

Документ подписан простой электронной подписью  
Информация о владельце:  
ФИО: Смирнов Сергей Николаевич  
Должность: врио ректора  
Дата подписания: 27.03.2025 15:09:44  
Уникальный программный ключ:  
69e375c64f7e975d4e8830e7b4fcc2ad1bf35f08

Министерство образования и науки Российской Федерации  
ФГБОУ ВО «Тверской государственный университет»

Утверждаю

Руководитель ООП

М.Л. Логунов

13 мая 2024 г.



Рабочая программа дисциплины (с аннотацией)

Технология обработки текста и звучащей речи

Направление подготовки

45.03.03. «Фундаментальная и прикладная лингвистика»

Направление (профиль)

Теория языка и языковые технологии

Для студентов: 4 курса очной формы обучения

Составитель: к.ф.н., доцент Е.П. Максимова

Тверь 2024

## **I. Аннотация**

### **1. Цель и задачи дисциплины**

Целью освоения дисциплины являются: теоретическое и практическое освоение будущими бакалаврами системы теоретических и практических знаний в области технологий обработки текста и звучащей речи.

Задачами освоения дисциплины являются: подготовка обучающихся к самостоятельному использованию технологий обработки текста и звучащей речи процессе обучения в вузе и в будущей профессиональной деятельности; привитие обучающимся навыков разработки элементов технологий обработки текста и звучащей речи (самостоятельно и в сотрудничестве с программистами и специалистами других предметных областей).

### **2. Место дисциплины в структуре ООП**

Дисциплина «Технологии обработки текста и звучащей речи» входит в часть, формируемой участниками образовательных отношений. Она опирается на знания и навыки, получаемые студентами в ходе освоения курсов «Основы прикладной лингвистики», «Общая фонетика», «Информатика и основы программирования», а также на практическое и теоретическое владение родным и иностранными языками. Дисциплина «Технологии обработки текста и звучащей речи» подготавливает обучающихся к использованию продуктов и технологических средств обработки текстов и звучащей речи в их учебной и будущей профессиональной деятельности.

**3. Объем дисциплины:** 2 зачетные единицы, 72 академических часов, в том числе:

**Контактная аудиторная работа:** лекции 22 часа;

**Самостоятельная работа:** 50 часов.

**4. Планируемые результаты обучения по дисциплине, соотнесенные с планируемыми результатами освоения образовательной программы**

Планируемые результаты освоения образовательной программы (формируемые компетенции)	Планируемые результаты обучения по дисциплине
ПК-2. Способность спланировать и провести лингвистический эксперимент, описать его результаты и сформулировать выводы. ПК-1. Владение основными	2.1. Систематизирует и классифицирует лингвистический материал для организации работ по созданию и редактированию контента сайта учреждения.

методами фонологического, морфологического, синтаксического, дискурсивного и семантического анализа с учетом языковых и экстралингвистических факторов.	1.1. Применяет при подготовке и размещении публикаций специальных текстов основные понятия фонетики и фонологии, морфологии, синтаксиса, дискурса и семантики с учётом языковых и экстралингвистических факторов;
---	---

**5. Форма промежуточной аттестации и семестр прохождения** зачет (7 семестр)

**6. Язык преподавания** русский.

**II. Содержание дисциплины, структурированное по темам с указанием отведенного на них количества академических часов и видов учебных занятий**

Учебная программа – наименование разделов и тем	Всего (час.)	Контактная работа (час.)			Самостоятельная работа, в том числе контроль (час.)
		Лекции	Семинарские/Практические занятия/Лабораторные	Контроль самостоятельной работы (в том числе лабораторные)	
1. Автоматическая обработка естественного языка. Общие принципы и сферы применения.	4	2			2
2. Оптимизация общения с ЭВМ. Системы обработки естественного языка. Базовые компоненты систем, обеспечивающих взаимодействие с ЭВМ.	6	2			4
3. Типология систем, обеспечивающих взаимодействие с ЭВМ. Вопросы - ответные системы	4	2			2

