedeutungsähnlicher Wörter</li> <li>- Linguistische Suchmaschinen.</li> </ul>
<b>Teilnahmevoraussetzungen</b>	keine
<b>Literaturangabe</b>	Hinweise zu Literaturangaben erfolgen in den Lehrveranstaltungen.
<b>Vergabe von Leistungspunkten</b>	Leistungspunkte werden mit erfolgreichem Abschluss des Moduls vergeben. Näheres regelt die Prüfungsordnung.

**Prüfungsleistungen und -vorleistungen**

<b>Modulprüfung: Klausur 60 Min., mit Wichtung: 1</b>	
	Vorlesung "Textdatenbanken" (2SWS)
	Übung "Textdatenbanken" (1SWS)

## Master of Science Data Science

Akademischer Grad	Modulnummer	Modulform
Master of Science	10-INF-DS101	Wahlpflicht

### Modultitel **Skalierbare Datenbanktechnologien 2**

**Modultitel (englisch)** Scalable Database Technologies 2

**Empfohlen für:** 2. Semester

**Verantwortlich** Abteilung Datenbanken

**Dauer** 1 Semester

**Modulturnus** jedes Sommersemester

**Lehrformen**

- Vorlesung "Skalierbare Datenbanktechnologien 3" (2 SWS) = 30 h Präsenzzeit und 30 h Selbststudium = 60 h
- Vorlesung "NoSQL-Datenbanken" (1 SWS) = 15 h Präsenzzeit und 15 h Selbststudium = 30 h
- Übung "NoSQL-Datenbanken" (1 SWS) = 15 h Präsenzzeit und 45 h Selbststudium = 60 h

**Arbeitsaufwand** 5 LP = 150 Arbeitsstunden (Workload)

**Verwendbarkeit** • M.Sc. Data Science

**Ziele** Nach der aktiven Teilnahme am Modul "Skalierbare Datenbanktechnologien 2" sind die Studierenden in der Lage verschiedene Architekturen aktueller Datenbankanwendungen anzugeben.

Sie können Verfahren zur Verarbeitung großer Datenmengen erläutern.

Die Studierenden sind imstande aktuelle Datenbanktechnologien und Verfahren zur Verarbeitung und Analyse großer Datenmengen anzuwenden.

Sie sind in der Lage, selbstständig Anfragen durchzuführen.

### Inhalt

Vorlesung NoSQL-Datenbanken

- Verwaltung großer Datenmengen in verteilten Clusterumgebungen
- Kategorisierung und Eigenschaften von NoSQL-Datenbanksystemen
- Vergleich von NoSQL-Systemen zu relationalen Datenbanksystemen
- Partitionierung, Konsistenz, Replikation
- Key-Value, Document Stores und Extensible Record Stores
- Graphdatenbanken

Übung NoSQL-Datenbanken

- Verständnisaufgaben zum Inhalt der Vorlesung
- Praktische Arbeit mit realen, teilweise sehr großen, Datensätzen
- Formulierung und Ausführung von Anfragen in den jeweiligen Sprachen verschiedener Datenbankimplementierungen
- Umgang mit den Anbindungen an gängige Programmiersprachen

Außerdem wird mindestens eine der folgenden Vorlesungen angeboten. Eine Vorlesung wird ausgewählt.

Vorlesung Data Warehousing

- Architektur von Data Warehouse-Systemen
- Mehrdimensionale Modellierung
- Datenintegration, Datenbereinigung, ETL-Werkzeuge
- Performance-Techniken: Indexstrukturen, materialisierte Sichten, parallele Datenbanken
- Data Mining-Verfahren
- Anwendungen von Datawarehouses

Vorlesung Implementierung von Datenbanksystemen II

- Synchronisation: Serialisierbarkeit, Sperrverfahren, Deadlock-Behandlung,
- Mehrversionenverfahren, sonstige Synchronisationsansätze
- Logging und Recovery: Fehlermodell, Logging-Strategien, Checkpoint-Ansätze, Crash-Recovery, Media-Recovery
- Erweiterte Transaktionsmodelle (geschachtelte Transaktionen, verkettete Transaktionen, etc.)
- DB-Benchmarks.

