



Modulkatalog

Computerlinguistik und Sprachtechnologie

Programmformat: Minor 60

Studienstufe: Bachelor

Gültig ab: Herbstsemester 2019

[Erstellt am 14.04.2022]

Modulgruppen des Programms

Einführung in die Computerlinguistik

Wissenschaftliche Vertiefung

Kernbereich Computerlinguistik und Sprachtechnologie

Informatik

Praxis der Computerlinguistik und Sprachtechnologie

Der Modulkatalog enthält alle beständigen Angaben zu den Pflicht- und Wahlpflichtmodulen des Programms, die semesterbezogenen Informationen dazu entnehmen Sie dem Vorlesungsverzeichnis.

Ebenfalls im Vorlesungsverzeichnis finden Sie das aktuelle Angebot an Wahlmodulen sowie weiterführende Informationen zu Modulen anderer Fakultäten.



Einführung in die Computerlinguistik

06SM521-001	Einführung in die Computerlinguistik 1	7
06SM521-002	Einführung in die Computerlinguistik 2	8
06SM521-003	Mathematische Grundlagen der Computerlinguistik	9
06SM521-004	Programmiertechniken der Computerlinguistik 1	10
06SM521-005	Programmiertechniken der Computerlinguistik 2	11



Wissenschaftliche Vertiefung

06SM521-s04 [Seminar]

12



Kernbereich Computerlinguistik und Sprachtechnologie

06SM521-009	Maschinelles Lernen in der Computerlinguistik	13
06SM521-006	Erstellung und Annotation von linguistischen Ressourcen	14
06SM521-007	Grammatikformalismen und Parsing	15
06SM521-008	Maschinelle Übersetzung	16
06SM521-010	Phonetic Data Analysis	17
06SM521-011	Semantic Web & Linked Data	18
06SM521-012	Sprachtechnologie als Beitrag zur Barrierefreiheit	19
06SM521-013	Sprachtechnologische Webapplikationen	20
06SM521-014	Text Analytics in the Digital Humanities	21
06SM521-015	Text Mining	22
06SM521-s10	[Sommerschule]	23



Informatik

Diese Modulgruppe enthält ausschliesslich Wahlmodule. Informieren Sie sich im Vorlesungsverzeichnis über das aktuelle Angebot.



Praxis der Computerlinguistik und Sprachtechnologie

06SM521-016	Praktikum extern	24
06SM521-017	Praktikum intern	25
06SM521-018	Programmierprojekt	26
06SM521-019	Übungsleitung/Tutorat	27
06SM521-021	Übungsleitung/Tutorat 2	28
06SM521-022	Programmierprojekt 2	29



Einführung in die Computerlinguistik 1

06SM521-001

Anbietende Organisationseinheit PhF: Institut für Computerlinguistik

ECTS Credits 6

Angebotsmuster 1-semesterig, jedes Herbstsemester

Bewertung/Benotung 1-6, in Halbschritten

Repetierbarkeit einmal wiederholbar, erneut buchen

Leistungsnachweis Portfolio (75% schriftliche Prüfung und 25% Nachweis von im Selbststudium erbrachten Studienleistung)

Unterrichtssprache Deutsch und/oder Englisch

Lehrformen Vorlesung mit integrierter Übung, Tutorat

Lernziel

Die Studierenden (1) kennen die grundlegende Thematik von Computerlinguistik, Sprachtechnologie und Linguistik sowie ihre interdisziplinäre Verknüpfung (2) verstehen die wichtigen linguistischen und formalen Begriffe und können sie abgrenzen und anwenden (3) ermitteln detaillierte Syntaxanalysen für deutsche Sätze, verstehen Grammatikregeln für Englisch und entwickeln diese weiter (4) kennen Zweck, Probleme und Lösungsansätze der textuellen Vorverarbeitung, automatischen Wortartenbestimmung, flachen Syntaxanalyse, Wortformenanalyse, Eigennamenerkennung, Volltextsuche, Informationsextraktion, Textklassifikation, Sprachsynthese, Spracherkennung, Maschineller Übersetzung (5) erkennen die für sprachtechnologische Anwendungen benötigten computerlinguistischen Techniken (6) können quantitative Evaluierungen interpretieren und anwenden

Allgemeine Beschreibung

Dieser Kurs vermittelt die wichtigsten praktischen und theoretischen Grundlagen der Computerlinguistik. Vorgestellt werden beispielhafte Systeme und Anwendungen wie maschinelle Übersetzung, Sprachsynthese und Spracherkennung, Textsuche, Informationsextraktion, Wortartenbestimmung und Textsegmentierung, sowie Methoden und Probleme der automatischen Syntaxanalyse von Sprachen. Im Kurs wird das notwendige linguistische Wissen erarbeitet.

Voraussetzungen

-



Einführung in die Computerlinguistik 2

06SM521-002

Anbietende Organisationseinheit

PhF: Institut für Computerlinguistik

ECTS Credits

3

Angebotsmuster

1-semesterig, jedes Frühlingsemester

Bewertung/Benotung

1-6, in Halbschritten

Repetierbarkeit

einmal wiederholbar, erneut buchen

Leistungsnachweis

Portfolio (75% schriftliche Prüfung und 25% Nachweis von im Selbststudium erbrachten Studienleistungen)

Unterrichtssprache

Deutsch und/oder Englisch

Lehrformen

Vorlesung mit integrierter Übung, Tutorat

Lernziel

Die Studierenden (1) kennen die Anwendungs- und Problembereiche der Semantik innerhalb der Computerlinguistik: lexikalische Semantik, insbesondere Wortnetze, Framesemantik und Satzsemantik (2) kennen Logik als Basis der Satzsemantik (3) gewinnen Einblick in weitere Probleme beim automatischen Verarbeiten der natürlichen Sprache (4) kennen computerlinguistische Verfahren zur Berechnung von semantischen Strukturen und Bedeutungsdesambiguierung (5) erarbeiten praktische Fähigkeiten für die Modellierung von semantischen Beziehungen mit Wortnetzen, für die einfache semantische Annotation von Sätzen in FrameNet und für die Übersetzung von Sätzen in logische Formen (6) erlangen Kenntnisse über die Modellierung von Wissen, die Grundlagen des Semantic Web und Basiskenntnisse in XML