, диалоговые системы решения задач, системы обработки связных текстов.					
4. Теория и практика информационно -поисковых систем. Основные понятия информационного поиска. Типы информационно -поисковых систем. Информационно -поисковые языки.	5	1			4
5. Моделирование общения. Компьютерные модели диалога. Программа «Элиза», ее алгоритмы	6	2			4
6. Компьютерное моделирование структуры сюжета. Морфологический, синтаксический и когнитивный подходы.	5	1			4
7. Гипертекстовые системы представления текста. История развития. Теоретические основы гипертекста. Элементы типологии гипертекста.	4	2			2
8. Компьютерное моделирование языка и речи. Статистико-дистрибутивный метод.	6	2			4
9. Авторизация текста. Множественная неопределенность, сравнение по образцу, конкуренция образцов	6	2			4

10.Машинный перевод. Основные этапы развития. Элементы периодизации стратегий, типология систем. Основные области использования машинного перевода. Перспективы развития систем машинного перевода.	6	2			4
11.Технологии обработки звучащей речи. Основные направления речевых технологий: анализ речи, синтез речи, шумоочистка, идентификация по голосу.	5	1			4
12. История создания систем анализа и синтеза речи.	5	1			4
13. Фонетические основания систем Анализа и синтеза речи. Экспериментальные данные современной фонетики и их роль в речевых технологиях.	5	1			4
14. Современное состояние и перспективы развития речевых технологий.	5	1			4
Итого	72	22			50

### III. Образовательные технологии

Учебная программа – наименование разделов и тем	Вид занятия	Образовательные технологии
1. Автоматическая обработка естественного языка. Общие принципы и сферы применения.	Лекция, с/р	Традиционная лекция с элементами лекции- презентации, работа с программными продуктами, работа в интернет
2. Оптимизация общения с ЭВМ. Системы обработки естественного языка. Базовые компоненты систем, обеспечивающих взаимодействие с ЭВМ.	Лекция, с/р	Традиционная лекция с элементами лекции- презентации, работа с программными продуктами, работа в интернет
3. Типология систем , обеспечивающих взаимодействие с ЭВМ. Вопросно - ответные системы , диалоговые системы решения задач, системы обработки связных текстов.	Лекция, с/р	Традиционная лекция с элементами лекции- презентации, работа с программными продуктами, работа в интернет.
4. Теория и практика информационно -поисковых систем. Основные понятия информационного поиска. Типы информационно -поисковых систем. Информационно -поисковые языки.	Лекция, с/р	Традиционная лекция, проблемная лекция, метод малых групп, различные виды тестов, коллоквиум, лекция- презентация, работа с программными продуктами, работа в интернет
5. Моделирование общения. Компьютерные модели диалога. Программа «Элиза», ее алгоритмы	Лекция, с/р	Традиционная лекция с элементами лекции- презентации, работа с программными продуктами, работа в интернет
6. Компьютерное	Лекция, с/р	Традиционная лекция с элементами лекции-

