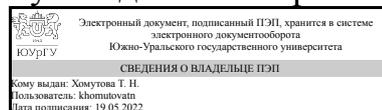


УТВЕРЖДАЮ:  
Руководитель направления



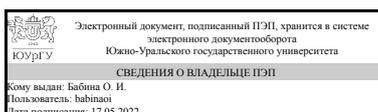
Т. Н. Хомутова

## РАБОЧАЯ ПРОГРАММА

**дисциплины** ФД.05 Автоматическая обработка естественного языка  
**для направления** 45.04.02 Лингвистика  
**уровень** Магистратура  
**форма обучения** очная  
**кафедра-разработчик** Лингвистика и перевод

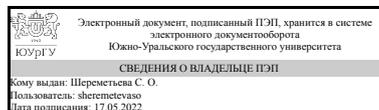
Рабочая программа составлена в соответствии с ФГОС ВО по направлению подготовки 45.04.02 Лингвистика, утверждённым приказом Минобрнауки от 12.08.2020 № 992

Зав.кафедрой разработчика,  
к.филол.н., доц.



О. И. Бабина

Разработчик программы,  
д.филол.н., доц., профессор



С. О. Шереметьева

## 1. Цели и задачи дисциплины

Целью курса является ознакомить студентов с понятиями автоматической обработки текста и звучащей речи  
Задачи: • ознакомить студентов с основными приложениями систем автоматической обработки текста и звучащей речи • рассмотреть основные сферы применения систем автоматической обработки текста и звучащей речи • рассмотреть основные подходы к разработке систем автоматической обработки текста и звучащей речи (лингвистический и статистический) • научить основным приемам морфологического, синтаксического и семантического анализа текстов • научить основным приемам синтеза текстов • научить основным приемам создания лингвистического обеспечения компьютерных систем обработки текстов и звучащей речи

## Краткое содержание дисциплины

Введение в автоматическую обработку текста Автоматический перевод  
Автоматический информационный поиск Автоматическое аннотирование и реферирование  
Автоматический анализ и синтез информации Компьютерная обработка речи

## 2. Компетенции обучающегося, формируемые в результате освоения дисциплины

Планируемые результаты освоения ОП ВО (компетенции)	Планируемые результаты обучения по дисциплине
ПК-1 Способен осуществлять автоматизацию переводческих проектов	Знает: основные задачи автоматизации обработки языка в переводческих проектах; базовые принципы автоматической обработки языковых данных; лингвистически-ориентированные интеллектуальные и информационные электронные системы и принципы работы с ними Умеет: использовать в переводческих проектах существующие системы обработки естественного языка, интеллектуальные и информационные электронные системы; проводить их сравнительный анализ; проектировать модули данных систем Имеет практический опыт: работы с системами обработки естественного языка; проектирования модулей таких систем

## 3. Место дисциплины в структуре ОП ВО

Перечень предшествующих дисциплин, видов работ учебного плана	Перечень последующих дисциплин, видов работ
Нет	ФД.02 Электронные ресурсы переводчика, 1.Ф.04 Практический курс машинного и синхронного перевода, ФД.06 Основы программирования лингвистических задач (язык Python)

Требования к «входным» знаниям, умениям, навыкам студента, необходимым при освоении данной дисциплины и приобретенным в результате освоения предшествующих дисциплин:

Нет

#### 4. Объём и виды учебной работы

Общая трудоемкость дисциплины составляет 3 з.е., 108 ч., 56,5 ч. контактной работы

Вид учебной работы	Всего часов	Распределение по семестрам в часах	
		Номер семестра	
		2	
Общая трудоёмкость дисциплины	108	108	
<i>Аудиторные занятия:</i>	48	48	
Лекции (Л)	16	16	
Практические занятия, семинары и (или) другие виды аудиторных занятий (ПЗ)	32	32	
Лабораторные работы (ЛР)	0	0	
<i>Самостоятельная работа (СРС)</i>	51,5	51,5	
с применением дистанционных образовательных технологий	0		
Анализ современной литературы по изучаемым темам	26	26	
Подготовка к экзамену	25,5	25,5	
Консультации и промежуточная аттестация	8,5	8,5	
Вид контроля (зачет, диф.зачет, экзамен)	-	экзамен	