Allgemeine Beschreibung

Die Vorlesung beschäftigt sich mit der Semantik der natürlichen Sprache, d.i. die lexikalische Semantik und die Satzsemantik. Wir beginnen mit dem Problem der Ambiguität, diskutieren die Grundlagen der lexikalischen Semantik und widmen uns dann den sog. Sinnrelationen (lexikalische Funktionen), die in den Wortnetzen zur Strukturierung des Wortschatzes eingesetzt werden. Wir werden ein probabilistisches Verfahren zur Bedeutungsdesambiguierung kennenlernen. Danach behandeln wir die Aussagen- und Prädikatenlogik und untersuchen, wie man Sätze mittels Prädikatenlogik semantisch repräsentieren kann. Schliesslich befassen wir uns mit dem Semantic Web, also dem Versuch, Internetseiten mit semantischer Information anzureichern, um die Informationssuche flexibler zu gestalten.

Voraussetzungen

Keine, jedoch Modul «Einführung in die Computerlinguistik 1» empfohlen.



Mathematische Grundlagen der Computerlinguistik

06SM521-003

Anbietende Organisationseinheit PhF: Institut für Computerlinguistik

ECTS Credits 6

Angebotsmuster 1-semesterig, jedes Frühlingsemester

Bewertung/Benotung 1-6, in Halbschritten

Repetierbarkeit einmal wiederholbar, erneut buchen

Leistungsnachweis Portfolio (90% schriftliche Prüfung und 10% Nachweis von im Selbststudium erbrachten Studienleistungen)

Unterrichtssprache Deutsch und/oder Englisch

Lehrformen Vorlesung mit integrierter Übung, Tutorat

Lernziel

Die Studierenden kennen die mathematischen Grundlagen für die quantitative Sprachverarbeitung; dazu gehören Wahrscheinlichkeitstheorie, Kombinatorik, Hypothesentesten und wichtige Konzepte der Analysis und der linearen Algebra. Ausserdem kennen sie die grundlegenden Verfahren des maschinellen Lernens und können diese auf ausgewählte Probleme der Sprachverarbeitung anwenden. Grundlage dafür ist eine Programmbibliothek wie z.B. sklearn.

Allgemeine Beschreibung

Dieser Kurs behandelt die mathematischen Grundlagen der Computerlinguistik. Es werden die zentralen Konzepte der Wahrscheinlichkeitstheorie und Kombinatorik eingeführt. Ausserdem behandeln wir Fragestellungen der Statistik (Deskriptive Statistik, Inferenzstatistik) und der linearen Algebra. Der Kurs vermittelt die theoretischen Grundlagen und dient zudem der praktischen Einübung und Anwendung der numerischen Verfahren. Der Kurs gibt ausserdem eine Einführung in traditionelle maschinelle Lernverfahren wie z.B. Logistische Regression oder das Perceptron und deren mathematisches Hintergrund.

Voraussetzungen

Programmierkenntnisse im Umfang von «Programmiertechniken in der Computerlinguistik 1».



Programmiertechniken der Computerlinguistik 1

06SM521-004

Anbietende Organisationseinheit PhF: Institut für Computerlinguistik

ECTS Credits 9

Angebotsmuster 1-semesterig, jedes Herbstsemester

Bewertung/Benotung 1-6, in Halbschritten

Repetierbarkeit einmal wiederholbar, erneut buchen

Leistungsnachweis Portfolio (75% Schriftliche Prüfung und 25% Nachweis von im Selbststudium erbrachten Studienleistungen)

Unterrichtssprache Deutsch und/oder Englisch

Lehrformen Vorlesung mit integrierter Übung, Tutorat

Lernziel

Die Studierenden (1) können mit Unix-basierten Systemen umgehen und sind fähig, die Unix-Textverarbeitungstools effizient einzusetzen (2) kennen die grundlegenden Datentypen, Kontrollstrukturen und Funktionen von Python (3) konzipieren und implementieren einfache Problemlösungen in Python (4) wenden grundlegende Werkzeuge für korpuslinguistische Fragestellungen an

Allgemeine Beschreibung

Der Kurs beginnt mit einer Einführung in Unix, speziell in dessen Textverarbeitungstools. Danach folgt eine Einführung in die Programmiersprache Python. Vermittelt werden grundlegende Operatoren und Funktionen, der Umgang mit Listen und Dictionaries sowie ein grundlegendes Verständnis für objektorientierte Programmierung. Es wird insbesondere Wert darauf gelegt, dass die Studierenden die Fähigkeit erwerben, Texte für die weitere Verarbeitung aufzubereiten. In regelmässigen Übungen soll das im Kurs vermittelte Wissen zur praktischen Programmierfähigkeit der Teilnehmenden werden.

Voraussetzungen

Paralleler Besuch der Vorlesung «Einführung in die Computerlinguistik 1» empfohlen.



Programmiertechniken der Computerlinguistik 2

06SM521-005

Anbietende Organisationseinheit PhF: Institut für Computerlinguistik

ECTS Credits 6

Angebotsmuster 1-semesterig, jedes Frühlingsemester

Bewertung/Benotung 1-6, in Halbschritten

Repetierbarkeit einmal wiederholbar, erneut buchen

Leistungsnachweis Portfolio (75% schriftliche Prüfung und 25% Nachweis von im Selbststudium erbrachten Studienleistungen)

Unterrichtssprache Deutsch und/oder Englisch

Lehrformen Vorlesung mit integrierter Übung, Tutorat

Lernziel

Die Studierenden beherrschen die benötigten Programmiertechniken und -kenntnisse zur Lösung einfacher sprachtechnologischer Aufgaben, insbesondere objektorientiertes Programmieren, Speicher-/Laufzeitoptimierung im Umgang mit grossen Datenmengen und einfaches maschinelles Lernen.