Vorlesung Datenintegration

- Überblick zur Integration verteilter, heterogener Datenbestände
- Verteilung, Autonomie und Heterogenität
- Eigenschaften von Integrationssystemen
- Architekturen von Integrationssystemen
- Anfrageverarbeitung
- Schemamanagement
- Datenfusion

**Teilnahmevoraussetzungen**

keine

**Literaturangabe**

Hinweise zu Literaturangaben erfolgen in den Lehrveranstaltungen.

**Vergabe von Leistungspunkten**

Leistungspunkte werden mit erfolgreichem Abschluss des Moduls vergeben. Näheres regelt die Prüfungsordnung.

**Prüfungsleistungen und -vorleistungen**

<b>Modulprüfung: Klausur 120 Min., mit Wichtung: 1</b>	
<i>Prüfungsvorleistung: Testat (60 Min) in der Übung "NoSQL-Datenbanken"</i>	
	Vorlesung "Skalierbare Datenbanktechnologien 3" (2SWS)
	Vorlesung "NoSQL-Datenbanken" (1SWS)
	Übung "NoSQL-Datenbanken" (1SWS)

## Master of Science Data Science

Akademischer Grad	Modulnummer	Modulform
Master of Science	10-INF-DS102	Wahlpflicht

### Modultitel **Big Data Praktikum**

**Modultitel (englisch)** Big Data Project

**Empfohlen für:** 2. Semester

**Verantwortlich** Abteilung Datenbanken

**Dauer** 1 Semester

**Modulturnus** jedes Sommersemester

**Lehrformen** • Praktikum "Big Data Praktikum" (2 SWS) = 30 h Präsenzzeit und 120 h Selbststudium = 150 h

**Arbeitsaufwand** 5 LP = 150 Arbeitsstunden (Workload)

**Verwendbarkeit** • M.Sc. Data Science

**Ziele** Nach der aktiven Teilnahme am Modul "Big Data Praktikum" sind die Studierenden in der Lage, aktuelle Datenbank-/Big-Data-Technologien anzuwenden.

Sie können selbständig Anfragen formulieren und ausführen.

Die Studierenden sind imstande, Algorithmen zur effizienten Verarbeitung großer Datenmengen zu implementieren.

**Inhalt**

- Entwurf und Realisierung einer Anwendung oder eines Algorithmus unter Verwendung von existierenden Big Data Frameworks und Technologien,
- Anfertigung eines Entwurfs zur Darstellung des Konzepts und der Architektur zur geplanten Implementierung unter Einbeziehung des jeweils verwendeten Frameworks,
- Implementierung einer Anwendung und zur Verfügung stellen eines dokumentierten, ausführbaren Programms oder Ablaufplans,
- Präsentation der Anwendung und Ergebnisse des Praktikums

**Teilnahmevoraussetzungen** keine

**Literaturangabe** Hinweise zu Literaturangaben erfolgen in den Lehrveranstaltungen.

**Vergabe von Leistungspunkten** Leistungspunkte werden mit erfolgreichem Abschluss des Moduls vergeben. Näheres regelt die Prüfungsordnung.

### Prüfungsleistungen und -vorleistungen

<b>Modulprüfung: Praktikumsleistung (3 Testate a 60 Min.), mit Wichtung: 1</b>	
	Praktikum "Big Data Praktikum" (2SWS)

## Master of Science Data Science

Akademischer Grad	Modulnummer	Modulform
Master of Science	10-INF-DS301	Wahlpflicht

### Modultitel **Aktuelle Trends in Data Science**

**Modultitel (englisch)** Current Trends in Data Science

**Empfohlen für:** 2. Semester

**Verantwortlich** Institut für Informatik

**Dauer** 1 Semester

**Modulturnus** jedes Sommersemester

**Lehrformen**

- Vorlesung "Aktuelle Trends in Data Science" (2 SWS) = 30 h Präsenzzeit und 70 h Selbststudium = 100 h
- Übung "Aktuelle Trends in Data Science" (1 SWS) = 15 h Präsenzzeit und 35 h Selbststudium = 50 h

**Arbeitsaufwand** 5 LP = 150 Arbeitsstunden (Workload)

**Verwendbarkeit** • M.Sc. Data Science

**Ziele** Nach der aktiven Teilnahme am Modul "Aktuelle Trends in Data Science" sind die Studierenden in der Lage, Begriffe und Verfahren eines aktuellen Themas im Bereich Data Science zu benennen und zu erklären.