<p>моделирование структуры сюжета. Морфологический, синтаксический и когнитивный подходы.</p>		<p>презентации, работа с программными продуктами, работа в интернет</p>
<p>7. Гипертекстовые системы представления текста. История развития. Теоретические основы гипертекста. Элементы типологии гипертекста.</p>	<p>Лекция, с/р</p>	<p>Традиционная лекция с элементами лекции-презентации, работа с программными продуктами, работа в интернет</p>
<p>8. Компьютерное моделирование языка и речи. Статистико-дистрибутивный метод.</p>	<p>Лекция, с/р</p>	<p>Традиционная лекция с элементами лекции-презентации, работа с программными продуктами, работа в интернет</p>
<p>9. Авторизация текста. Множественная неопределенность, сравнение по образцу, конкуренция образцов</p>	<p>Лекция, с/р</p>	<p>Традиционная лекция с элементами лекции-презентации, работа с программными продуктами, работа в интернет</p>
<p>10. Машинный перевод. Основные этапы развития. Элементы периодизации стратегий, типология систем. Основные области использования машинного перевода. Перспективы развития систем машинного перевода.</p>	<p>Лекция, с/р</p>	<p>Традиционная лекция с элементами лекции-презентации, работа с программными продуктами, работа в интернет</p>
<p>11. Технологии обработки звучащей речи. Основные направления речевых технологий: анализ речи, синтез речи, шумоочистка,</p>	<p>Лекция, с/р</p>	<p>Традиционная лекция с элементами лекции-презентации, работа с программными продуктами, работа в интернет</p>

идентификация по голосу.		
12. История создания систем анализа и синтеза речи.	Лекция, с/р	Традиционная лекция, работа в интернет
13. Фонетические основания систем Анализа и синтеза речи. Экспериментальные данные современной фонетики и их роль в речевых технологиях.	Лекция, с/р	Традиционная лекция, работа в интернет
14. Современное состояние и перспективы развития речевых технологий.	Лекция, с/р	Традиционная лекция с элементами лекции-презентации, работа с программными продуктами, работа в интернет

#### **IV. Оценочные материалы для проведения текущей и промежуточной аттестации**

##### Примерные аттестующие задания

##### Модуль 1

Ответить на вопросы:

1. Сферы применения автоматической обработки естественного языка.
2. Принципы работы анализатора и синтезатора речи в автоматических системах распознавания речи.
3. Классификация видов автоматизированной обработки текста в зависимости от цели использования.
4. Методы воспроизведения и обработки данных.
5. Данные. Информация. Знания
6. Знания. Виды знаний.
7. Системы продукций.
8. Фреймы.
9. Семантические сети.
10. Исчисление предикатов.
11. Нечеткая логика. Ответить на вопросы:
  1. Типы текстовых структур в системах АОТ.
  2. Графематический анализ текста.

3. Морфологический анализ текста.
4. Синтаксический анализ текста.
5. Синтаксический анализ текста.
6. Семантический анализ текста.

## Модуль 2

Ответить на вопросы:

1. Принципы работы автоматизированных информационных систем.
2. Информационно-поисковые системы.
3. Автоматическое реферирование/фрагментирование текстов
4. Компьютерные автоматические словари.
- 5.Тезаурусы Стратегии машинного перевода.
- 6.Принципы работы экспертных систем.
- 7.Системы «вопрос-ответ»
- 8.Определение и функции гипертекста.
- 9.Формы и виды гипертекста.
- 10.Области применения гипертекстовых систем
11. Области применения систем анализа и синтеза речи.
12. Алгоритмы систем анализа речи.
13. Алгоритмы систем синтеза реч

Требования к зачету:

уровень освоения студентами учебного материала;

умения студента использовать теоретические знания при выполнении практических задач;

сформированность общеучебных умений;

обоснованность и четкость изложения ответа;

умение четко формулировать проблему, предложив ее решение, критически оценить решение и его последствия;

умение показать, проанализировать альтернативные возможности, варианты действий;

умение сформировать свою позицию, оценку и аргументировать ее.

