

Linguistische Informatik

Bachelor of Arts

1. Studienkonzept und Qualifikationsprofil

Die maschinelle Bearbeitung und Erschließung von geschriebenen und gesprochenen Texten sowie die Interaktion von Mensch und Computer in natürlicher Sprache konstituieren einen Objektbereich, der unter verschiedenen Perspektiven und mit verschiedenen Schwerpunktsetzungen untersucht werden kann. Entsprechend gibt es mehrere Bezeichnungen für dieses neue Gebiet: Computerlinguistik, Linguistische Datenverarbeitung, Informationslinguistik, Linguistische Informatik. Nach allgemeinem Konsens fußt die maschinelle Sprachverarbeitung auf dem Wissensstand der Linguistik und Informatik und umfasst außerdem einen Bereich eigener Theorien und Methoden. Eine erschöpfende Ausbildung auf allen drei Gebieten kann im Rahmen eines einzigen Studienganges kaum angeboten werden. Es gibt zwei grobe Zielvorstellungen für eine Differenzierung der Ausbildung auf dem Sektor maschinelle Sprachverarbeitung: den voll ausgebildeten Informatiker einerseits, der ein gewisses Verständnis für die Probleme der natürlichen Sprachen hat (Diplomstudiengang Informatik mit Anwendungsfach Linguistik), und den ausgesprochenen Sprachfachmann andererseits, der hinreichende Informatikkenntnisse hat, um bei einer Umsetzung der linguistischen Theorie in Programme mitwirken zu können (Computerlinguistik). Diese grobe Zweiteilung kann durch Bezug zu einem speziellen Anwendungsfeld (wie der Information und Dokumentation, der Lexikographie und Terminologiearbeit, der maschinellen Übersetzung, des computerunterstützten Unterrichts u.a.) weiter differenziert werden. Da die Kenntnisse des jeweils anderen Gebiets jedoch meist beschränkt bleiben, gibt es Transfer-Probleme, die dazu führen, dass optimale Lösungen von Problemen der Verarbeitung natürlicher Sprache ausbleiben. Der Linguist hat den Eindruck, dass der Informatiker die Phänomene der natürlichen Sprache nur als störend empfindet, während der Informatiker die Programmierversuche des Linguisten für obsolet erachtet. Wünschenswert wären daher Doppelqualifikationen.

Das Erlanger Modell

Eine solche Doppelqualifikation eröffnet ein Studium mit einer obligatorischen Kombination aus einem sprachwissenschaftlichen Fach und dem Fach „Linguistische Informatik“. Im ersten Fach soll der Student eine gründliche theoretische und empirische sprachwissenschaftliche Ausbildung erhalten. Das zweite Fach sollte etwa zur Hälfte aus Informatik und zur anderen Hälfte aus Computerlinguistik bestehen. Dadurch verleiht es dem Absolventen sowohl eine beachtliche Qualifikation in Informatik, die seinen Berufsaussichten förderlich sein wird, wie auch die Befähigung, die speziellen Sprachverarbeitungsaufgaben zu lösen. Für ein solches Modell ist die Situation in Erlangen günstig, da die Voraussetzungen für eine gute Ausbildung in Informatik bestehen und der Anteil an Computerlinguistik durch die Professur für Linguistische Informatik am Institut für Germanistik abgedeckt werden kann.

Berufsperspektiven

Die Berufsperspektiven, die ein Studium der Linguistischen Informatik eröffnet, sind auf dem Hintergrund der allgemein schlechten Beschäftigungssituation für Sprach- und Literaturwissenschaftler außerhalb von Schule und Hochschule zu sehen. Es ist wohl mit einiger Sicherheit vorherzusagen, dass die Informationserschließung mit Hilfe von EDV weiterhin voranschreiten wird und dass dabei die Erschließung von in natürli-