Allgemeine Beschreibung

Ziel dieses Kurses ist die Vertiefung von Programmierkenntnissen in Python. Relevante Themen und Techniken (s. Lernziele) werden anhand von ausgewählten Anwendungen aus dem Bereich der Computerlinguistik vermittelt.

Voraussetzungen

Modul «Programmiertechniken in der Computerlinguistik 1» (PCL1) oder nachgewiesene Kenntnisse im gleichen Umfang; Kenntnisse im Umfang von «Einführung in die Computerlinguistik 1» (ECL1). Interessierte Studierende, die PCL1 oder ECL1 nicht absolviert/bestanden haben, müssen sich vor der Buchung bei der Studienberatung melden.



[Seminar]

06SM521-s04

Anbietende Organisationseinheit PhF: Institut für Computerlinguistik

ECTS Credits 6

Angebotsmuster 1-semesterig, einmalig

Bewertung/Benotung 1-6, in Halbschritten

Repetierbarkeit keine Wiederholungsmöglichkeit

Leistungsnachweis Portfolio (50% schriftliche Arbeit und 50% Referat)

Unterrichtssprache

Lehrformen Seminar

Lernziel

Die Studierenden (1) vertiefen ihr Wissen in einem spezifischen Gebiet (2) erlernen die methodischen Grundkompetenzen wissenschaftlicher Arbeit (3) üben das verständliche Präsentieren komplexer Sachverhalte (4) können eine wissenschaftliche Arbeit verfassen

Allgemeine Beschreibung

Das Seminar dient der wissenschaftlichen Bearbeitung eines spezifischen Themenbereichs. Dabei werden die Methoden wissenschaftlichen Arbeitens erlernt, z.B. das Rezipieren von Fachliteratur, die Informationssuche, das Strukturieren von Fakten und Theorien, sowie das eigenständige Beurteilen empirischer Befunde. Daneben lernen die Studierenden einen Vortrag zu konzipieren und zu halten. Sie lernen mit anderen über das Gehörte zu diskutieren und es zu bewerten. Schliesslich erwerben sie die Fähigkeit zur schriftlichen Ausarbeitung ihres Vortrags.

Voraussetzungen

-



Maschinelles Lernen in der Computerlinguistik

06SM521-009

Anbietende Organisationseinheit PhF: Institut für Computerlinguistik

ECTS Credits 6

Angebotsmuster 1-semesterig, jedes Herbstsemester

Bewertung/Benotung 1-6, in Halbschritten

Repetierbarkeit einmal wiederholbar, erneut buchen

Leistungsnachweis Portfolio (90% schriftliche Prüfung und 10% Nachweis von im Selbststudium erbrachten Studienleistungen)

Unterrichtssprache Deutsch und/oder Englisch

Lehrformen Vorlesung mit integrierter Übung, Tutorat

Lernziel

Die Studierenden verstehen die mathematischen Details der grundlegenden Ansätze im Bereich Deep Learning. Sie kennen die gängigsten Netzwerkarchitekturen und sind in der Lage, eigenständig Modelle zur Bearbeitung typischer sprachtechnologischer Probleme in einer auf Python basierenden Programmierumgebung zu realisieren.

Allgemeine Beschreibung

Der Kurs ist eine Einführung in das sogenannte Deep Learning: das Lernen von neuronalen Netzen. Es werden die mathematischen Grundlagen behandelt und verschiedene Ansätze wie RNN, LSTM etc. vorgestellt. Neben der theoretischen gibt es auch eine praktische Herangehensweise. Wir werden eine aktuelle Programmbibliothek (Framework) wie z.B. Pytorch verwenden, um prototypische Aufgabenstellungen bei der Verarbeitung menschlicher Sprache zu untersuchen.

Voraussetzungen

Erfolgreich bestandene Module «Programmiertechniken in der Computerlinguistik 1», «Programmiertechniken in der Computerlinguistik 2», «Mathematische Grundlagen der Computerlinguistik» oder nachgewiesene Kenntnisse im gleichen Umfang.



Erstellung und Annotation von linguistischen Ressourcen

06SM521-006

Anbietende Organisationseinheit PhF: Institut für Computerlinguistik

ECTS Credits 6

Angebotsmuster 1-semesterig, jedes Frühlingsemester

Bewertung/Benotung bestanden/nicht bestanden

Repetierbarkeit einmal wiederholbar, erneut buchen

Leistungsnachweis Nachweis von im Selbststudium erbrachten Studienleistungen

Unterrichtssprache Deutsch und/oder Englisch

Lehrformen Vorlesung mit integrierter Übung

Lernziel

Die Studierenden (1) kennen und beherrschen alle Teilschritte, welche für die nachhaltige Erstellung von sprachtechnologischen Ressourcen notwendig sind (2) kennen gängige Verfahren, um manuelle Annotation effizient und in guter Qualität zu erzeugen (3) produzieren gemeinsam eine annotierte sprachtechnologische Ressource, welche wissenschaftliche Standards erfüllt

Allgemeine Beschreibung

Manuell oder automatisch annotierte Ressourcen sind mittlerweile die Grundlage für viele computerlinguistische Systeme. Dieses Modul vermittelt die Grundlagen, Probleme und Best-Practice-Verfahren der Erstellung von Korpora und (ihrer) Annotation anhand eines praktischen Projekts zur Ressourcenerstellung, welche alle Teilschritte umfasst. Bezüglich Ressourcen werden Themen angesprochen wie Rohdatenbeschaffung (Crawling), automatische Datenaufbereitung, Verwendung von nachhaltigen XML-Standards als Austauschformat, Lizenzierung. Das Thema Annotation umfasst neben den klassischen automatischen computerlinguistischen Verfahren die Erstellung von Guidelines und Dokumentation, manuelle Annotation (Crowdsourcing) und Inter-Annotator-Übereinstimmung sowie effiziente halbautomatische Verfahren.

