Georg-August-Universität Göttingen **Modul B.Inf.1247: Introduction to Information Retrieval and Natural Language Processing** *English title: Introduction to Information Retrieval and Natural Language Processing*	6 C 4 SWS

Lernziele/Kompetenzen: After successfully completing the course, students should be able to: • Summarize major IR and NLP applications • Explain important IR and NLP algorithms and data structures • Determine the conceptual requirements of specific IR and NLP problems • Compare the suitability of algorithms and data structures for specific tasks • Devise solutions for complex IR and NLP tasks by implementing and adapting suitable algorithms and data structures • Evaluate IR and NLP methods and systems quantitatively and qualitatively	**Arbeitsaufwand:** Präsenzzeit: 56 Stunden Selbststudium: 124 Stunden

Lehrveranstaltung: Lecture Introduction to Information Retrieval and Natural Language Processing (Vorlesung) *Inhalte*: The lecture will cover the following topics: • Basics: Background, Text Preprocessing, Documents, Terms, Vocabulary, Inverted Index • Boolean Retrieval, Positional Retrieval, Tolerant Retrieval • Efficient Index Construction, Index Compression • Term Weighting, Relevance Scoring, Ranked Retrieval • Semantic Text Analysis, Link Analysis • Complete Retrieval Systems • Results Visualization and Exploration • Evaluation of Retrieval Systems Please visit www.gipplab.org/teaching for details on this course.	2 SWS
Prüfung: Written test (90 min.) or oral exam (approx. 20 min.) **Prüfungsvorleistungen:** Successful completion of the examination in the practical course component of this module. **Prüfungsanforderungen:** • Knowledge of major IR and NLP applications • Ability to explain important IR and NLP algorithms and data structures • Ability to analyze the conceptual requirements of specific IR and NLP problems • Ability to compare the suitability of algorithms and data structures for specific tasks • Ability to evaluate IR and NLP methods and systems quantitatively and qualitatively	2 C

Lehrveranstaltung: Practical Course Introduction to Information Retrieval and Natural Language Processing (Laborpraktikum) *Inhalte*:	2 SWS

In the practical course, students work on applied research projects (teamwork is possible) that address complex information retrieval tasks. Using the programming language Python and presenting the intermediate and final results of the projects is mandatory. Please visit www.gipplab.org/teaching for details on this course.	
Prüfung: Präsentation (ca. 20 Minuten) **Prüfungsvorleistungen:** Successful completion of an applied research project including at least one intermediate milestone or presentation. **Prüfungsanforderungen:** • Ability to analyze the conceptual requirements of specific IR and NLP problems • Ability to compare the suitability of algorithms and data structures for specific tasks • Ability to determine the conceptual requirements of specific IR and NLP problems • Ability to devise solutions for complex IR and NLP tasks by implementing and adapting suitable algorithms • Ability to evaluate IR and NLP methods and systems quantitatively and qualitatively	4 C

Zugangsvoraussetzungen: keine	**Empfohlene Vorkenntnisse:** Knowledge of at least one object-oriented programming language, preferably Python, is required to complete the course. Python is used as part of the exercise sessions. For participants who are unfamiliar with Python, a fast-paced introduction into the essentials of the language will be provided.
Sprache: Englisch	**Modulverantwortliche[r]:** Prof. Dr. Bela Gipp
Angebotshäufigkeit: irregular	**Dauer:** 1 Semester
Wiederholbarkeit: zweimalig	**Empfohlenes Fachsemester:**
Maximale Studierendenzahl: 30	

Bemerkungen:
This course provides a good foundation for a bachelor's or master's thesis in our group. Visit www.gipplab.org/students-corner/graduation-projects for our current theses proposals.