

МИНОБРАЗОВАНИЯ РОССИИ



Федеральное государственное автономное образовательное учреждение  
высшего образования  
«Российский государственный гуманитарный университет»  
(ФГАОУ ВО «РГГУ»)

ИНСТИТУТ ЛИНГВИСТИКИ  
УНЦ компьютерной лингвистики

## КОМПЬЮТЕРНАЯ ЛИНГВИСТИКА

### РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

---

45.03.03 Фундаментальная и прикладная лингвистика

*Код и наименование направления подготовки/специальности*

---

Фундаментальная и прикладная лингвистика

*Наименование направленности (профиля)/ специализации*

Уровень высшего образования: *бакалавриат*

Форма обучения: *очная*

РПД адаптирована для лиц  
с ограниченными возможностями  
здоровья и инвалидов

Москва 2025

*Компьютерная лингвистика*  
Рабочая программа дисциплины

Составитель:

*к. филол. н., доцент УНЦ компьютерной лингвистики А.Ч. Пиперски*

УТВЕРЖДЕНО

Протокол заседания УНЦ компьютерной лингвистики

№5 от 16.12.2024

**ОГЛАВЛЕНИЕ**

1. Пояснительная записка.....	4
1.1. Цель и задачи дисциплины.....	4
1.2. Перечень планируемых результатов обучения по дисциплине, соотнесенных с индикаторами достижения компетенций.....	5
1.3. Место дисциплины в структуре образовательной программы.....	6
2. Структура дисциплины.....	6
3. Содержание дисциплины.....	7
4. Образовательные технологии.....	8
5. Оценка планируемых результатов обучения.....	10
5.1 Система оценивания.....	10
5.2 Критерии выставления оценки по дисциплине.....	10
5.3 Оценочные средства (материалы) для текущего контроля успеваемости, промежуточной аттестации обучающихся по дисциплине.....	12
6. Учебно-методическое и информационное обеспечение дисциплины.....	13
6.1 Список источников и литературы.....	13
6.2 Перечень ресурсов информационно-телекоммуникационной сети «Интернет».....	13
7. Материально-техническое обеспечение дисциплины.....	13
8. Обеспечение образовательного процесса для лиц с ограниченными возможностями здоровья и инвалидов.....	13
9. Методические материалы.....	15
9.1 Планы семинарских занятий.....	15
9.2 Иные материалы.....	15
Приложение 1. Аннотация рабочей программы дисциплины.....	16

## 1. Пояснительная записка

### 1.1. Цель и задачи дисциплины

Цель дисциплины — освоение студентами базовых понятий компьютерной и базовых принципов автоматической обработки естественного языка.

#### Задачи

- познакомить студентов с основными задачами, стоящими в современной компьютерной лингвистике, и с методами их решения;

- указать на связь между содержательными характеристиками языковых явлений и способами их автоматической обработки при решении практических задач компьютерной лингвистики;

- научить студентов пользоваться базовыми программными продуктами, разработанными в компьютерной лингвистике, знать их области применения;

В результате освоения дисциплины обучающийся должен:

#### ***знать:***

- основные понятия и методы современной компьютерной лингвистики
- базовые принципы лингвистической разметки

#### ***уметь:***

- анализировать различные уровни языковой структуры
- решать конкретные компьютерно-лингвистические задачи

#### ***владеть:***

- современной терминологией компьютерной лингвистики
- методами решения компьютерно-лингвистических задач