Voraussetzungen

Programmierkenntnisse



Grammatikformalismen und Parsing

06SM521-007

Anbietende Organisationseinheit PhF: Institut für Computerlinguistik

ECTS Credits 6

Angebotsmuster 1-semesterig, jedes Herbstsemester

Bewertung/Benotung 1-6, in Halbschritten

Repetierbarkeit einmal wiederholbar, erneut buchen

Leistungsnachweis Portfolio (90% schriftliche Prüfung und 10% Nachweis von im Selbststudium erbrachten Studienleistungen)

Unterrichtssprache Deutsch und/oder Englisch

Lehrformen Vorlesung mit integrierter Übung

Lernziel

Die Studierenden (1) kennen die Fragestellungen, Probleme und Lösungsansätze der Syntaxanalyse (2) verstehen ausgewählte Grammatikformalismen und können sie anwenden, sowie ausgewählte Grammatikphänomene des Deutschen modellieren (3) arbeiten sich ein in verschiedene Ansätze zum (probabilistischen) Parsing und ihrem Training anhand von Baumbanken

Allgemeine Beschreibung

Der Kurs behandelt die automatische Syntaxanalyse. Dazu werden die zentralen Begriffe der syntaktischen Analyse eingeführt, es werden verschiedene Grammatikformalismen behandelt und anhand einer ausgewählten Theorie wird Grammatikmodellierung betrieben. Daneben widmen wir uns den probabilistischen Verfahren der Syntaxanalyse (Parsing), ihren Grundlagen, ihrem Training anhand von Baumbanken und ihrer Evaluierung.

Voraussetzungen

Programmierkenntnisse im Umfang von «Programmiertechniken in der Computerlinguistik 1» oder nachgewiesene Kenntnisse im gleichen Umfang.



Maschinelle Übersetzung

06SM521-008

Anbietende Organisationseinheit PhF: Institut für Computerlinguistik

ECTS Credits 6

Angebotsmuster 1-semesterig, jedes Frühlingsemester

Bewertung/Benotung 1-6, in Halbschritten

Repetierbarkeit einmal wiederholbar, erneut buchen

Leistungsnachweis Portfolio (75% schriftliche Prüfung und 25% Nachweis von im Selbststudium erbrachten Studienleistungen)

Unterrichtssprache Deutsch und/oder Englisch

Lehrformen Vorlesung mit integrierter Übung

Lernziel

Die Studierenden gewinnen praktische Erfahrungen und ein vertieftes Verständnis im Bereich der Maschinellen Übersetzung.

Allgemeine Beschreibung

Der Kurs besteht aus Vorlesungen und praktischen Übungen in Maschinellem Übersetzung. Die Teilnehmer bauen und evaluieren MÜ-Systeme über unterschiedlichen Korpora. Themenbeispiele: (1) Einführung in die maschinelle Übersetzung (2) Parallele Korpora und Alignierung (3) Aktuelle Methoden der Maschinellen Übersetzung (4) Hybride maschinelle Übersetzungssysteme (5) Gesellschaftliche Auswirkungen der Maschinellen Übersetzung

Voraussetzungen

Erfolgreich bestandene Module «Mathematische Grundlagen der Computerlinguistik» und «Maschinelles Lernen in der Computerlinguistik» oder nachgewiesene Kenntnisse im gleichen Umfang; Erfolgreich bestandenes Modul «Programmiertechniken in der Computerlinguistik 1» oder nachgewiesene Kenntnisse im gleichen Umfang; Modul «Einführung in die Computerlinguistik 1» empfohlen.



Phonetic Data Analysis

06SM521-010

Anbietende Organisationseinheit PhF: Institut für Computerlinguistik

ECTS Credits 6

Angebotsmuster 2-semesterig, Beginn im Herbstsemester

Bewertung/Benotung 1-6, in Halbschritten

Repetierbarkeit einmal wiederholbar, erneut buchen

Leistungsnachweis Portfolio (Referat 1/3, schriftliche Arbeit 2/3)

Unterrichtssprache Englisch

Lehrformen Seminar

Lernziel

Students know how to make high-quality audio recordings for phonetic research purposes and how to organize them efficiently for analysis. They can also annotate them manually and use automatic tools for segmentation. They perform reliable measurements (formants, pitch, intensity, etc.) and produce meaningful visualizations (wave forms, spectra, spectrograms, etc.) with suitable software. They also understand how to read spectrograms so as to draw informed conclusions about the temporal and spectral characteristics of speech events. Moreover, students understand the most important key notions and concepts in automatizing measurements and in making them replicable (scripting).

Allgemeine Beschreibung

Research on language increasingly relies on the analysis of large databases and a large number of such corpora already exist, in particular for English. In spoken language, data analysis is more complicated compared to written language as the speech signal is a highly variable continuous stream of events which first needs to be processed to arrive at the underlying linguistic units. For this reason, phoneticians regularly resort to the aid of technical devices in order to record, file, edit and analyze large speech corpora. In this module we look at the technical side of phonetic database compilation, processing and analysis. The students acquire and develop skills and techniques necessary to carry out recordings of speech, organize and edit them efficiently and learn how to analyze them. We will look at sound recording equipment and software solutions geared specifically towards the need of phonetic processing and analysis.

Voraussetzungen

Students are required to have passed an introductory module in linguistics at bachelor level.