#### 5. Содержание дисциплины

№ раздела	Наименование разделов дисциплины	Объем аудиторных занятий по видам в часах			
		Всего	Л	ПЗ	ЛР
1	Введение в автоматическую обработку текста	2	2	0	0
2	Автоматический перевод	22	6	16	0
3	Автоматический информационный поиск	6	2	4	0
4	Автоматическое аннотирование и реферирование	6	2	4	0
5	Автоматический анализ и синтез информации	6	2	4	0
6	Компьютерная обработка речи	6	2	4	0

##### 5.1. Лекции

№ лекции	№ раздела	Наименование или краткое содержание лекционного занятия	Кол-во часов
1	1	Введение в автоматическую обработку текста	2
2	2	Приемы компьютерного анализа текстов	2
3	2	Автоматическая генерация текстов	2

4	2	Разработка автоматических многоязычных лексиконов	2
5	3	Автоматический информационный поиск	2
6	4	Автоматическое аннотирование и реферирование	2
7	5	Автоматический анализ и синтез информации	2
8	6	Компьютерная обработка речи	2

## 5.2. Практические занятия, семинары

№ занятия	№ раздела	Наименование или краткое содержание практического занятия, семинара	Кол-во часов
1-3	2	Морфологический анализ текстов	6
4-6	2	Синтаксический анализ текстов	6
7-8	2	Разработка двуязычного лексикона	4
9-10	3	Извлечение ключевых слов и поиск информации	4
11-12	4	Составление рефератов и аннотаций по формальным критериям	4
13-14	5	Моделирование генерации текста на естественном языке	4
15-16	6	Инструментальный анализ речевого сигнала	4

## 5.3. Лабораторные работы

Не предусмотрены

## 5.4. Самостоятельная работа студента

Выполнение СРС			
Подвид СРС	Список литературы (с указанием разделов, глав, страниц) / ссылка на ресурс	Семестр	Кол-во часов
Анализ современной литературы по изучаемым темам	Шипицына, Л.Ю. Информационные технологии в лингвистике: учебное пособие. - М.: Флинта : Наука, 2013. Боярский, К.К. Введение в компьютерную лингвистику: учебное пособие. - СПб: НИУ ИТМО, 2013. Автоматическая обработка текстов на естественном языке и компьютерная лингвистика : учеб. пособие / Большакова Е.И., Клышинский Э.С., Ландэ Д.В., Носков А.А., Пескова О.В., Ягунова Е.В.— М.: МИЭМ, 2011. — 272 с. Марчук Ю.Н. Компьютерная лингвистика: учебное пособие. М., 2007 Зубов А.В., Зубова И.И. Основы искусственного интеллекта для лингвистов. М., 2007 С.В. Рыбин СИНТЕЗ РЕЧИ Учебное пособие <a href="https://books.ifmo.ru/file/pdf/1925.pdf">https://books.ifmo.ru/file/pdf/1925.pdf</a> Автоматический анализ и синтез звучащей речи этапы автоматического анализа речи.	2	26
Подготовка к экзамену	Шипицына, Л.Ю. Информационные технологии в лингвистике: учебное пособие. - М.: Флинта : Наука, 2013. Боярский, К.К. Введение в компьютерную	2	25,5

	лингвистику: учебное пособие. - СПб: НИУ ИТМО, 2013. Автоматическая обработка текстов на естественном языке и компьютерная лингвистика : учеб. пособие / Большакова Е.И., Клышинский Э.С., Ландэ Д.В., Носков А.А., Пескова О.В., Ягунова Е.В.— М.: МИЭМ, 2011. — 272 с. Марчук Ю.Н. Компьютерная лингвистика: учебное пособие. М., 2007 Зубов А.В., Зубова И.И. Основы искусственного интеллекта для лингвистов. М., 2007 С.В. Рыбин СИНТЕЗ РЕЧИ Учебное пособие <a href="https://books.ifmo.ru/file/pdf/1925.pdf">https://books.ifmo.ru/file/pdf/1925.pdf</a> Автоматический анализ и синтез звучащей речи этапы автоматического анализа речи.		
--	---	--	--

## 6. Фонд оценочных средств для проведения текущего контроля успеваемости, промежуточной аттестации

Контроль качества освоения образовательной программы осуществляется в соответствии с Положением о балльно-рейтинговой системе оценивания результатов учебной деятельности обучающихся.

