

## Studium

Computerlinguistik wird am Centrum für Informations- und Sprachverarbeitung (CIS) als Bachelor- und Masterstudiengang (Haupt- oder Nebenfach) angeboten. Es besteht darüber hinaus die Möglichkeit zur Promotion.

Die Computerlinguistik ist ein interdisziplinäres Forschungsfeld. Daher erwerbt ihr im Studium Kenntnisse aus sehr unterschiedlichen Disziplinen, z.B.

- aus der **Informatik** die Verwendung von geeigneten Programmier Techniken und -sprachen sowie das Design effizienter Algorithmen und Speichertechniken,
- aus der **Sprachwissenschaft** die Begriffe zur Beschreibung von Wortbildung (Morphologie), Satz- und Textaufbau (Syntax), Bedeutung (Semantik) und Verwendungsweise (Pragmatik),
- aus der **Statistik** Techniken zur Unterstützung unterschiedlicher Formen des maschinellen Lernens,
- aus der **Logik** Techniken zur formalen Repräsentation von Bedeutungen.

Computerlinguistik ist am CIS kein Massenstudiengang – der persönliche Kontakt zum Dozenten ist Normalität. Viele Kurse sind klein und ermöglichen einen dialogorientierten Lehrstil. Viele Studenten kommen aus Osteuropa, Asien oder dem Nahen Osten. Es besteht ein Austausch mit unterschiedlichen europäischen Universitäten.

## Anfahrt

Das CIS hat Lehrstühle an zwei Orten in München:

- Oettingenstr. 67 (direkt am Englischen Garten)  
Das Sekretariat und die meisten Mitarbeiter sind im ersten Stock, Flügel C, zu finden. Die meisten Vorlesungen finden in diesem Gebäude statt.  
🚶 16 (Tivolistr.) von der U-Bahn-Station Lehel  
🚶 58 & 154 (Tivolistr.), 54 (Hirschauer Str.) von der U-Bahn-Station Giselastraße
- Akademiestr. 7 (beim LMU-Hauptgebäude)  
🚶 U3 & U6 (Universität)  
🚶 27 & 28 (Nordendstr.)  
🚶 58, 68 & 154 (Siegstor), 153 (Universität)

[www.cis.lmu.de](http://www.cis.lmu.de)  
[www.cis.lmu.de/fachschaft](http://www.cis.lmu.de/fachschaft)



# COMPUTER LINGUISTIK

BSc/MSc/PhD

Centrum für  
Informations- und  
Sprachverarbeitung  
(CIS)



# Was ist Computerlinguistik?

Computerlinguistik erforscht die maschinelle Verarbeitung natürlicher Sprachen. Sie erarbeitet die theoretischen Grundlagen der Analyse, Erkennung und Erzeugung geschriebener und gesprochener Sprache durch Maschinen. Diese Aufgabenstellung führt auf natürliche Weise zu einer interdisziplinären Ausrichtung.

Nachbarfächer von besonderer Bedeutung für die Computerlinguistik sind Informatik, Statistik, Sprachwissenschaften, Phonetik sowie die Philosophie und Logik.

Praktische Anwendungen der Computerlinguistik sind u.a.:

- Maschinelle und computerunterstützte Übersetzung
- Umwandlung von gesprochener in geschriebene Sprache (Spracherkennung) und Dialogsysteme
- Informationsextraktion aus Texten und Informationssuche in Texten
- Suchmaschinentechnologie
- Digitale Bibliotheken und Archive
- Klassifikation von Texten und automatische Extraktion von Schlagwörtern
- Sentimentanalyse (z.B. für Bewertungsportale, Blogs)
- Automatische Fragenbeantwortung (Question Answering)
- Rechtschreibkorrektur und Grammatikprüfung
- Unterstützung datenorientierter Forschung in den Geisteswissenschaften

# Forschung

Trotz aller Fortschritte erreichen Rechner auch heute bei weitem nicht das menschliche Verständnis beim Umgang mit natürlicher Sprache. Gleichzeitig ist im Zeitalter des Internets ein dramatisches Wachstum des verfügbaren Bestands „roher“ (weitgehend unanalysierter) Sprache in elektronisierter Form zu verzeichnen.

Die *theoretische Computerlinguistik* versucht, Grundlagen für ein vertieftes Verständnis natürlicher Sprache durch Rechner bereitzustellen. Die *praktische Computerlinguistik* versucht, menschliche Nutzer bei allen Formen des Umgangs mit elektronischen Sprachbeständen aktiv zu unterstützen sowie eine korrekte Umwandlung gesprochener oder gedruckter Sprache in ein elektronisches Format zu erreichen. Die sich hieraus ergebenden Teilaufgaben und Anwendungen sind extrem vielfältig. Eine ständige Belebung ergibt sich durch neue Internetanwendungen.

Zentraler Forschungsschwerpunkt am CIS ist der Einsatz statistischer und lernender Verfahren bei der Analyse natürlicher Sprache. Wichtige Anwendungsfelder und Forschungsbereiche am CIS sind derzeit beispielsweise die automatische Übersetzung, die Verminderung von sozialen Stereotypen in automatisch generierten Texten, die automatische Erkennung von Hasskommentaren, die automatische Verarbeitung von Sprachen, für die es nur wenige Trainingsdaten gibt, die linguistische Analyse neuronaler Netze und die computergestützte Untersuchung von sprachwissenschaftlichen Fragestellungen (sowie viele weitere Themen).

# Berufsbild

Suchmaschinen, digitale Bibliotheken und Webportale für Service, Verkauf und Handel sind auf computerlinguistisches Know-How angewiesen, um für die jeweiligen Kunden geeignete Such und Zugriffsmöglichkeiten auf Webdokumente, Artikel, Angebote und Produktbeschreibungen zu schaffen.

In vielen wirtschaftlich wichtigen Bereichen erfordert die Umstellung auf digitalisierte Arbeitsprozesse computerlinguistische Hilfe (teilautomatisierte Analyse von Patientenreports und Abrechnungen im Bereich Medizin, teilautomatische Beantwortung von Korrespondenz bei Versicherungen, usw.). Große Softwarehäuser sowie Verlage beschäftigen Computerlinguisten, die an der Erschließung der Sprachdatenbestände (z.B. für elektronische Taschenwörterbücher) sowie an der Übersetzung oder Verarbeitung von Dokumenten arbeiten.

Interessante Arbeitsmöglichkeiten bieten auch Universitäten in der gesamten Welt, an denen die Computerlinguistik mittlerweile einen festen Platz einnimmt.

Der Technologiestandort München verheißt gute Aussichten auf Praktika und Arbeitsplätze im näheren Umfeld der Universität. Niederlassungen diverser Suchmaschinenbetreiber, Verlage und Medienunternehmen sowie die Bayerische Staatsbibliothek bieten weitere Arbeitsmöglichkeiten in München und Umgebung. Nicht zuletzt bietet München eine sehr aktive Startup-Szene – auch aus dem CIS sind bereits verschiedene Firmengründungen erfolgt.