Semantic Web & Linked Data

06SM521-011

Anbietende Organisationseinheit PhF: Institut für Computerlinguistik

ECTS Credits 6

Angebotsmuster 1-semesterig, jedes 2. Frühlingsemester

Bewertung/Benotung 1-6, in Halbschritten

Repetierbarkeit einmal wiederholbar, erneut buchen

Leistungsnachweis Portfolio (60% Schriftliche Prüfung und 40% Nachweis von im Selbststudium erbrachten Studienleistungen)

Unterrichtssprache Deutsch und/oder Englisch

Lehrformen Vorlesung mit integrierter Übung

Lernziel

The students (1) understand the fundamental ideas of knowledge representation (entities, concepts, instances, ontologies, etc) and are aware of the historical approaches (e.g. Conceptual Graphs) (2) are very familiar with the Semantic Web standards (RDF, SPARQL, OWL) and are able to apply them for the construction and querying of Knowledge Bases (3) understand what Linked Data is, and are able to make use of the concepts and goals of the Linked Open Data movement

Allgemeine Beschreibung

A formalized computable representation of human knowledge has long been a goal of data science. Several approaches have been proposed since the dawn of computer science and artificial intelligence. In this course we will revise some of the historical approaches (First Order Logic, Semantic Networks, Frames, Conceptual Graphs, etc.) and then introduce the most recent approaches based on the semantic web standards (RDF, SPARQL, OWL). We will then describe how these standards can be used to achieve the goals of a distributed shared knowledge infrastructure (Linked Open Data).

Voraussetzungen

Erfolgreich bestandenes Modul «Programmiertechniken in der Computerlinguistik 1» oder nachgewiesene Kenntnisse im gleichen Umfang.



Sprachtechnologie als Beitrag zur Barrierefreiheit

06SM521-012

Anbietende Organisationseinheit PhF: Institut für Computerlinguistik

ECTS Credits 6

Angebotsmuster 1-semesterig, jedes Frühlingssemester

Bewertung/Benotung 1-6, in Halbschritten

Repetierbarkeit einmal wiederholbar, erneut buchen

Leistungsnachweis schriftliche Prüfung

Unterrichtssprache Deutsch und/oder Englisch

Lehrformen Vorlesung mit integrierter Übung

Lernziel

Die Studierenden (1) kennen die verschiedenen Zielgruppen von Sprachtechnologiesystemen im Kontext von Barrierefreiheit (2) können verschiedene Anwendungen von Sprachtechnologiesystemen im Kontext von Barrierefreiheit beschreiben (3) können den sprachtechnologischen Gehalt dieser Systeme genauer erläutern

Allgemeine Beschreibung

Gehörlose und hörbehinderte, blinde und sehbehinderte Menschen und Menschen mit kognitiven Beeinträchtigungen stehen im Alltag immer wieder vor Hindernissen. Um diese Hindernisse abzubauen, werden laufend neue Werkzeuge entwickelt. Die Vorlesung beschäftigt sich mit Sprachtechnologie-Werkzeugen, die einen Beitrag zur Barrierefreiheit leisten: Wir untersuchen unter anderem die Funktionsweise von Screen-Readern, Systemen zur Untertitelung für Hörbehinderte (Respeaking mit Spracherkennung), Systemen zur Audiodeskription, Gebärdensprach-Avataren, Systemen zur automatischen Übersetzung in Gebärdensprache, Systemen zur automatischen Textvereinfachung (Herstellung von leichter Sprache) und Systemen zur Konvertierung in Braille.

Voraussetzungen

Kenntnisse im Umfang der Lehrveranstaltung "Einführung in die Computerlinguistik 1".



Sprachtechnologische Webapplikationen

06SM521-013

Anbietende Organisationseinheit PhF: Institut für Computerlinguistik

ECTS Credits 9

Angebotsmuster 1-semesterig, jedes Herbstsemester

Bewertung/Benotung 1-6, in Halbschritten

Repetierbarkeit einmal wiederholbar, erneut buchen

Leistungsnachweis Portfolio (1/3 schriftliche Prüfung [Take-Home-Exam], 2/3 Nachweis von im Selbststudium erbrachten Studienleistungen)

Unterrichtssprache Deutsch und/oder Englisch

Lehrformen Vorlesung mit integrierter Übung, Tutorat

Lernziel

Die Studierenden (1) können verschiedene Arten der Daten- und Softwaremodellierung interpretieren, erläutern und selbst anwenden (2) sind in der Lage, Datenstrukturen in Datenbankschemata zu übertragen, diese mit Daten zu füllen und einfache Anfragen an die Datenbank zu formulieren, sowie die Resultate mittels Webservice-Schnittstelle in einer Webapplikation darzustellen (3) kennen grundlegende Konzepte der Datenvisualisierung und können sie anwenden (4) erwerben die Fähigkeit im Rahmen eines Programmierprojektes die anfallenden Arbeitsschritte zu planen, die Aufgaben untereinander aufzuteilen und die Ergebnisse zu integrieren (5) sind imstande, ein komplexes System verständlich zu präsentieren und dokumentieren

Allgemeine Beschreibung

Die Veranstaltung widmet sich Datenstrukturen und Benutzerschnittstellen für computerlinguistische Anwendungen. Einen Schwerpunkt bildet die Datenmodellierung und die Abbildung dieser Modelle in Datenbankschemata, einen anderen die Konzipierung und Implementierung von interaktiven Webanwendungen und die Visualisierung von Daten. Die Teilnehmer erstellen während des Kurses ein Programmierprojekt in Gruppenarbeit und präsentieren dieses am Ende des Semesters.

Voraussetzungen

Erfolgreich bestandenes Modul «Programmiertechniken in der Computerlinguistik 1» oder nachgewiesene Kenntnisse im gleichen Umfang; ausserdem empfohlen: «Programmiertechniken in der Computerlinguistik 2».



Text Analytics in the Digital Humanities

06SM521-014

Anbietende Organisationseinheit PhF: Institut für Computerlinguistik

ECTS Credits 6

Angebotsmuster 1-semesterig, jedes Frühlingsemester

Bewertung/Benotung 1-6, in Halbschritten

Repetierbarkeit einmal wiederholbar, erneut buchen

Leistungsnachweis dokumentierte praktische Arbeit

Unterrichtssprache Englisch

Lehrformen Seminar

Lernziel

Participants will (1) know the fundamental questions, problems, challenges and methods in Digital Humanities (2) be able to apply the methods to their own scientific questions and datasets (3) acquire an overview of existing resources.