**1.2. Перечень планируемых результатов обучения по дисциплине, соотнесенных с индикаторами достижения компетенций**

<b>Компетенция (код и наименование)</b>	<b>Индикаторы компетенции (код и наименование)</b>	<b>Результаты обучения</b>
ПК-1. Владеет основными методами лингвистического анализа с учетом языковых и экстралингвистических факторов.	ПК-1.1. Применяет современные методы лингвистического анализа для исследования явлений различных уровней языковой структуры.	Умеет решать лингвистические задачи с помощью методов компьютерной лингвистики; знает преимущества и недостатки конкретных методов при решении конкретных лингвистических задач.
	ПК-1.2. Знаком с основными параметрами языкового разнообразия и базовыми теоретическими концепциями описания языковых явлений в собственно лингвистическом и экстралингвистическом аспектах.	Знает основные понятия и методы компьютерной лингвистики; знаком с основным кругом лингвистических задач, поддающихся решению при помощи методов компьютерной автоматизации.
	ПК-1.3. Способен осуществлять самостоятельную научно-исследовательскую деятельность с опорой на методы современной лингвистики.	Умеет использовать доступные программные продукты для решения самостоятельно поставленных исследовательских задач в области теоретической и прикладной лингвистики; способен применять знания, полученные в рамках курсов программирования, для решения своих исследовательских задач.
ПК-7. Способен пользоваться лингвистически ориентированными программными продуктами и знает принципы создания электронных языковых ресурсов.	ПК-7.1. Использует программные продукты и электронные языковые ресурсы для лексического, морфологического и синтаксического анализа текста.	Ориентируется в современных программных ресурсах автоматической обработки естественного языка, умеет их использовать и, при необходимости, минимально дорабатывать для решения исследовательских задач в различных областях лингвистики.
	ПК-7.2. Пользуется	Умеет подбирать

	основными методами, способами и средствами получения, хранения, и обработки лингвистической информации, а также лингвистически ориентированными программными продуктами.	необходимые лингвистические ресурсы для различных задач лингвистического обеспечения систем (например, лексикографических, задач морфологического анализа и т.п.).
	ПК-7.3. Разрабатывает и создает различные типы электронных языковых ресурсов (словари, корпуса, базы данных) и способен оценить эффективность и точность работы лингвистических программных продуктов для решения исследовательских задач.	Владеет навыками оценки качества работы лингвистических программных продуктов; способен оценить полноту и точность выполнения лингвистически значимой задачи методами компьютерного моделирования.

### 1.3. Место дисциплины в структуре образовательной программы

Дисциплина «Компьютерная лингвистика» относится к части, формируемой участниками образовательных отношений, блока дисциплин учебного плана.

Для освоения дисциплины необходимы знания, умения и владения, сформированные в ходе изучения следующих дисциплин и прохождения практик: «Понятийный аппарат математики», «Вероятностные модели», «Введение в теорию языка», «Общая морфология», «Общий синтаксис», «Корпусная лингвистика», «Программирование в лингвистике».

В результате освоения дисциплины формируются знания, умения и владения, необходимые для изучения следующих дисциплин и прохождения практик: «Автоматический перевод», преддипломная практика.

## 2. Структура дисциплины

Общая трудоёмкость дисциплины составляет 2 з.е., 72 академических часа.

### Структура дисциплины для очной формы обучения

Объем дисциплины в форме контактной работы обучающихся с педагогическими работниками и (или) лицами, привлекаемыми к реализации образовательной программы на иных условиях, при проведении учебных занятий:

Семестр	Тип учебных занятий	Количество часов
	Лекции	12
	Семинары	16
	Всего:	28

Объем дисциплины (модуля) в форме самостоятельной работы обучающихся составляет 44 академических часа.

### 3. Содержание дисциплины

№	Наименование раздела дисциплины	Содержание
1	Компьютерная лингвистика и автоматическая обработка естественного языка	Компьютерная лингвистика и автоматическая обработка естественного языка. Лингвистический и инженерный подход к компьютерной лингвистике. Задачи компьютерной лингвистики. Сложности при обработке естественного языка: омонимия, синонимия, проблемы с пониманием прагматики и т. п.
2	Сегментация текста на токены и предложения. Проблемы токенизации	Сегментация текста на токены ( $\approx$ слова) и предложения. Проблемы токенизации и деления на предложения в языках с различными системами графики. Образец простейшего токенизатора с использованием регулярных выражений.
3	$n$ -граммные языковые модели. Сглаживание	$n$ -граммные языковые модели. Оценка вероятности для последовательности слов. Оценка $n$ -граммных моделей. Перплексивность. Сглаживание: метод Лапласа, интерполяция и откат.
4	Стемминг, лемматизация и морфологическая разметка	Понятия стемминга, лемматизации, частеречная разметка и морфологическая разметка. Стандарты морфологической разметки для русского и английского языка. Омонимия и её разрешение. Скрытые марковские модели. Алгоритм Витерби. Таггер Брилла.
5	Формальное представление синтаксиса. Основные алгоритмы парсинга	Формальное представление синтаксиса: структура зависимостей и структура составляющих. Синтаксически аннотированные корпуса. Типология формальных грамматик. Основные алгоритмы парсинга. Stanford Parser, MaltParser.
6	Решение конкретных компьютерно-лингвистических задач	Оценка качества в компьютерной лингвистике. Автоматическая проверка орфографии. Машинный перевод. Классификация и кластеризация текстов. Чат-боты. Информационный поиск.

