



МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
федеральное государственное автономное образовательное учреждение высшего
образования
«Дальневосточный федеральный университет»
(ДВФУ)

ИНСТИТУТ МАТЕМАТИКИ И КОМПЬЮТЕРНЫХ ТЕХНОЛОГИЙ (ШКОЛА)

«СОГЛАСОВАНО»

Руководитель ОП

Артемьева И.Л.

«УТВЕРЖДАЮ»

И.о. директора департамента

Смагин С.В.



«15» июля 2021 г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ
МЕТОДЫ РАБОТЫ С ЕСТЕСТВЕННЫМ ЯЗЫКОМ
Направление подготовки 09.03.04 Программная инженерия
(Программная инженерия)

Форма подготовки очная

курс 4 семестр 7
лекции 36 час.
практические занятия 00 час.
лабораторные работы 36 час.
в том числе с использованием МАО лек. 0 / пр. 0 / лаб. 18 час.
всего часов аудиторной нагрузки 72 час.
в том числе с использованием МАО 18 час.
самостоятельная работа 72 час.
в том числе на подготовку к экзамену 00 час.
контрольные работы (количество) не предусмотрены
курсовая работа / курсовой проект не предусмотрены
зачет 7 семестр
экзамен _____ семестр

Рабочая программа составлена в соответствии с требованиями Федерального государственного образовательного стандарта высшего образования по направлению подготовки 09.03.04 Программная инженерия, утвержденный приказом Министерства образования и науки Российской Федерации от 19 сентября 2017 г. № 920 (с изменениями и дополнениями).

Рабочая программа обсуждена на заседании кафедры прикладной математики, механики, управления и программного обеспечения, протокол № 7.1 от «04» июля 2019 г.

Заведующий кафедрой прикладной математики, механики, управления и программного обеспечения д.т.н., профессор Артемьева И.Л.

Составители: к.ф.-м.н., доцент Л.И. Прудникова

Владивосток
2021

Оборотная сторона титульного листа РПД

I. Рабочая программа пересмотрена на заседании кафедры прикладной математики, механики, управления и программного обеспечения:

Протокол от «09» июля 2021 г. № 7.1

Заведующий кафедрой _____ Артемяева И.Л.
(подпись) (И.О. Фамилия)

II. Рабочая программа пересмотрена на заседании департамента программной инженерии и искусственного интеллекта:

Протокол от «17» сентября 2021 г. № 9.1

И.о. директора департамента _____ Смагин С.В.
(подпись) (И.О. Фамилия)

III. Рабочая программа пересмотрена на заседании кафедры:

Протокол от « ____ » _____ 20__ г. № _____

Заведующий кафедрой _____
(подпись) (И.О. Фамилия)

IV. Рабочая программа пересмотрена на заседании кафедры:

Протокол от « ____ » _____ 20__ г. № _____

Заведующий кафедрой _____
(подпись) (И.О. Фамилия)

ЦЕЛИ И ЗАДАЧИ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Цель: курса «Методы работы с естественным языком» - сформировать у бакалавров систему знаний, связанную со всей сферой применения компьютерных моделей языка в лингвистике и смежных дисциплинах; сформировать представление о методах автоматической обработки естественно-языкового текста на всех уровнях лингвистического анализа; ознакомить студентов с практическими методами автоматической обработки естественно-языковых текстов.

Задачи дисциплины:

1. Обучение студентов методам формального представления и описания структур и закономерностей естественных языков;
2. Ознакомить с лингвистическим обеспечением различных прикладных систем (автоматического индексирования, реферирования, и т.д.).
3. Раскрыть структуру лингвистических процессоров.
4. Совершенствование знаний в области теории языка.

Планируемые результаты обучения по данной дисциплине (знания, умения, владения), соотнесенные с планируемыми результатами освоения образовательной программы.

Общепрофессиональные компетенции выпускников и индикаторы их достижения:

Наименование категории (группы) общепрофессиональных	Код и наименование общепрофессиональной компетенции	Код и наименование индикатора достижения общепрофессиональной компетенции
--	---	---

	<p>ОПК-2. Способен понимать принципы работы современных информационных технологий и программных средств, в том числе отечественного производства, и использовать их при решении задач профессиональной деятельности</p>	<p>ОПК-2.1. Знает типичные программные системы (приложения) в области АОТ и их архитектурные особенности</p> <p>ОПК-2.2. Умеет разрабатывать системы, которые позволяли бы взаимодействовать с ЭВМ в конкретной проблемной области на естественном языке или каком-то его ограниченном варианте</p> <p>ОПК-2.3. Владеет навыками работы со специальными программными средствами автоматизированной обработки текстов</p>
	<p>ОПК-7. Способен применять в практической деятельности основные концепции, принципы, теории и факты, связанные с информатикой</p>	<p>ОПК-7.1. Знает сферу применения и перспективы развития информационно-поисковых тезаурусов, применение тезаурусного метода поиска в Интернет существенные отличия естественных языков от искусственных и особенности компьютерных моделей естественного языка, какие лингвистические технологии следует применять в зависимости от условий задачи, понимать, на какие лингвистические данные и ресурсы может опираться</p> <p>ОПК-7.2. Умеет интерпретировать результаты автоматической обработки лингвистических данных разрабатывать системы, которые позволяли бы взаимодействовать с ЭВМ в конкретной проблемной области на естественном языке или каком-то его ограниченном</p> <p>ОПК-7.3. Владеет методикой использования компьютерных технологий в предметной области приемами работы с прикладным программным обеспечением</p>