Sie können ausgewählte Verfahren und Algorithmen analysieren, beurteilen und diese selbstständig auf Problemstellungen anwenden.

Die Studierenden sind imstande, verschiedene Ansätze in der Gruppe zu diskutieren.

**Inhalt** Wechselnde aktuelle Themen und Trends im Bereich Data Science, die sich besonderer Aktualität erfreuen. Die Veranstaltung kann auch von Gästen des Instituts für ein eigenes Lehrangebot genutzt werden, das nicht in die bestehenden Module passt. Der konkrete Inhalt wird im Vorlesungsverzeichnis bekannt gegeben.

**Teilnahmevoraussetzungen** keine

**Literaturangabe** Hinweise zu Literaturangaben erfolgen in den Lehrveranstaltungen.

**Vergabe von Leistungspunkten** Leistungspunkte werden mit erfolgreichem Abschluss des Moduls vergeben. Näheres regelt die Prüfungsordnung.

### Prüfungsleistungen und -vorleistungen

Modulprüfung: Mündliche Prüfung 30 Min., mit Wichtung: 1	
	Vorlesung "Aktuelle Trends in Data Science" (2SWS)
	Übung "Aktuelle Trends in Data Science" (1SWS)



## Master of Science Data Science

Akademischer Grad	Modulnummer	Modulform
Master of Science	10-MAT-DSTAT	Wahlpflicht

<b>Modultitel</b>	<b>Mathematische Statistik</b>
<b>Modultitel (englisch)</b>	Mathematical Statistics
<b>Empfohlen für:</b>	2. Semester
<b>Verantwortlich</b>	Professuren für Angewandte Mathematik und Stochastik
<b>Dauer</b>	1 Semester
<b>Modulturnus</b>	jedes Sommersemester
<b>Lehrformen</b>	<ul style="list-style-type: none"> <li>• Vorlesung "Mathematische Statistik" (4 SWS) = 60 h Präsenzzeit und 140 h Selbststudium = 200 h</li> <li>• Übung "Statistik am Rechner" (2 SWS) = 30 h Präsenzzeit und 70 h Selbststudium = 100 h</li> </ul>
<b>Arbeitsaufwand</b>	10 LP = 300 Arbeitsstunden (Workload)
<b>Verwendbarkeit</b>	• M.Sc. Data Science
<b>Ziele</b>	<p>Nach der aktiven Teilnahme am Modul Mathematische Statistik sind die Studierenden in der Lage, grundlegende Begriffe der Mathematischen Statistik (wie z.B. Schätzer, Statistisches Modell, Stichprobenraum, Statistik,...) zu definieren. Sie kennen die Standardfragen der Statistik (wie z.B. Schätzen, Testen, Entscheiden, Regredieren) und können entsprechende Aufgaben auf der Grundlage mathematischer Modelle lösen.</p> <p>Die Studierenden kennen die Standardprogramme zur Bearbeitung Statistischer Probleme und können diese zur Lösung von Aufgaben aus dem Bereich der Mathematischen Statistik einsetzen.</p>
<b>Inhalt</b>	<ul style="list-style-type: none"> <li>- Einführung in die Statistik</li> <li>- Parameterschätzung</li> <li>- Verteilungen rund um die Normalverteilung</li> <li>- Testen</li> <li>- Nichtparametrische Modelle</li> <li>- Lineare Modelle</li> </ul>
<b>Teilnahmevoraussetzungen</b>	Teilnahme am Modul Wahrscheinlichkeitstheorie (10-201-1802) oder äquivalente Vorkenntnisse
<b>Literaturangabe</b>	Hans-Otto Georgii, "Stochastik", de Gruyter (2. Auflage)
<b>Vergabe von Leistungspunkten</b>	Leistungspunkte werden mit erfolgreichem Abschluss des Moduls vergeben. Näheres regelt die Prüfungsordnung.