Вопросы к зачету

1. Что такое автоматическая обработка естественного языка
2. Экспертные системы, их лингвистическое обеспечение
3. Автоматический анализ и синтез речи
4. Области применения гипертекстовых систем
5. Автоматические системы распознавания речи
6. Структурные элементы гипертекста.
7. Автоматическая обработка текста. Виды систем АОТ. Области применения систем АОТ
8. Гипертекстовые текстовые технологии. Формы и виды гипертекста.
9. Основные задачи и классы систем АОТ.
10. Основные области использования систем машинного перевода. Некоторые существующие системы МП.
11. Методы воспроизведения и обработки данных
12. Стратегии машинного перевода. Классификация систем МП.
13. Информация. Свойства информации.
14. Системы машинного перевода: история создания, периодизация.
15. Данные. Основные структуры данных.
16. Компьютерные программы поддержки лексикографических работ. Автоматические словари.
17. Знания. Виды знаний. Приобретение и извлечение знаний.
18. Информационно-поисковые системы.
19. Представление знаний в системах АОТ. Проблема предметной области.
20. Автоматическое аннотирование текстов. Построение аннотаций.
21. Способы формального представления знаний. Системы продукций.
22. Системы автоматического реферирования / фрагментирования текстов.
23. Способы формального представления знаний. Семантические сети.
24. Автоматические системы взаимодействия человека и машины.
25. Способы формального представления знаний. Фреймы.
26. Семантический анализ текста в системах АОТ.
27. Способы формального представления знаний. Исчисление предикатов.
28. Синтаксический компонент в АОТ.
29. Способы формального представления знаний. Нечеткая логика.
30. Морфологический анализ текста в системах АОТ.
31. Текст и его основные характеристики в информатике.
32. Компонент первичного автоматического анализа текстов (графематический анализ)
33. Особенности систем обработки связных текстов.
34. Состав компонентов стандартной системы понимания текстов.
35. Типы текстовых структур в системах автоматического понимания текстов.
36. Системы автоматического извлечения знаний из текстов.

37. Анализатор речи и синтезатор речи в системах автоматического распознавания речи.
38. Необходимость создания компьютерных систем обработки текстов и их особенности.
39. История создания гипертекста и его отличие от обычного текста.
40. Области применения экспертных систем. Некоторые существующие экспертные системы.
41. Области применения систем анализа и синтеза речи.
42. Алгоритмы систем анализа речи.
43. Алгоритмы систем синтеза речи.

## **V. Учебно-методическое и информационное обеспечение дисциплины**

### 1) Рекомендуемая литература

#### а) Основная литература

Левицкий Ю. А. Общее языкознание : учебное пособие. - М.: Директ-Медиа, 2013. - 361 с. - [Электронный ресурс]. - РЕЖИМ ДОСТУПА: <http://biblioclub.ru/index.php?page=book&id=210686>

Актуальные проблемы современной лингвистики : учебное пособие / сост. Л.Н. Чурилина. - 10-е изд., стереотип. - М. : Флинта, 2017. - 412 с. - ISBN 978-5-89349-892-9 ; То же [Электронный ресурс]. - РЕЖИМ ДОСТУПА: <http://biblioclub.ru/index.php?page=book&id=103797>

Хроленко А. Т. , Денисов А. В. Современные информационные технологии для гуманитария: практическое руководство Москва: Издательство «Флинта», 2018  
[https://biblioclub.ru/index.php?page=book\\_red&id=363413](https://biblioclub.ru/index.php?page=book_red&id=363413)

#### б) Дополнительная литература

1. Звегинцев, В.А. Язык как фактор компьютерной революции // НТИ. Сер. 2. 1985. N 9. Искусственный интеллект. В 3-х кн.
2. Системы общения и экспертные системы. М., 1990.