Allgemeine Beschreibung

After an introduction to Text Analytics and Digital Humanities, we learn to program key methods ourselves and to use existing tools. The methods comprise concordancing, collocations, machine learning, topic modelling, keyword detection, named entity recognition, network analysis, stylometry, sentiment detection, visualization.

Voraussetzungen

Keine, jedoch Modul «Einführung in die Computerlinguistik 1» empfohlen.



Text Mining

06SM521-015

Anbietende Organisationseinheit

PhF: Institut für Computerlinguistik

ECTS Credits

6

Angebotsmuster

1-semesterig, jedes Frühjahrssemester

Bewertung/Benotung

1-6, in Halbschritten

Repetierbarkeit

einmal wiederholbar, erneut buchen

Leistungsnachweis

Portfolio (75% schriftliche Prüfung und 25% Nachweis von im Selbststudium erbrachten Studienleistungen). Die genaue notenrelevante Gewichtung der Einzelteile wird Anfang des Semesters durch die Dozierenden bekannt gegeben und kann von den oben genannten Werten abweichen.

Unterrichtssprache

Deutsch und/oder Englisch

Lehrformen

Vorlesung mit integrierter Übung

Lernziel

Die Studierenden (1) kennen die wichtigen Fragestellungen der Erschliessung von strukturierter Information aus natürlichsprachlichen Texten, die damit verbundenen Probleme und die computerlinguistischen Ansätze zu ihrer Lösung (2) kennen insbesondere die Teilgebiete der Sentimentanalyse und des automatischen Media-Monitoring und haben eine Übersicht, welche sprachtechnologischen Komponenten und Ressourcen benötigt werden (3) kennen verschiedene aktuelle Forschungsansätze für die Teilprobleme und gewinnen praktische Erfahrung mit eigener Programmierarbeit

Allgemeine Beschreibung

Das Modul führt ins Text-Mining ein. Es vermittelt dabei die relevanten sprachtechnologischen Ansätze für Sentimentanalyse und Media-Monitoring, zwei wichtige Anwendungsbereiche der Computerlinguistik. Sentimentanalyse, d.h. die Erkennung von positiven/negativen Bewertungen in klassischen oder sozialen Medien (insbesondere nutzergenerierte Inhalte wie Kommentare, Blogs, Kurznachrichten), ist zu einem zentralen Bereich der Informationsextraktion geworden. Sie ist eine zentrale Komponente von Media-Monitoring-Systemen, welche sich damit beschäftigen, welche Themen und Entitäten (Personen, Orte, Produkte) prominent sind. Das Modul beginnt mit den Fragestellungen, Methoden und Problemen. Anschliessend werden ausgewählte aktuelle Forschungsansätze vertiefter behandelt. Relevant sind dabei Themen wie lexikalische Ressourcen, Multilingualität, informelle Sprache, Named-Entity-Erkennung und Relationserkennung. Die Teilnehmenden implementieren in Übungen einzelne Aspekte.

Voraussetzungen

Erfolgreich bestandenes Modul «Programmiertechniken in der Computerlinguistik 2» oder nachgewiesene Kenntnisse im gleichen Umfang.



[Sommerschule]

06SM521-s10

Anbietende Organisationseinheit PhF: Institut für Computerlinguistik

ECTS Credits 6

Angebotsmuster 1-semesterig, einmalig

Bewertung/Benotung bestanden/nicht bestanden

Repetierbarkeit keine Wiederholungsmöglichkeit

Leistungsnachweis Nachweis von im Selbststudium erbrachten Studienleistungen

Unterrichtssprache

Lehrformen Selbststudium

Lernziel

Die Studierenden (1) repetieren und verfestigen Gelerntes (2) erarbeiten neue Inhalte/Themenbereiche in kompakter Form (3) lernen aktuelle Trends kennen (4) tauschen Erfahrungen aus mit Studierenden anderer Universitäten (5) machen Networking auf internationaler Ebene

Allgemeine Beschreibung

Sommerschulen dienen dazu, den Studierenden vertiefenden Einblick in spezifische computerlinguistische Themenbereiche zu geben. Sie festigen das bisher im Studium Gelernte, erweitern ihre Kenntnis der zentralen Inhalte und bieten ihnen in kompakter Weise Zugang zu neuen Ansätzen. Sie werden auf aktuelle Trends aufmerksam, tauschen Erfahrungen und Einschätzungen mit Studierenden anderer Universitäten aus, und bekommen auf diese Weise die Möglichkeit, das internationale Niveau kennenzulernen und gleichzeitig Beziehungen zu knüpfen, die über das Studium hinaus hilfreich sein können. Dieses Modul kann gebucht werden, um den Besuch von Sommerschulen im Bereich der Sprachverarbeitung anrechnen zu lassen. Dieses Modul kann mit einer Punktzahl von 3 oder 6 gebucht werden. Die Punktzahl wird in Absprache mit dem Modulverantwortlichen festgelegt.

Voraussetzungen

Dieses Modul kann nicht selbst gebucht werden, die Buchung des Moduls muss vom Modulverantwortlichen autorisiert werden. Bevor die Sommerschule besucht wird, muss die Anrechnung mit dem Modulverantwortlichen abgesprochen werden.