**Prüfungsleistungen und -vorleistungen**

<b>Modulprüfung: Klausur 90 Min., mit Wichtung: 1</b>	
	Vorlesung "Mathematische Statistik" (4SWS)
	Übung "Statistik am Rechner" (2SWS)

## Master of Science Data Science

Akademischer Grad	Modulnummer	Modulform
Master of Science	10-201-2219S	Wahlpflicht

### Modultitel Grundlagen der Parallelverarbeitung (S)

**Modultitel (englisch)** Foundations of Parallel Processing (S)

**Empfohlen für:** 3. Semester

**Verantwortlich** Professur für Schwarmintelligenz und Komplexe Systeme

**Dauer** 1 Semester

**Modulturnus** jedes Wintersemester

**Lehrformen**

- Vorlesung "Grundlagen der Parallelverarbeitung" (2 SWS) = 30 h Präsenzzeit und 45 h Selbststudium = 75 h
- Seminar "Grundlagen der Parallelverarbeitung" (2 SWS) = 30 h Präsenzzeit und 45 h Selbststudium = 75 h

**Arbeitsaufwand** 5 LP = 150 Arbeitsstunden (Workload)

**Verwendbarkeit** • M.Sc. Data Science

**Ziele** Nach der aktiven Teilnahme am Modul "Grundlagen der Parallelverarbeitung" sind die Studierenden in Lage:

- grundlegende Begriffe und Konzepte der Parallelverarbeitung zu formulieren und zu erklären,
- grundlegende parallele algorithmische Verfahren und Rechnermodelle (u.a. Sortieralgorithmen, Hardware- Addition) zu analysieren und zu vergleichen und
- für grundlegende algorithmische Probleme selbständig parallele Lösungsverfahren zu entwerfen.

**Inhalt** Es wird eine der folgenden Vorlesungen und das Seminar gewählt.

Vorlesung "Parallele Algorithmen": Grundlegende Konzepte und Bewertungskriterien für parallele Algorithmen, PRAM-Modell, Parallele Algorithmen für grundlegende Probleme wie Sortieren oder Mergen, Grundlagen von Hardware Algorithmen.

Vorlesung "Parallele Berechnungsmodelle": Grundlegender Aufbau von Parallelrechnern, Einführung in realistische Parallelerrechnermodelle, Varianten des BSP-Modells, Varianten des LogP-Modells, Auswirkungen der Modelle auf den Entwurf von Algorithmen, Algorithmische Lösung von Beispielproblemen.

Vorlesung "Rekonfigurierbare Rechensysteme": Einsatzbereiche rekonfigurierbarer Rechensysteme, Typen rekonfigurierbarer Rechensysteme, Aufbau von Field Programmable Gate Arrays (FPGAs), Theoretische Konzepte der Rekonfigurierbarkeit, Grundlegende Algorithmen zu dynamischer Rekonfiguration

Vorlesung "Entwurf und Implementierung paralleler Algorithmen": Parallele Plattformen, Entwurfsprinzipien, Analytische Modellierung, Parallele Programmierung für nachrichtengekoppelte und speichergekoppelte Parallelrechner, Matrixmultiplikation, Sortieren, Graphenalgorithmen, Diskrete

Optimierung, Dynamische Programmierung.

Seminar "Grundlagen der Parallelverarbeitung": Das Seminar behandelt ergänzende Arbeiten zu den Themen der Vorlesungen und befasst sich mit aktuellen wissenschaftlichen Abhandlungen aus der Parallelverarbeitung.

**Teilnahmevoraussetzungen**

Teilnahme am Modul "Algorithmen und Datenstrukturen 1" (10-201-2001-1) oder gleichwertige Kenntnisse.  
Die Module 10-201-2219S und -2219V schließen sich gegenseitig aus.

**Literaturangabe**

Hinweise zu Literaturangaben erfolgen in den Lehrveranstaltungen.

**Vergabe von Leistungspunkten**

Leistungspunkte werden mit erfolgreichem Abschluss des Moduls vergeben.  
Näheres regelt die Prüfungsordnung.