### 6.1. Контрольные мероприятия (КМ)

№ КМ	Се-местр	Вид контроля	Название контрольного мероприятия	Вес	Макс. балл	Порядок начисления баллов	Учитывается в ПА
1	2	Текущий контроль	Контрольная работа 1	1	9	Контрольная работа содержит три задания. Каждое задание оценивается максимум в три балла. Три балла выставляется, если задание выполнено в полном объеме и корректно. Два балла выставляется, если задание выполнено в полном объеме и с незначительными ошибками. 1 балл выставляется, если задание выполнено с многочисленными ошибками. 0 баллов выставляется, если задание не выполнено .	экзамен
2	2	Текущий контроль	Контрольная работа 2	1	9	Контрольная работа содержит три задания. Каждое задание оценивается максимум в три балла. Три балла выставляется, если задание выполнено в полном объеме и корректно. Два балла выставляется, если задание выполнено в полном объеме и с незначительными ошибками. 1 балл выставляется, если задание выполнено с многочисленными ошибками. 0 баллов выставляется, если задание не выполнено	экзамен
3	2	Текущий	Контрольная	1	9	Контрольная работа содержит три задания.	экзамен

		контроль	работа 3			Каждое задание оценивается максимум в три балла. Три балла выставляется, если задание выполнено в полном объеме и корректно. Два балла выставляется, если задание выполнено в полном объеме и с незначительными ошибками. 1 балл выставляется, если задание выполнено с многочисленными ошибками. 0 баллов выставляется, если задание не выполнено	
4	2	Промежуточная аттестация	экзамен	-	12	Экзамен содержит четыре задания. Каждое задание оценивается максимум в три балла. Три балла выставляется, если задание выполнено в полном объеме и корректно. Два балла выставляется, если задание выполнено в полном объеме и с незначительными ошибками. 1 балл выставляется, если задание выполнено с многочисленными ошибками. 0 баллов выставляется, если задание не выполнено .	экзамен

## 6.2. Процедура проведения, критерии оценивания

Вид промежуточной аттестации	Процедура проведения	Критерии оценивания
экзамен	Студенты получают билеты с вопросами по пройденным темам. Отвечают на них письменно и сдают для проверки преподавателю, который с учетом баллов за контрольные работы выставляет окончательную оценку.	В соответствии с пп. 2.5, 2.6 Положения

## 6.3. Паспорт фонда оценочных средств

Компетенции	Результаты обучения	№ КМ			
		1	2	3	4
ПК-1	Знает: основные задачи автоматизации обработки языка в переводческих проектах; базовые принципы автоматической обработки языковых данных; лингвистически-ориентированные интеллектуальные и информационные электронные системы и принципы работы с ними	+	+	+	+
ПК-1	Умеет: использовать в переводческих проектах существующие системы обработки естественного языка, интеллектуальные и информационные электронные системы; проводить их сравнительный анализ; проектировать модули данных систем	+	+	+	+
ПК-1	Имеет практический опыт: работы с системами обработки естественного языка; проектирования модулей таких систем	+	+	+	+

Типовые контрольные задания по каждому мероприятию находятся в приложениях.

## 7. Учебно-методическое и информационное обеспечение дисциплины

### Печатная учебно-методическая документация

а) основная литература:

1. Прикладное языкознание Текст учеб. С. А. Аверина и др.; отв. ред. А. С. Герд ; Санкт-Петербург. гос. ун-т. - СПб.: Издательство С.-Петербургского университета, 1996. - 525, [1] с. ил.

*б) дополнительная литература:*

1. Зубов, А. В. Информационные технологии в лингвистике Текст учеб. пособие по специальности 021800 "Теорет. и приклад. лингвистика" А. В. Зубов, И. И. Зубова. - М.: Академия, 2004. - 205, [1] с. ил.

2. Потапова, Р. К. Речь : коммуникация, информация, кибернетика [Текст] учеб. пособие для вузов по специальностям "Автоматизир. системы обраб. информ. и упр.", "Лингвистика" Р. К. Потапова. - 2-е изд., доп. - М.: Эдиториал УРСС, 2001. - 564 с. ил.