Praktikum extern

06SM521-016

Anbietende Organisationseinheit PhF: Institut für Computerlinguistik

ECTS Credits 6

Angebotsmuster 1-semesterig, jedes Semester

Bewertung/Benotung bestanden/nicht bestanden

Repetierbarkeit einmal wiederholbar, erneut buchen

Leistungsnachweis dokumentierte praktische Arbeit

Unterrichtssprache Deutsch und/oder Englisch

Lehrformen Praktikum

Lernziel

Die Studierenden (1) erhalten Einblick in die Arbeitswelt (2) arbeiten an der Verbindung von Theorie und Praxis (3) lernen die Abläufe in Firmen kennen (4) wenden Gelerntes an (5) erweitern ihr Wissen um praxisrelevante Fragestellungen

Allgemeine Beschreibung

Die Studierenden sammeln Erfahrung in der Praxis. Sie erhalten Einblick in die Abläufe in Betrieben, sind beteiligt an der Umsetzung computerlinguistischer Methoden zur Lösung firmenspezifischer Probleme. Sie wenden das im Studium Gelernte auf kommerzielle Fragestellungen an und passen es an die Anforderungen der Arbeitswelt an. Externe Praktika sind in der Regel Aufenthalte in Firmen oder öffentlichen Organisationen, welche auf dem Gebiet der Sprachtechnologie tätig sind. Die Praktika müssen fachbezogen sein und müssen selber organisiert werden. Dieses Modul kann mit einer Punktzahl von 3 oder 6 gebucht werden. Die Punktzahl wird in Absprache mit dem Modulverantwortlichen festgelegt.

Voraussetzungen

Dieses Modul kann nicht selbst gebucht werden, die Buchung des Moduls muss vom Modulverantwortlichen autorisiert werden. Damit externe Praktika angerechnet werden können, bedarf es der Bewilligung eines vorgängigen Gesuchs durch den Modulverantwortlichen. Externe Praktika sind für fortgeschrittene Studierende gedacht.



Praktikum intern

06SM521-017

Anbietende Organisationseinheit PhF: Institut für Computerlinguistik

ECTS Credits 6

Angebotsmuster 1-semesterig, jedes Semester

Bewertung/Benotung bestanden/nicht bestanden

Repetierbarkeit einmal wiederholbar, erneut buchen

Leistungsnachweis dokumentierte praktische Arbeit

Unterrichtssprache Deutsch und/oder Englisch

Lehrformen Praktikum

Lernziel

Die Studierenden (1) erhalten Einblick in die wissenschaftliche Projektarbeit (2) rezipieren Fachliteratur (3) sind beteiligt an Evaluations- und Revisionszyklen (4) übernehmen Teilarbeiten im Projekt (5) sind an Publikationsvorhaben beteiligt (6) gewinnen insgesamt Erfahrung in praktischer Arbeit (7) vertiefen ihr Wissen in einem Thema

Allgemeine Beschreibung

Das Modul ist gedacht als praktisches Kennenlernen von wissenschaftlicher Projektarbeit. Die Studierenden werden an wissenschaftliches Arbeiten herangeführt. Sie recherchieren, lesen Fachliteratur, bereiten Daten vor, annotieren, wenden statistische Methoden und Verfahren des Maschinellen Lernens an, sind beteiligt an Publikationsarbeiten und diskutieren Lösungsansätze. Die Studierenden arbeiten selbständig an einer Teilaufgabe im Kontext eines wissenschaftlichen Themas oder Projektes. Dieses Modul kann gebucht werden, um Mitarbeit bei wissenschaftlichen Projekten der UZH anrechnen zu lassen. Dieses Modul kann mit einer Punktzahl von 6 oder 9 gebucht werden. Die Punktzahl wird in Absprache mit dem Modulverantwortlichen festgelegt.

Voraussetzungen

Dieses Modul kann nicht selbst gebucht werden, die Buchung des Moduls muss vom Modulverantwortlichen autorisiert werden. Es besteht kein Anspruch auf ein internes Praktikum, sondern das Modul wird nur angeboten, wenn in einem Projekt eine entsprechende Möglichkeit besteht. Die Voraussetzungen werden dem Thema entsprechend festgelegt. In der Regel werden Programmierkenntnisse erwartet und/oder der erfolgreiche Besuch bestimmter Lehrveranstaltungen.



Programmierprojekt

06SM521-018

Anbietende Organisationseinheit PhF: Institut für Computerlinguistik

ECTS Credits 6

Angebotsmuster 1-semesterig, jedes Semester

Bewertung/Benotung bestanden/nicht bestanden

Repetierbarkeit einmal wiederholbar, erneut buchen

Leistungsnachweis dokumentierte praktische Arbeit

Unterrichtssprache Deutsch und/oder Englisch

Lehrformen Selbststudium

Lernziel

Die Studierenden (1) planen ein Projekt (2) setzen einen Projektplan um (3) realisieren Software (4) halten Dokumentationsstandards ein (5) evaluieren (6) nutzen Softwarerepositorien

Allgemeine Beschreibung

Programmierprojekte dienen der Festigung der Programmierkenntnisse und dem Erwerb von Fertigkeiten, wie sie im Kontext des Softwareengineering erforderlich sind. Ausgehend von einer computerlinguistischen Fragestellung und ggfs. relevanter Fachliteratur wird ein Lösungskonzept erarbeitet, es werden die Meilensteine fixiert, es werden Daten beschafft und/oder eigens annotiert, es wird ein Programm implementiert und schliesslich anhand der Daten evaluiert. Dieses Modul kann gebucht werden, um ein Programmierprojekt anrechnen zu lassen. Dieses Modul kann mit einer Punktzahl von 3, 6 oder 9 gebucht werden. Die Punktzahl wird in Absprache mit dem Modulverantwortlichen festgelegt.

Voraussetzungen

Dieses Modul kann nicht selbst gebucht werden, die Buchung des Moduls muss vom Modulverantwortlichen autorisiert werden. Bevor ein Programmierprojekt erarbeitet wird, muss zwingend Rücksprache mit dem Modulverantwortlichen genommen werden (per E-Mail). Die Voraussetzungen werden dem Thema entsprechend festgelegt.