cher Sprache vorliegenden Quellen immer wichtiger wird. Es wird daher in Zukunft neue Bereiche geben, in denen die Verarbeitung von Sprachdaten eine Rolle spielt. Neben der Befähigung zur Pflege und Anpassung entsprechender Programme beim Anwender schafft ein Studium der Linguistischen Informatik die Qualifikation, an der Entwicklung von neuen Produkten mitzuwirken. Die Einrichtung entsprechender Studiengänge an den Universitäten wird also selbst dazu beitragen, dass die Berufsaussichten auf diesem Gebiet mit der Zeit immer besser werden. Gegenwärtig gibt es bereits in folgenden Bereichen einen Bedarf an Linguisten mit EDV-Ausbildung (bzw. ist ein Bedarf zu vermuten):

- an Hochschulen in Forschung und Lehre auf dem Gebiet linguistische Datenverarbeitung,
- in außeruniversitären Forschungseinrichtungen wie der Gesellschaft für Mathematik und Datenverarbeitung (GMD), in staatl. Förderungsgremien, in den Entwicklungsabteilungen bei Computerfirmen und Software-Häusern,
- im Bereich der Fachinformation und Dokumentation auf den Gebieten der Indexierung, der Referaterstellung und der Recherche,
- beim Aufbau von Datenbanken und Archiven (z.B. bei Parlamenten und Presseorganen),
- im Bereich der Texterstellung und Textredaktion bei Verlagen, Zeitungen sowie im Rahmen der Büroautomatisierung,
- bei Informationsanbietern im Rahmen der neuen Medien und des „electronic publishing“,
- auf dem Gebiet ein- und mehrsprachiger Lexikographie und der Terminologiearbeit (z.B. bei der EU),
- auf dem Gebiet der computerunterstützten Übersetzung (bei internat. Organisationen sowie in Übersetzungsbüros),
- bei der Entwicklung neuer Produkte (z.B. Wandlern von geschriebener in gesprochene Sprache und umgekehrt),
- bei der Entwicklung von computerunterstütztem Sprachunterricht.

Abschluss B.A. und Inhalt des Studiengangs

Der Studiengang vermittelt solide sprachwissenschaftliche Kenntnisse und Fähigkeiten in der Programmierung, die die Grundlage der Entwicklung neuer Anwendungen und Methoden der automatischen Sprachverarbeitung bilden. Durch die Auseinandersetzung mit verschiedenen Grammatiktheorien und ihrer praktischen Anwendung auf natürliche und formale Sprachen erhalten die Studierenden einen Überblick über den Stand von Theorie und Technik dieses Fachgebiets, der sie zu einem kritischen Umgang damit befähigt.

Der Studiengang Linguistische Informatik selbst beschränkt sich dabei auf die technischen und praktischen Aspekte speziell der maschinellen Sprachverarbeitung; um Redundanzen zu vermeiden und den unterschiedlichen Interessen der Studierenden an den bestimmten Sprachen Rechnung zu tragen, sollen die hierfür notwendigen Grundkenntnisse der theoretischen und empirischen Linguistik in einem in Kombination zu studierenden sprachwissenschaftlichen Fach erworben werden.

2. Struktur des Studiengangs

„Linguistische Informatik“ kann im modularisierten Zwei-Fach-Bachelorstudiengang als Erst- oder Zweifach (je 70 ECTS, plus Bachelorarbeit bei Studium als Erstfach) studiert werden. Nahelegend ist eine Kombination mit einem sprachwissenschaftli-

chen Fach, z.B.: Indogermanistik und Indoiranistik, Griechische Philologie, Lateinische Philologie, Mittel- und Neulatein, Sinologie, Japanologie, Frankoromanistik, Iberoromanistik, English and American Studies, Nordische Philologie, Germanistik und Orientalistik. Bei einer geplanten Kombination mit Italoromanistik ist nicht gewährleistet, dass die Regelstudienzeit eingehalten werden kann. Vor der Einschreibung ist eine Beratung bei der Studienberatung (IBZ oder Studien-Service-Center) notwendig. Bei Belegung als Erstfach kommen außerdem noch 30 ECTS-Punkte aus dem Bereich der Schlüsselqualifikationen hinzu. Voraussetzung für das Bestehen der **Grundlagen- und Orientierungsprüfung** ist das erfolgreiche Ablegen der Module „Grundlagen der Computerlinguistik I“ sowie „Grundlagen der Informatik“.