## 4. Образовательные технологии

№ п/п	Наименование раздела	Виды учебных занятий	Образовательные технологии
1	Компьютерная лингвистика и автоматическая обработка естественного языка	Лекция 1	Вводная лекция с использованием презентации.
2	Сегментация текста на токены и предложения. Проблемы токенизации	Лекция 2	Лекция с использованием презентации.
		Семинар 1	Развёрнутое обсуждение прочитанной научной литературы. Совместное выполнение заданий с помощью компьютера с доступом к сети «Интернет» и электронным ресурсам, в т.ч. текстовым корпусам.
3	n-граммные языковые модели. Сглаживание	Лекция 3	Лекция с использованием презентации.
		Семинар 2	Развёрнутое обсуждение прочитанной научной литературы. Совместное выполнение заданий с помощью компьютера с доступом к сети «Интернет» и электронным ресурсам, в т.ч. текстовым корпусам.
		Семинар 3	Индивидуальное выполнение заданий с помощью компьютера с доступом к сети «Интернет» и электронным ресурсам, в т.ч. текстовым корпусам, и совместное обсуждение после.
4	Стемминг, лемматизация и морфологическая разметка	Лекция 4	Лекция с использованием презентации.
		Семинар 4	Развёрнутое обсуждение прочитанной научной литературы. Индивидуальное выполнение заданий с помощью компьютера с доступом к сети «Интернет» и электронным ресурсам, в т.ч. текстовым корпусам, и совместное обсуждение после.
5	Формальное представление синтаксиса. Основные алгоритмы парсинга	Лекция 5	Лекция с использованием презентации.

		Семинар 5	Развёрнутое обсуждение прочитанной научной литературы. Индивидуальное выполнение заданий с помощью компьютера с доступом к сети «Интернет» и электронным ресурсам, в т.ч. текстовым корпусам, и совместное обсуждение после.
6	Решение конкретных компьютерно-лингвистических задач	Лекция 6	Лекция с использованием презентации.
		Семинар 6	Обсуждение существующих компьютерно-лингвистических задач (оценка качества в компьютерной лингвистике, информационный поиск) и способов их решения. Индивидуальное выполнение заданий с помощью компьютера с доступом к сети «Интернет» и электронным ресурсам, в т.ч. текстовым корпусам, и совместное обсуждение после.
		Семинар 7	Обсуждение существующих компьютерно-лингвистических задач (автоматическая проверка орфографии, машинный перевод) и способов их решения. Индивидуальное выполнение заданий с помощью компьютера с доступом к сети «Интернет» и электронным ресурсам, в т.ч. текстовым корпусам, и совместное обсуждение после.
		Семинар 8	Обсуждение существующих компьютерно-лингвистических задач (классификация и кластеризация текстов, чат-боты) и способов их решения. Индивидуальное выполнение заданий с помощью компьютера с доступом к сети «Интернет» и электронным ресурсам, в т.ч. текстовым корпусам, и совместное обсуждение после.
7	Экзамен		Защита и обсуждение докладов на ранее заданные темы.

В период временного приостановления посещения обучающимися помещений и территории РГГУ для организации учебного процесса с применением электронного обучения и дистанционных образовательных технологий могут быть использованы следующие образовательные технологии:

- видео-лекции;
- онлайн-лекции в режиме реального времени;
- электронные учебники, учебные пособия, научные издания в электронном виде и доступ к иным электронным образовательным ресурсам;
- системы для электронного тестирования;

– консультации с использованием телекоммуникационных средств.

## 5. Оценка планируемых результатов обучения

### 5.1 Система оценивания

Форма контроля	Макс. количество баллов	
	За одну работу	Всего
Текущий контроль:		
- участие в дискуссиях в ходе лекций (в т.ч. в обсуждении прочитанной литературы)	3 балла	12 баллов
- выполнение заданий в ходе семинаров (темы 2-6)	8 баллов	48 баллов
Промежуточная аттестация – экзамен		40 баллов
<b>Итого за семестр</b>		<b>100 баллов</b>

Полученный совокупный результат конвертируется в традиционную шкалу оценок и в шкалу оценок Европейской системы переноса и накопления кредитов (European Credit Transfer System; далее – ECTS) в соответствии с таблицей:

100-балльная шкала	Традиционная шкала		Шкала ECTS
95 – 100	отлично	зачтено	A
83 – 94			B
68 – 82	хорошо		C
56 – 67	удовлетворительно		D
50 – 55		E	
20 – 49	неудовлетворительно	не зачтено	FX
0 – 19			F

### 5.2 Критерии выставления оценки по дисциплине

Баллы/ Шкала ECTS	Оценка по дисциплине	Критерии оценки результатов обучения по дисциплине
100-83/ A,B	отлично/ зачтено	Выставляется обучающемуся, если он глубоко и прочно усвоил теоретический и практический материал, может продемонстрировать это на занятиях и в ходе промежуточной аттестации. Обучающийся исчерпывающе и логически стройно излагает учебный материал, умеет увязывать теорию с практикой, справляется с решением задач профессиональной направленности высокого уровня сложности, правильно обосновывает принятые решения.

Баллы/ Шкала ECTS	Оценка по дисциплине	Критерии оценки результатов обучения по дисциплине
		<p>Свободно ориентируется в учебной и профессиональной литературе.</p> <p>Оценка по дисциплине выставляются обучающемуся с учётом результатов текущей и промежуточной аттестации.</p> <p>Компетенции, закреплённые за дисциплиной, сформированы на уровне – «высокий».</p>
82-68/ С	хорошо/ зачтено	<p>Выставляется обучающемуся, если он знает теоретический и практический материал, грамотно и по существу излагает его на занятиях и в ходе промежуточной аттестации, не допуская существенных неточностей.</p> <p>Обучающийся правильно применяет теоретические положения при решении практических задач профессиональной направленности разного уровня сложности, владеет необходимыми для этого навыками и приёмами.</p> <p>Достаточно хорошо ориентируется в учебной и профессиональной литературе.</p> <p>Оценка по дисциплине выставляются обучающемуся с учётом результатов текущей и промежуточной аттестации.</p> <p>Компетенции, закреплённые за дисциплиной, сформированы на уровне – «хороший».</p>
67-50/ D,E	удовлетво- рительно/ зачтено	<p>Выставляется обучающемуся, если он знает на базовом уровне теоретический и практический материал, допускает отдельные ошибки при его изложении на занятиях и в ходе промежуточной аттестации.</p> <p>Обучающийся испытывает определённые затруднения в применении теоретических положений при решении практических задач профессиональной направленности стандартного уровня сложности, владеет необходимыми для этого базовыми навыками и приёмами.</p> <p>Демонстрирует достаточный уровень знания учебной литературы по дисциплине.</p> <p>Оценка по дисциплине выставляются обучающемуся с учётом результатов текущей и промежуточной аттестации.</p> <p>Компетенции, закреплённые за дисциплиной, сформированы на уровне – «достаточный».</p>
49-0/ F,FX	неудовлет- ворительно/ не зачтено	<p>Выставляется обучающемуся, если он не знает на базовом уровне теоретический и практический материал, допускает грубые ошибки при его изложении на занятиях и в ходе промежуточной аттестации.</p> <p>Обучающийся испытывает серьёзные затруднения в применении теоретических положений при решении практических задач профессиональной направленности стандартного уровня сложности, не владеет необходимыми для этого навыками и приёмами.</p> <p>Демонстрирует фрагментарные знания учебной литературы по дисциплине.</p> <p>Оценка по дисциплине выставляются обучающемуся с учётом результатов текущей и промежуточной аттестации.</p> <p>Компетенции на уровне «достаточный», закреплённые за дисциплиной, не сформированы.</p>

### 5.3 Оценочные средства (материалы) для текущего контроля успеваемости, промежуточной аттестации обучающихся по дисциплине

#### *Примеры заданий к семинарам*

1. Таргет Penn: найти ошибки
  - I/PRP need/VBP a/DT flight/NN from/IN Atlanta/NN
  - Does/VBZ this/DT flight/NN serve/VB dinner/NNS
  - I/PRP have/VB a/DT friend/NN living/VBG in/IN Denver/NNP
  - What/WDT flights/NNS do/VBP you/PRP have/VB from/IN Milwaukee/NNP to/IN Tampa/NNP
  - Can/VBP you/PRP list/VB the/DT nonstop/JJ afternoon/NN flights/NNS
  
2. Таргет Penn: разметить
  - It is a nice night.
  