Профессиональные компетенции выпускников и индикаторы их достижения:

Задача профессиональной деятельности	Объект или область знания	Код и наименование профессиональной компетенции	Код и наименование индикатора достижения профессиональной компетенции	Основание (ПС, анализ иных требований, предъявляемых к выпускникам)
Тип задач профессиональной деятельности: научно-исследовательский				

<p>Участие в научно-исследовательских и опытно-конструкторских работах в области программной инженерии. Анализ и выбор программно-технологических платформ, сервисов и информационных ресурсов программной инженерии; подготовка обзоров, аннотаций, составление рефератов и докладов, публикаций и библиографии по научно-исследовательской работе в области программной инженерии</p>	<p>Прикладные и информационные процессы. Информационные технологии и. Программное обеспечение</p>	<p>ПК-4. Готовность к использованию методов и инструментальных средств исследования объектов профессиональной деятельности</p>	<p>ПК-4.1. Знает технологию автоматизированной обработки текстовой информации, основные принципы представления знаний о предметной области в виде рубрикаторов, тезаурусов, онтологий особенности реализации естественно-языковых систем, назначение лингвистического ПК-4.2. Умеет проектировать процесс обработки языковых данных, работать с данными: понимать, какие именно данные нужны для решения поставленной задачи, как взаимодействуют лингвистические и иные данные, откуда их можно получить; определять основные классы ЕЯ-систем; ПК-4.3. Владеет широким диапазоном различных информационно-коммуникационных технологий; методами поиска информации по работе с естественным языком</p>	<p>06.028 Системный программист 06.022 Системный аналитик</p>
---	---	--	---	--

I. СТРУКТУРА И СОДЕРЖАНИЕ ТЕОРЕТИЧЕСКОЙ ЧАСТИ КУРСА

Лекционные занятия (36 час)

Раздел.1 Компьютерная лингвистика. (16ч.)

Тема 1. Моделирование в компьютерной лингвистике. (8ч.)

Задачи компьютерной лингвистики. Особенности системы ЕЯ: уровни и связи. Моделирование в компьютерной лингвистике. Автоматическая обработка звучащей речи. Автоматическая обработка письменного текста. Представление о модели "смысл \Leftrightarrow текст".

Тема 2. Инструментальные системы разработки приложений по автоматической обработке текстов на естественном языке. (8ч.)

Программные средства лингвистической обработки. Представление лингвистических данных. Корпусная лингвистика. Подходы к представлению данных. Лингвистическая разметка. Лингвистические аннотации. Представления, основанные на абстракции. Архитектура инструментальных ЕЯ-систем. Процессы обработки текста.

Раздел.2 Модель лингвистической онтологии для автоматической обработки текстов (10ч.)

Тема 1. Типы онтологий: верхнего уровня, предметных областей, прикладные онтологии. Лексические онтологии. (4ч.)

Определение понятий: онтология, концепт, отношение, аксиомы. Типы онтологий: верхнего уровня, предметных областей, прикладных онтологий. Лексические онтологии. Примеры онтологий.

Тема 2. Задачи, решаемые с помощью онтологий и тезаурусов (информационный поиск, интеграция гетерогенных источников данных, SemanticWeb) (6ч.)

Назначение онтологий. Информационный поиск. Интеграция разнородных источников данных. SemanticWeb. Онтологии предметных областей и прикладные онтологии: назначение, отличительные черты, решаемые задачи (примеры проектов). Языки описания онтологий. Основные синтаксические структуры: классы, отношения, аксиомы. Тезаурусы. Основные принципы разработки, создания и использования традиционных информационно-поисковых тезаурусов. Примеры тезаурусов. Информационно-поисковые тезаурусы в условиях сверхбольших

электронных коллекций и автоматической обработки текстов. Тезаурус для автоматического концептуального индексирования как особый вид тезауруса.

Раздел.3 Примеры систем (10ч.)

Тема 1. Программы анализа и лингвистической обработки текстов.(2ч.)

Программа для синтаксического и морфологического анализа русскоязычных текстов. Автоматизированная система поиска и аналитической обработки информации.

Тема 2. Программы преобразования текстов.(2ч.)