Bachelorarbeit

Die Bachelorarbeit soll maximal 40 Seiten Text umfassen und wird mit 10 ECTS-Punkten bewertet. Die Zulassung zur Bachelorarbeit erfolgt nach erfolgreichem Ablegen sämtlicher in

der Fachprüfungsordnung Linguistische Informatik angegebenen Module außer „Vertiefungsmodul Computerlinguistik“.

3. Adressen

Abteilung für Computerlinguistik

Bismarckstr. 6, 91054 Erlangen

Studienfachberatung / Anrechnung von Studienleistungen

Prof. Dr. Stefan Evert

Raum 4.000, Tel.: 09131/85-22426

E-Mail: stefan.evert@fau.de

Sprechstunde: Mo. 14-15 Uhr; Termine in der vorlesungsfreien Zeit: nach Vereinbarung per E-Mail

Prüfungsamt

Halbmondstr. 6, Zi. 1.033 / 1.034 / 1.059, 91054 Erlangen

Tel. 09131/85-24813, -26716, -24834, -24835

geöffnet: Mo – Fr, 8.30 – 12.00 Uhr

4. Übersicht über die Module in Linguistischer Informatik

Sem.	Modul	Lehrveranstaltung	SWS	ECTS	Leistungsnachweis
1	Grundlagen der Computerlinguistik I	Vorlesung Computerlinguistik 1	2	2	Klausur (90 Min.)
		Übung zu Computerlinguistik 1	2	3	
		Arbeitstechniken	2	2,5	
1	Grundlagen der Informatik (Importmodul)			7,5	<i>Je nach Maßgabe des Faches</i>
2	Grundlagen der Computerlinguistik II	Vorlesung Computerlinguistik 2	2	2	Klausur (90 Min.)
		Übung zu Computerlinguistik 2	2	3	
2-3	Programmierung	Grundseminar Programmierung	2	5	Portfolio: Programmieraufgaben (4 x 8h, 0%) und Klausur (90 Min., 100%)
		Aufbauseminar Programmierung	2	5	
3	Vertiefungsmodul Computerlinguistik I	Proseminar	2	10	Portfolio: Hausarbeit (ca. 10 S., 50%) und Projekt (ca. 50 Stunden, 50%)
		Werkzeuge und Infrastrukturen	2		
2 od. 3	Konzeptionelle Modellierung (Importmodul)			5	<i>Je nach Maßgabe des Faches</i>
4 od. 6	Vertiefungsmodul Computerlinguistik II	HS theoretisch	2	10	Portfolio: Hausarbeit (ca. 15 S., 50%) und Projekt (ca. 80 Stunden) mit Projektbericht (10 S., 50%)
5		HS praktisch	2		
4	Korpuslinguistik	HS Korpuslinguistik	2	5	Klausur (90 Min.)
		Übung Statistik	2	5	
5 od. 6	Praktikum	ca. vierwöchiges Praktikum		5	Präsentation (30 Min.) und Praktikumsbescheinigung
6	Bachelorarbeit	<i>(Nur Erstfach)</i>		10	<i>Bachelorarbeit (25-40 S.)</i>

5. Informationen im Internet

Linguistische Informatik <http://www.linguistik.fau.de/>

Studien- und Prüfungsordnungen <https://www.fau.de/studium/im-studium/pruefungen-studienordnungen/>

Vorlesungsverzeichnis <http://www.vorlesungsverzeichnis.uni-erlangen.de>

Infos des IBZ (Infos zu Fächern, Zulassung, Career Service...) <http://www.fau.de/studium/>

Berufsbezogene Informationen der Bundesagentur für Arbeit <http://www.berufenet.de>