3. Построить регулярные выражения, которые будут находить в тексте:
  - IPv4 адреса
  - телефонные номера
  - инициалы и фамилии (А.С. Пушкин)
  
4. Оценить перплексивность униграммной, биграммной и триграммной модели, обученной на Национальном корпусе русского языка.

#### *Примеры статей для докладов*

- Kuhn, Tobias. 2014. A survey and classification of controlled natural languages. *Computational Linguistics* 40.1: 121–170.
- Kernighan, Mark, Kenneth Church & William Gale. 1990. A spelling correction program based on a noisy channel model.
- Johannes Schaback, Fang Li. 2007. Multi-level feature extraction for spelling correction.
- Hobbs, Jerry R. & Ellen Riloff. 2010. Information extraction. In: Indurkha & Damerau (2010) (eds.). 511–532.
- Liu, Bing. 2010. Sentiment analysis and subjectivity. In: Indurkha & Damerau (2010) (eds.). 627–666.
- Knight, Kevin & Graehl Jonathan. 1999. Machine transliteration. *Computational Linguistics* 24.4: 599–612.

## 6. Учебно-методическое и информационное обеспечение дисциплины

### 6.1 Список источников и литературы

#### Основная литература

- Шунейко, А. А. Квантитативная лингвистика и новые информационные технологии : учебник для вузов / А. А. Шунейко, И. А. Авдеенко. — Москва : Издательство Юрайт, 2023. — 347 с. — (Высшее образование). — ISBN 978-5-534-15446-7. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/543983>
- Jurafsky, Dan & James H. Martin. 2017. *Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition*. 3rd edition draft. <https://web.stanford.edu/~jurafsky/slp3/ed3book.pdf>

### 6.2 Перечень ресурсов информационно-телекоммуникационной сети «Интернет».

- Национальный корпус русского языка (НКРЯ): <http://ruscorpora.ru/>
- Гайдлайны по стандартизованному представлению текстов разных видов <https://teic.org/Guidelines/P5/>
- Regular Expression Cheat Sheet <https://www.cheatography.com/davechild/cheat sheets/regular expressions>

## 7. Материально-техническое обеспечение дисциплины

Лекционные занятия по проводятся с использованием компьютерных презентаций, поэтому в аудитории необходимы компьютер и проектор, а также соответствующее освещение. В ходе семинарских занятий студенты должны быть обеспечены компьютерами с лицензионным программным обеспечением и выходом в Интернет, т.е. занятия должны проходить в компьютерных классах.

## 8. Обеспечение образовательного процесса для лиц с ограниченными возможностями здоровья и инвалидов

В ходе реализации дисциплины используются следующие дополнительные методы обучения, текущего контроля успеваемости и промежуточной аттестации обучающихся в зависимости от их индивидуальных особенностей:

- для слепых и слабовидящих: лекции оформляются в виде электронного документа, доступного с помощью компьютера со специализированным программным обеспечением; письменные задания выполняются на компьютере со специализированным программным обеспечением или могут быть заменены устным ответом; обеспечивается индивидуальное равномерное освещение не менее 300 люкс; для выполнения задания при необходимости предоставляется увеличивающее устройство; возможно также использование собственных увеличивающих устройств; письменные задания оформляются увеличенным шрифтом; экзамен и зачёт проводятся в устной форме или выполняются в письменной форме на компьютере.

- для глухих и слабослышащих: лекции оформляются в виде электронного документа, либо предоставляется звукоусиливающая аппаратура индивидуального пользования; письменные задания выполняются на компьютере в письменной форме; экзамен и зачёт проводятся в письменной форме на компьютере; возможно проведение в форме тестирования.

- для лиц с нарушениями опорно-двигательного аппарата: лекции оформляются в виде электронного документа, доступного с помощью компьютера со специализированным программным обеспечением; письменные задания выполняются на компьютере со специализированным программным обеспечением; экзамен и зачёт проводятся в устной форме или выполняются в письменной форме на компьютере.

При необходимости предусматривается увеличение времени для подготовки ответа.

Процедура проведения промежуточной аттестации для обучающихся устанавливается с учётом их индивидуальных психофизических особенностей. Промежуточная аттестация может проводиться в несколько этапов.

При проведении процедуры оценивания результатов обучения предусматривается использование технических средств, необходимых в связи с индивидуальными особенностями обучающихся. Эти средства могут быть предоставлены университетом, или могут использоваться собственные технические средства.