Übungsleitung/Tutorat

06SM521-019

Anbietende Organisationseinheit PhF: Institut für Computerlinguistik

ECTS Credits 6

Angebotsmuster 1-semesterig, jedes Semester

Bewertung/Benotung bestanden/nicht bestanden

Repetierbarkeit einmal wiederholbar, erneut buchen

Leistungsnachweis dokumentierte praktische Arbeit

Unterrichtssprache Deutsch und/oder Englisch

Lehrformen Sonstiges

Lernziel

Die Studierenden (1) verstehen fachspezifische Inhalte aus einer Lehrperspektive (2) lernen Vorlesungsinhalte zielgruppengerecht aufzubereiten und als Übungsmaterial zu konzipieren (3) lernen zu korrigieren und Rückmeldung zu geben

Allgemeine Beschreibung

Ein Tutorat/Übungsleitung dient der Erlernung von ersten didaktischen Fertigkeiten. Dies setzt ein reflektiertes Verständnis fachspezifischer Inhalte voraus und erfordert die Aufbereitung von Lernmaterialien für das Repetieren der Inhalte aus der Vorlesung, sowie das Konzipieren und Korrigieren von Übungsblättern. Dieses Modul kann mit einer Punktzahl von 3 oder 6 gebucht werden. Die Punktzahl wird in Absprache mit dem Modulverantwortlichen festgelegt.

Voraussetzungen

Dieses Modul kann nicht selbst gebucht werden. Es muss zwingend Rücksprache mit dem Modulverantwortlichen genommen werden (per E-Mail). Die Dozierenden der Lehrveranstaltung müssen in den E-Mail-Verkehr einbezogen sein. Übungsleitungen werden normalerweise einige Wochen vor Semesterbeginn auf der Mailingliste der Computerlinguistik (cl-list@lists.ifi.uzh.ch) ausgeschrieben. Die Studierenden können sich auch gerne jederzeit spontan bei den Dozierenden bewerben, für deren Lehrveranstaltung sie eine Übungsleitung übernehmen möchten. Für Übungsleitungen/Tutorate muss das entsprechende Modul selbst erfolgreich bestanden sein.



Übungsleitung/Tutorat 2

06SM521-021

Anbietende Organisationseinheit PhF: Institut für Computerlinguistik

ECTS Credits 6

Angebotsmuster 1-semesterig, jedes Semester

Bewertung/Benotung bestanden/nicht bestanden

Repetierbarkeit einmal wiederholbar, erneut buchen

Leistungsnachweis dokumentierte praktische Arbeit

Unterrichtssprache Deutsch und/oder Englisch

Lehrformen Sonstiges

Lernziel

Die Studierenden (1) verstehen fachspezifische Inhalte aus einer Lehrperspektive (2) lernen Vorlesungsinhalte zielgruppengerecht aufzubereiten und als Übungsmaterial zu konzipieren (3) lernen zu korrigieren und Rückmeldung zu geben

Allgemeine Beschreibung

Dieses Modul kann gebucht werden, um eine zweite Übungsleitung anrechnen zu lassen. Ein Tutorat/Übungsleitung dient der Erlernung von ersten didaktischen Fertigkeiten. Dies setzt ein reflektiertes Verständnis fachspezifischer Inhalte voraus und erfordert die Aufbereitung von Lernmaterialien für das Repetieren der Inhalte aus der Vorlesung, sowie das Konzipieren und Korrigieren von Übungsblättern. Dieses Modul kann mit einer Punktzahl von 3 oder 6 gebucht werden. Die Punktzahl wird in Absprache mit dem Modulverantwortlichen festgelegt.

Voraussetzungen

Dieses Modul kann nicht selbst gebucht werden. Es muss zwingend Rücksprache mit dem Modulverantwortlichen genommen werden (per E-Mail). Die Dozierenden der Lehrveranstaltung müssen in den E-Mail-Verkehr einbezogen sein. Übungsleitungen werden normalerweise einige Wochen vor Semesterbeginn auf der Mailingliste der Computerlinguistik (cl-list@lists.ifi.uzh.ch) ausgeschrieben. Die Studierenden können sich auch gerne jederzeit spontan bei den Dozierenden bewerben, für deren Lehrveranstaltung sie eine Übungsleitung übernehmen möchten. Für Übungsleitungen/Tutorate muss das entsprechende Modul selbst erfolgreich bestanden sein.



Programmierprojekt 2

06SM521-022

Anbietende Organisationseinheit PhF: Institut für Computerlinguistik

ECTS Credits 6

Angebotsmuster 1-semesterig, jedes Semester

Bewertung/Benotung bestanden/nicht bestanden

Repetierbarkeit einmal wiederholbar, erneut buchen

Leistungsnachweis dokumentierte praktische Arbeit

Unterrichtssprache Deutsch und/oder Englisch

Lehrformen Selbststudium

Lernziel

Die Studierenden (1) planen ein Projekt (2) setzen einen Projektplan um (3) realisieren Software (4) halten Dokumentationsstandards ein (5) evaluieren (6) nutzen Softwarerepositorien

Allgemeine Beschreibung

Dieses Modul kann gebucht werden, um ein zweites Programmierprojekt anrechnen zu lassen. Programmierprojekte dienen der Festigung der Programmierkenntnisse und dem Erwerb von Fertigkeiten, wie sie im Kontext des Softwareengineering erforderlich sind. Ausgehend von einer computerlinguistischen Fragestellung und ggfs. relevanter Fachliteratur wird ein Lösungskonzept erarbeitet, es werden die Meilensteine fixiert, es werden Daten beschafft und/oder eigens annotiert, es wird ein Programm implementiert und schliesslich anhand der Daten evaluiert. Dieses Modul kann mit einer Punktzahl von 3, 6 oder 9 gebucht werden. Die Punktzahl wird in Absprache mit dem Modulverantwortlichen festgelegt.

Voraussetzungen

Dieses Modul kann nicht selbst gebucht werden, die Buchung des Moduls muss vom Modulverantwortlichen autorisiert werden. Bevor ein Programmierprojekt erarbeitet wird, muss zwingend Rücksprache mit dem Modulverantwortlichen genommen werden (per E-Mail). Die Voraussetzungen werden dem Thema entsprechend festgelegt.