**Prüfungsleistungen und -vorleistungen**

<b>Modulprüfung: Mündliche Prüfung 20 Min., mit Wichtung: 1</b>	
	Vorlesung "Grundlagen der Parallelverarbeitung" (2SWS)
Referat (45 Min.) mit schriftlicher Ausarbeitung (2 Wochen), mit Wichtung: 1	Seminar "Grundlagen der Parallelverarbeitung" (2SWS)

## Master of Science Data Science

Akademischer Grad	Modulnummer	Modulform
Master of Science	10-201-2219V	Wahlpflicht

### Modultitel Grundlagen der Parallelverarbeitung (V)

**Modultitel (englisch)** Foundations of Parallel Processing (V)

**Empfohlen für:** 3. Semester

**Verantwortlich** Professur für Schwarmintelligenz und Komplexe Systeme

**Dauer** 1 Semester

**Modulturnus** jedes Wintersemester

**Lehrformen**

- Vorlesung "Grundlagen der Parallelverarbeitung" (2 SWS) = 30 h Präsenzzeit und 45 h Selbststudium = 75 h
- Vorlesung mit seminaristischem Anteil "Grundlagen der Parallelverarbeitung 2" (2 SWS) = 30 h Präsenzzeit und 45 h Selbststudium = 75 h

**Arbeitsaufwand** 5 LP = 150 Arbeitsstunden (Workload)

**Verwendbarkeit** • M.Sc. Data Science

**Ziele** Nach der aktiven Teilnahme am Modul „Grundlagen der Parallelverarbeitung“ sind die Studierenden in Lage:

- grundlegende Begriffe und Konzepte der Parallelverarbeitung zu formulieren und zu erklären,
- grundlegende parallele algorithmische Verfahren und Rechnermodelle (u.a. Sortieralgorithmen, Hardware- Addition) zu analysieren und zu vergleichen und
- für grundlegende algorithmische Probleme selbständig parallele Lösungsverfahren zu entwerfen.

**Inhalt** Studierende wählen die grundlegende Vorlesung "Parallele Algorithmen" und eine weiterführende Vorlesung.

Parallele Algorithmen: Grundlegende Konzepte und Bewertungskriterien für parallele Algorithmen, PRAM-Modell, Parallele Algorithmen für grundlegende Probleme wie Sortieren oder Mergen, Grundlagen von Hardware Algorithmen.

Weiterführende Vorlesungen

1. Parallele Berechnungsmodelle: Grundlegender Aufbau von Parallelrechnern, Einführung in realistische Parallelerechnermodelle, Varianten des BSP-Modells, Varianten des LogP-Modells, Auswirkungen der Modelle auf den Entwurf von Algorithmen, Algorithmische Lösung von Beispielproblemen.

2. Rekonfigurierbare Rechensysteme: Einsatzbereiche rekonfigurierbarer Rechensysteme, Typen rekonfigurierbarer Rechensysteme, Aufbau von Field Programmable Gate Arrays (FPGAs), Theoretische Konzepte der Rekonfigurierbarkeit, Grundlegende Algorithmen zu dynamischer Rekonfiguration

In unregelmäßigen Abständen wird die grundlegende Vorlesung durch die folgende Vorlesung ersetzt:

Entwurf und Implementierung paralleler Algorithmen: Parallele Plattformen, Entwurfsprinzipien, Analytische Modellierung, Parallele Programmierung für

nachrichtengekoppelte und speichergekoppelte Parallelrechner, Matrixmultiplikation, Sortieren, Graphenalgorithmen, Diskrete Optimierung, Dynamische Programmierung.

**Teilnahmevoraussetzungen** Teilnahme am Modul "Algorithmen und Datenstrukturen 1" (10-201-2001-1) oder gleichwertige Kenntnisse.  
Die Module 10-201-2219S und -2219V schließen sich gegenseitig aus.

**Literaturangabe** Hinweise zu Literaturangaben erfolgen in den Lehrveranstaltungen.

**Vergabe von Leistungspunkten** Leistungspunkte werden mit erfolgreichem Abschluss des Moduls vergeben.  
Näheres regelt die Prüfungsordnung.