*в) отечественные и зарубежные журналы по дисциплине, имеющиеся в библиотеке:*

1. Вестник ЮУрГУ. Серия Лингвистика
2. Научно-техническая информация. Серия 2

*г) методические указания для студентов по освоению дисциплины:*

1. Шереметьева, С.О. Информационные технологии в помощь переводчику Текст Ч. 2 учеб. пособие для фак. лингвистики по направлению 45.03.02 и др. направлениям С. О. Шереметьева, П. Г. Осминин ; Юж.-Урал. гос. ун-т, Каф. Лингвистика и межкультур. коммуникация ; ЮУрГУ . - Челябинск Издательский Центр ЮУрГУ 2014 . - 39 с.

2. Бабина, О.И. Информационные технологии в помощь переводчику Текст Ч. 1 учеб. пособие по направлению 032700.68 и др. направлениям О. И. Бабина ; Юж.-Урал. гос. ун-т, Каф. Лингвистика и межкультур. коммуникация ; ЮУрГУ . - Челябинск Издательский Центр ЮУрГУ 2013 - 59 с.

*из них: учебно-методическое обеспечение самостоятельной работы студента:*

1. Шереметьева, С.О. Информационные технологии в помощь переводчику Текст Ч. 2 учеб. пособие для фак. лингвистики по направлению 45.03.02 и др. направлениям С. О. Шереметьева, П. Г. Осминин ; Юж.-Урал. гос. ун-т, Каф. Лингвистика и межкультур. коммуникация ; ЮУрГУ . - Челябинск Издательский Центр ЮУрГУ 2014 . - 39 с.

2. Бабина, О.И. Информационные технологии в помощь переводчику Текст Ч. 1 учеб. пособие по направлению 032700.68 и др. направлениям О. И. Бабина ; Юж.-Урал. гос. ун-т, Каф. Лингвистика и межкультур. коммуникация ; ЮУрГУ . - Челябинск Издательский Центр ЮУрГУ 2013 - 59 с.

**Электронная учебно-методическая документация**

№	Вид литературы	Наименование ресурса в электронной форме	Библиографическое описание
1	Дополнительная литература	Электронно-библиотечная система издательства	Ганегедара, Т. Обработка естественного языка с TensorFlow : руководство / Т. Ганегедара ; перевод с английского В. С. Яценкова. — Москва : ДМК Пресс, 2020. — 382 с. — ISBN 978-5-97060-756-5. — Текст : электронный // Лань :