3. Молчанова Л.В. Качественный и количественный аспекты лексико-семантической прогностики. – Воронеж: Издательский дом ВГУ, 2016. – 173 с.
4. Прикладная лингвистика: Материалы к семинарским занятиям и лабораторным работам / Сост. С.П. Хижняк. – Саратов: Саратовск. юрид. академия, 2012. – 212 с.
5. Гердт А.С. Научно-техническая лексикография // Прикладное языкознание. Спб., 1996.
6. Гердт А.С. Предмет и основные направления прикладной лингвистики // Прикладное языкознание. Спб., 1996.
7. Зубов А.В., Зубова И.И. Информационные технологии в лингвистике: учебное пособие для студентов лингвистических факультетов вузов. М., 2004. – 208 с.
8. Кибрик А.Е. Очерки по общим и прикладным вопросам языкознания. М., 1992.
9. Леонтьева Н.Н. Автоматическое понимание текстов: системы, модели, ресурсы. Учебное пособие для студентов лингвистических факультетов вузов. М., 2006. – 304 с.
10. Марусенко М.А. Атрибуция анонимных и псевдонимных текстов методами прикладной лингвистики // прикладное языкознание. Спб, 1996.
11. Панков И.П., Захаров В.П. Информационно-поисковые системы // Прикладное языкознание. Спб., 1996.
12. Потапова Р.К. Речь: коммуникация, информация, кибернетика. М., 1997.

## 2) Программное обеспечение

а) Лицензионное программное обеспечение

б) Свободно распространяемое программное обеспечение

3) Современные профессиональные базы данных и информационные справочные системы

Автоматическая обработка текста. Технологии. [www. aot. ru](http://www.aot.ru)

The Association for Computational Linguistics [http://www. aclweb. org/](http://www.aclweb.org/)

Компьютерная лингвистика и интеллектуальные технологии // Конференция «Диалог» [http://www. dialog-21. ru/](http://www.dialog-21.ru/)

Англо-русский словарь по вычислительной технике, информационным технологиям и связи под общей редакцией [http://www. morepc. ru/dict/](http://www.morepc.ru/dict/)

Технологии синтеза речи и распознавание голоса <http://speech-soft. ru/>

Современные технологии распознавания речи  
[http://index. php? nma=catalog&fla=stat&cat\\_id=3&page=1&nums=9](http://index. php? nma=catalog&fla=stat&cat_id=3&page=1&nums=9)

Обзор компьютерных программных решений распознавания и синтеза речи  
[http://index. php? nma=catalog&fla=stat&cat\\_id=3&page=1&nums=7](http://index. php? nma=catalog&fla=stat&cat_id=3&page=1&nums=7)  
<http://www. cti. ru/>

СТИ – Communications. Technology. Innovations.//российский системный интегратор и разработчик программного обеспечения, специализирующийся в области IP-коммуникаций.

Online Dictionaries– Лингво <http://slovari. yandex. ru>

Webster Dictionary

<http://www. /- Oxford Dictionaries>

<http://www. /- Cambridge Advanced Learner's Dictionary>

<http://dictionary. cambridge. org/>

4) Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины

Автоматическая обработка текста. Технологии. [www. aot. ru](http://www. aot. ru)

The Association for Computational Linguistics <http://www. aclweb. org/>

Компьютерная лингвистика и интеллектуальные технологии  
//Конференция «Диалог» <http://www. dialog-21. ru/>

Англо-русский словарь по вычислительной технике, информационным технологиям и связи под общей редакцией [http://www. morepc. ru/dict/](http://www.morepc.ru/dict/)

Технологии синтеза речи и распознавание голоса [http://speech-soft. ru/](http://speech-soft.ru/)

Современные технологии распознавания речи

[http://index. php? nma=catalog&fla=stat&cat\\_id=3&page=1&nums=9](http://index.php? nma=catalog&fla=stat&cat_id=3&page=1&nums=9)

Обзор компьютерных программных решений распознавания и синтеза речи

[http://index. php? nma=catalog&fla=stat&cat\\_id=3&page=1&nums=7](http://index.php? nma=catalog&fla=stat&cat_id=3&page=1&nums=7)

[http://www. cti. ru/](http://www.cti.ru/)

CTI – Communications. Technology. Innovations.//российский системный интегратор и разработчик программного обеспечения, специализирующийся в области IP-коммуникаций.