Проведение процедуры оценивания результатов обучения допускается с использованием дистанционных образовательных технологий.

Обеспечивается доступ к информационным и библиографическим ресурсам в сети Интернет для каждого обучающегося в формах, адаптированных к ограничениям их здоровья и восприятия информации:

- для слепых и слабовидящих: в печатной форме увеличенным шрифтом, в форме электронного документа, в форме аудиофайла.
- для глухих и слабослышащих: в печатной форме, в форме электронного документа.
- для обучающихся с нарушениями опорно-двигательного аппарата: в печатной форме, в форме электронного документа, в форме аудиофайла.

Учебные аудитории для всех видов контактной и самостоятельной работы, научная библиотека и иные помещения для обучения оснащены специальным оборудованием и учебными местами с техническими средствами обучения:

- для слепых и слабовидящих: устройством для сканирования и чтения с камерой SARA CE; дисплеем Брайля PAC Mate 20; принтером Брайля EmBraille ViewPlus;
- для глухих и слабослышащих: автоматизированным рабочим местом для людей с нарушением слуха и слабослышащих; акустический усилитель и колонки;
- для обучающихся с нарушениями опорно-двигательного аппарата: передвижными, регулируемые эргономическими партами СИ-1; компьютерной техникой со специальным программным обеспечением.

## 9. Методические материалы

### 9.1 Планы семинарских занятий

#### **Семинар 1. Сегментация текста на токены и предложения. Проблемы токенизации**

##### Вопросы для обсуждения:

Сегментация текста на токены ( $\approx$  слова) и предложения. Проблемы токенизации и деления на предложения в языках с различными системами графики. Образец простейшего токенизатора с использованием регулярных выражений.

#### **Семинары 2-3. n-граммные языковые модели. Сглаживание**

##### Вопросы для обсуждения:

n-граммные языковые модели. Оценка вероятности для последовательности слов. Оценка n-граммных моделей. Перплексивность. Сглаживание: метод Лапласа, интерполяция и откат.

#### **Семинар 4. Стемминг, лемматизация и морфологическая разметка**

##### Вопросы для обсуждения:

Понятия стемминга, лемматизации, частеречная разметка и морфологическая разметка. Стандарты морфологической разметки для русского и английского языка. Омонимия и её разрешение. Скрытые марковские модели. Алгоритм Витерби. Таггер Брилла..

#### **Семинар 5. Формальное представление синтаксиса. Основные алгоритмы парсинга**

##### Вопросы для обсуждения:

Формальное представление синтаксиса: структура зависимостей и структура составляющих. Синтаксически аннотированные корпуса. Типология формальных грамматик. Основные алгоритмы парсинга. Stanford Parser, MaltParser.

#### **Семинары 6-8. Решение конкретных компьютерно-лингвистических задач**

##### Вопросы для обсуждения:

Оценка качества в компьютерной лингвистике. Автоматическая проверка орфографии. Машинный перевод. Классификация и кластеризация текстов. Чат-боты. Информационный поиск.

### 9.2 Иные материалы

#### *Рекомендуемая литература для более глубокого освоения программы*

- Николаев И. С., Митренина О. В., Ландо Т. М. (ред.). 2016. Прикладная и компьютерная лингвистика. М.: URSS.
- Indurkha, Nitin & Fred J. Damerau (eds.). 2010. *Handbook of Natural Language Processing*. 2nd ed. Boca Raton: Chapman & Hall/CRC.

## АННОТАЦИЯ РАБОЧЕЙ ПРОГРАММЫ ДИСЦИПЛИНЫ

Цель дисциплины — освоение студентами базовых понятий компьютерной и базовых принципов автоматической обработки естественного языка.

### Задачи

- познакомить студентов с основными задачами, стоящими в современной компьютерной лингвистике, и с методами их решения;
- указать на связь между содержательными характеристиками языковых явлений и способами их автоматической обработки при решении практических задач компьютерной лингвистики;
- научить студентов пользоваться базовыми программными продуктами, разработанными в компьютерной лингвистике, знать их области применения;

В результате освоения дисциплины обучающийся должен:

### ***знать:***

- основные понятия и методы современной компьютерной лингвистики
- базовые принципы лингвистической разметки

### ***уметь:***

- анализировать различные уровни языковой структуры
- решать конкретные компьютерно-лингвистические задачи

### ***владеть:***

- современной терминологией компьютерной лингвистики
- методами решения компьютерно-лингвистических задач