		Лань	электронно-библиотечная система. — URL: <a href="https://e.lanbook.com/book/140584">https://e.lanbook.com/book/140584</a> (дата обращения: 31.10.2021). — Режим доступа: для авториз. пользователей.
2	Дополнительная литература	Электронно-библиотечная система издательства Лань	Гольдберг, Й. Нейросетевые методы в обработке естественного языка : руководство / Й. Гольдберг ; перевод с английского А. А. Слинкина. — Москва : ДМК Пресс, 2019. — 282 с. — ISBN 978-5-97060-754-1. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <a href="https://e.lanbook.com/book/131704">https://e.lanbook.com/book/131704</a> (дата обращения: 31.10.2021). — Режим доступа: для авториз. пользователей.
3	Дополнительная литература	Электронно-библиотечная система издательства Лань	Бонцанини, М. Анализ социальных медиа на Python. Извлекайте и анализируйте данные из всех уголков социальной паутины на Python / М. Бонцанини ; перевод с английского А. В. Логунова. — Москва : ДМК Пресс, 2018. — 288 с. — ISBN 978-5-97060-574-5. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <a href="https://e.lanbook.com/book/108129">https://e.lanbook.com/book/108129</a> (дата обращения: 31.10.2021). — Режим доступа: для авториз. пользователей.
4	Дополнительная литература	Электронно-библиотечная система издательства Лань	Душкин, Р. В. Искусственный интеллект / Р. В. Душкин. — Москва : ДМК Пресс, 2019. — 280 с. — ISBN 978-5-97060-787-9. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <a href="https://e.lanbook.com/book/131703">https://e.lanbook.com/book/131703</a> (дата обращения: 31.10.2021). — Режим доступа: для авториз. пользователей.
5	Дополнительная литература	Электронно-библиотечная система издательства Лань	Риз, Р. Обработка естественного языка на Java : учебное пособие / Р. Риз ; перевод с английского А. В. Снастина. — Москва : ДМК Пресс, 2016. — 264 с. — ISBN 978-5-97060-331-4. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <a href="https://e.lanbook.com/book/93272">https://e.lanbook.com/book/93272</a> (дата обращения: 31.10.2021). — Режим доступа: для авториз. пользователей.
6	Основная литература	Электронно-библиотечная система издательства Лань	Миронов, С. В. Формальные языки и грамматики : учебное пособие / С. В. Миронов. — Саратов : СГУ, 2019. — 80 с. — ISBN 978-5-292-04612-7. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <a href="https://e.lanbook.com/book/148854">https://e.lanbook.com/book/148854</a> (дата обращения: 31.10.2021). — Режим доступа: для авториз. пользователей.
7	Дополнительная литература	Электронно-библиотечная система издательства Лань	Митчелл, Р. Скрапинг веб-сайтов с помощью Python : руководство / Р. Митчелл ; перевод с английского А. В. Груздев. — Москва : ДМК Пресс, 2016. — 280 с. — ISBN 978-5-97060-223-2. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <a href="https://e.lanbook.com/book/100903">https://e.lanbook.com/book/100903</a> (дата обращения: 31.10.2021). — Режим доступа: для авториз. пользователей.
8	Основная литература	Электронно-библиотечная система издательства Лань	Гаврилова, Т. А. Инженерия знаний. Модели и методы : учебник для вузов / Т. А. Гаврилова, Д. В. Кудрявцев, Д. И. Муромцев. — 4-е изд., стер. — Санкт-Петербург : Лань, 2021. — 324 с. — ISBN 978-5-8114-8793-6. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <a href="https://e.lanbook.com/book/180874">https://e.lanbook.com/book/180874</a> (дата обращения: 31.10.2021). — Режим доступа: для авториз. пользователей.
9	Дополнительная литература	Электронно-библиотечная система издательства	Паттерсон, Д. Глубокое обучение с точки зрения практика / Д. Паттерсон, А. Гибсон. — Москва : ДМК Пресс, 2018. — 418 с. — ISBN 978-5-97060-481-6. — Текст : электронный // Лань : электронно-библиотечная система. — URL:

		Лань	<a href="https://e.lanbook.com/book/116122">https://e.lanbook.com/book/116122</a> (дата обращения: 31.10.2021). — Режим доступа: для авториз. пользователей.
10	Дополнительная литература	Электронно-библиотечная система издательства Лань	Гущин, А. Н. Личностно-ориентированные информационные системы : учебное пособие / А. Н. Гущин. — Санкт-Петербург : БГТУ "Военмех" им. Д.Ф. Устинова, 2012. — 120 с. — ISBN 978-5-85546-715-4. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <a href="https://e.lanbook.com/book/63719">https://e.lanbook.com/book/63719</a> (дата обращения: 31.10.2021). — Режим доступа: для авториз. пользователей.

Перечень используемого программного обеспечения:

1. Microsoft-Windows(бессрочно)
2. Microsoft-Office(бессрочно)

Перечень используемых профессиональных баз данных и информационных справочных систем:

1. EBSCO Information Services-EBSCOhost Research Databases(28.02.2017)
2. -База данных ВИНТИ РАН(бессрочно)

## 8. Материально-техническое обеспечение дисциплины

Вид занятий	№ ауд.	Основное оборудование, стенды, макеты, компьютерная техника, предустановленное программное обеспечение, используемое для различных видов занятий
Самостоятельная работа студента	478 (1)	Компьютерный класс, оснащенный компьютерной техникой с возможностью подключения к сети "Интернет" и обеспечением доступа в электронную информационно-образовательную среду университета
Практические занятия и семинары	478 (1)	Компьютерный класс, презентационные технологии (проектор, PowerPoint)
Лекции	462 (1)	Компьютер, презентационные технологии (проектор, PowerPoint)
Экзамен	478 (1)	Компьютерный класс, оснащенный компьютерной техникой с возможностью подключения к сети "Интернет" и обеспечением доступа в электронную информационно-образовательную среду университета