### **Prüfungsleistungen und -vorleistungen**

<b>Modulprüfung: Mündliche Prüfung 20 Min., mit Wichtung: 1</b>	
	Vorlesung "Grundlagen der Parallelverarbeitung" (2SWS)
	Vorlesung mit seminaristischem Anteil "Grundlagen der Parallelverarbeitung 2" (2SWS)

## Master of Science Data Science

Akademischer Grad	Modulnummer	Modulform
Master of Science	10-INF-DS103	Wahlpflicht

### Modultitel **Praktikum für Data Warehousing und Data Mining**

**Modultitel (englisch)** Data Warehousing and Data Mining Project

**Empfohlen für:** 3. Semester

**Verantwortlich** Abteilung Datenbanken

**Dauer** 1 Semester

**Modulturnus** jedes Wintersemester

**Lehrformen** • Praktikum "Praktikum für Data Warehousing und Data Mining" (2 SWS) = 30 h Präsenzzeit und 120 h Selbststudium = 150 h

**Arbeitsaufwand** 5 LP = 150 Arbeitsstunden (Workload)

**Verwendbarkeit** • M.Sc. Data Science

**Ziele** Nach der aktiven Teilnahme am Modul "Praktikum für Data Warehousing und Data Mining" sind die Studierenden in der Lage, aktuelle Datenbank-/Big-Data-Technologien anzuwenden.

Sie können selbständig Anfragen formulieren und ausführen.

Die Studierenden sind imstande, Integrationsverfahren sowie Techniken zur Anfragebearbeitung und -optimierung durchzuführen.

Sie sind in der Lage, Algorithmen zur effizienten Verarbeitung großer Datenmengen anzuwenden bzw. zu implementieren.

**Inhalt**

- Praktische Realisierung eines Data-Warehousing- und Data-Mining-Projektes
- Datenimport und -bereinigung (Objektkonsolidierung, Normalisierung, ...)
- Definition und Erstellung eines Data Cubes und OLAP-Analysen
- Anwendung und Implementierung verschiedener Data Mining Algorithmen unter Verwendung von existierenden Big Data Frameworks und Technologien

Je nach Interesse des Studierenden kann der Schwerpunkt des Praktikums entweder auf Data Warehousing oder auf Data Mining gelegt werden.

**Teilnahmevoraussetzungen** keine

**Literaturangabe** Hinweise zu Literaturangaben erfolgen in den Lehrveranstaltungen.

**Vergabe von Leistungspunkten** Leistungspunkte werden mit erfolgreichem Abschluss des Moduls vergeben. Näheres regelt die Prüfungsordnung.

**Prüfungsleistungen und -vorleistungen**

<b>Modulprüfung: Praktikumsleistung (3 Testate a 60 Min.), mit Wichtung: 1</b>	
	Praktikum "Praktikum für Data Warehousing und Data Mining" (2SWS)



## Master of Science Data Science

Akademischer Grad	Modulnummer	Modulform
Master of Science	10-INF-DS02	Pflicht

### Modultitel **Masterseminar Data Science**

**Modultitel (englisch)** Master's Seminar: Data Science

**Empfohlen für:** 4. Semester

**Verantwortlich** Studiengangsverantwortlicher Master of Science Data Science

**Dauer** 1 Semester

**Modulturnus** jedes Semester

**Lehrformen** • Seminar "Masterseminar Data Science" (2 SWS) = 30 h Präsenzzeit und 120 h Selbststudium = 150 h

**Arbeitsaufwand** 5 LP = 150 Arbeitsstunden (Workload)

**Verwendbarkeit** • M.Sc. Data Science

**Ziele** Nach der aktiven Teilnahme am Modul "Masterseminar Data Science" sind die Studierenden in der Lage sich selbstständig in aktuelle Forschungsthemen der Data Science einzuarbeiten.

Sie können selbstständig die zugehörige Literatur auswählen und analysieren, sowie das Thema angemessen in einer Gruppe präsentieren.

**Inhalt** In jedem Semester bieten mehrere Abteilungen Absolventenseminar an, das im Rahmen des Masterseminars belegt werden kann. Die Seminare behandeln aktuelle Forschungsthemen aus einem Forschungsgebiet der jeweiligen Abteilung.

**Teilnahmevoraussetzungen** keine

**Literaturangabe** Hinweise zu Literaturangaben erfolgen in den Lehrveranstaltungen.

**Vergabe von Leistungspunkten** Leistungspunkte werden mit erfolgreichem Abschluss des Moduls vergeben. Näheres regelt die Prüfungsordnung.

### Prüfungsleistungen und -vorleistungen

<b>Modulprüfung: Referat 60 Min., mit Wichtung: 1</b>	
	Seminar "Masterseminar Data Science" (2SWS)