Online Dictionaries– Лингво [http://slovari. yandex. ru](http://slovari.yandex.ru)

Webster Dictionary

<http://www. /– Oxford Dictionaries>

<http://www. /– Cambridge Advanced Learner's Dictionary>

[http://dictionary. cambridge. org/](http://dictionary.cambridge.org/)

## **VI. Методические материалы для обучающихся по освоению дисциплины**

1. При выполнении задания обратитесь к материалам соответствующих глав основной литературы, рекомендованных интернет-источников.

2. Обязательно указать цели, задачи, актуальность, новизну проекта и методы, которые были использованы при работе над материалом.
3. Также необходимо указать возможные гуманитарные сферы использования презентуемой разработки.

### **Методические рекомендации по подготовке к зачету**

1. При подготовке к зачету повторите весь теоретический материал курса, используя лекции, рекомендованные научные источники.
2. Отвечая на теоретический вопрос, обязательно употребляйте профессиональные термины, раскрывая их содержание, устанавливайте системные связи между основными понятиями.
3. При выполнении реконструктивного задания необходимо продемонстрировать умение грамотно вести анализ, правильно выбирая факты из предложенного материала, на их основании представлять целостную картину описания, обязательно аргументируя свои выводы.

### **Требования к рейтинг-контролю**

#### 1-ая контрольная точка

I	Текущая работа	Количество баллов
1.	Посещение лекций и работа на занятии	20 б.
2.	Выполнение самостоятельной работы	20 б.
	Всего:	40 б.
II.	Итоговая контрольная работа	10 б.
	Всего:	50 б.

#### 2-ая контрольная точка.

I.	Текущая работа	
1.	Посещение лекций и работа на занятии	10 б.
2.	Выполнение самостоятельной работы	10 б.
	Всего:	20 б.
II.	Итоговая контрольная работа	10 б.
III	Зачет	40 б.
	Всего:	100 б.

## **VII. Материально-техническое обеспечение**

Наименование помещений	Оснащенность помещений для самостоятельной работы	Перечень лицензионного программного обеспечения.  Реквизиты подтверждающего документа
<p>№ 28 компьютерный класс  (170002, Тверская область, Тверь, просп. Чайковского, д.70)</p>	<p>Компьютеры Norbel. Мониторы BENQ</p>	<p>ABC Pascal  Adobe Acrobat Reader DC - Russian  GIMP 2.8.16  Google Chrome  Kaspersky Endpoint Security 10 для Windows  Microsoft Office профессиональный плюс 2013 (Акт приема-передачи № 369 от 21 июля 2017)  Visual Studio 2015 (Акт предоставления прав № Tr035055 от 19.06.2017)  VLC media player  WinDjView 2.0.2  Microsoft Windows 10 Enterprise (Акт приема-передачи № 369 от 21 июля 2017)</p>

Помещения для самостоятельной работы:

Наименование помещений	Оснащенность помещений для самостоятельной работы	Перечень лицензионного программного обеспечения.  Реквизиты подтверждающего документа
<p>№ 28 компьютерный класс</p> <p>(170002, Тверская область, Тверь, просп. Чайковского, д.70)</p>	<p>Компьютеры Norbel. Мониторы BENQ</p>	<p>ABCPascal</p> <p>Adobe Acrobat Reader DC - Russian</p> <p>GIMP 2.8.16</p> <p>Google Chrome</p> <p>Kaspersky Endpoint Security 10 для Windows</p> <p>Microsoft Office профессиональный плюс 2013 (Акт приема-передачи № 369 от 21 июля 2017)</p> <p>Visual Studio 2015 (Акт предоставления прав № Tr035055 от 19.06.2017)</p> <p>VLC media player</p> <p>WinDjView 2.0.2</p> <p>Microsoft Windows 10 Enterprise (Акт приема-передачи № 369 от 21 июля 2017)</p>

### VIII. Сведение об обновлении рабочей программы дисциплины

№ п.п.	Обновленный раздел рабочей программы дисциплины	Описание внесенных изменений	Реквизиты документа, утвердившего изменения