Программы переформатирования текстовых или html-файлов (Windows). Программы поиска и замены текста по заданным маскам. Массовые операции поиска и замены во множестве текстовых файлов.

Тема 3. Генераторы текстов и "говорящие" программы.(2ч.)

Система общения на русском языке. Автоматическое составление и перевод деловых писем и документов.

Тема 4. Системы обработки естественного языка (NL-Processing).(2ч.)

Программы поиска и анализ в текстовых массивах информации, реализованные с использованием технологии нейронных сетей. Встраивание в разрабатываемые приложения функции полнотекстового поиска и морфологического анализа текстов.

Тема 5. Словари и тезаурусы. (1ч.)

Универсальные среды для систем автоматической обработки русского текста. Словари исходных словоформ русского языка в виде zip-архивов.

Тема 6. Психолингвистические программы.(1ч.)

Программы психолингвистического анализа текстов. Лексический и контент-анализ текстов.

II. СТРУКТУРА И СОДЕРЖАНИЕ ПРАКТИЧЕСКОЙ ЧАСТИ КУРСА И САМОСТОЯТЕЛЬНОЙ РАБОТЫ

Практические занятия (0 часов)

Не предусмотрены

Лабораторные работы (36 час.)

Лабораторная работа №1. (18час.).

Квазиреферирование (статистические методы).

Лабораторная работа №2. (18час.).

Квазиреферирование (нейронные сети).

III. УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ САМОСТОЯТЕЛЬНОЙ РАБОТЫ ОБУЧАЮЩИХСЯ

Самостоятельная работа обучающихся подразумевает обязательную подготовку к лабораторным занятиям (оформление отчетов), изучение основной и дополнительно литературы по дисциплине, подготовку к текущему контролю и промежуточной аттестации в конце семестра, консультации преподавателей.

Студент должен планировать график самостоятельной работы по дисциплине и придерживаться его.

План-график выполнения самостоятельной работы по дисциплине

№ п/п	Дата/сроки выполнения	Вид самостоятельной работы	Примерные нормы времени на выполнение	Форма контроля
1.	1-3 недели	Работа с первоисточниками по выбранной теме	16ч.	Опрос (анализ и обсуждение проработанных первоисточников в ходе лабораторных работ)
2.	4- 7недели	ПО для ИПС	16ч.	Демонстрация в ходе лабораторной работы
3.	8-9 недели	Подготовка презентации и демонстрации работы с программами	4 ч.	Демонстрация в ходе лабораторной работы
4.	10-12недели	Работа с первоисточниками по выбранной теме	16 ч.	Опрос (анализ и обсуждение проработанных первоисточников в ходе лабораторных работ)
5.	13-16 недели	ПО для ИПС	16 ч.	Демонстрация в ходе лабораторной работы
6.	17-18 недели	Подготовка презентации и	4 ч.	Демонстрация в ходе лабораторной

		демонстрации работы с программами		работы
--	--	--------------------------------------	--	--------

Рекомендации по самостоятельной работе студентов

Самостоятельная работа студентов состоит в работе с литературой, подготовке к лабораторным работам и выполнении индивидуальных заданий по темам.

Работа с литературой

В процессе подготовки к лабораторным работам студентам необходимо обратить особое внимание на самостоятельное изучение рекомендованной учебно-методической литературы. Самостоятельная работа с учебниками, учебными пособиями, научной, справочной литературой, материалами Интернета является наиболее эффективным методом получения знаний, позволяет значительно активизировать процесс овладения информацией, способствует более глубокому усвоению изучаемого материала, формирует у студентов свое отношение к конкретной проблеме.

Подготовка к лабораторным работам

Подготовку к каждой лабораторной работе каждый студент должен начать с изучения теоретического материала и ознакомления с имеющимися реализациями.

Индивидуальные задания

По теме лабораторной работы студентам выдаются индивидуальные задания, предназначенные для закрепления умений и навыков, полученных при выполнении лабораторной работы. Задания выполняются самостоятельно. По каждому заданию предполагается подготовка индивидуального документа (проекта). Выполненный проект демонстрируется преподавателю. Структура индивидуального задания определяется темой лабораторной работы.

Подготовка презентации и доклада

Для подготовки презентации рекомендуется использовать: PowerPoint, MS Word, Acrobat Reader, LaTeX-овский пакет beamer. Последовательность подготовки презентации:

1. Четко сформулировать цель презентации: вы хотите свою аудиторию мотивировать, убедить, заразить какой-то идеей или просто формально отчитаться.

2. Определить каков будет формат презентации: живое выступление (тогда, сколько будет его продолжительность) или электронная рассылка (каков будет контекст презентации).

3. Отобрать всю содержательную часть для презентации и выстроить логическую цепочку представления.

4. Определить ключевые моменты в содержании текста и выделить их.

5. Определить виды визуализации (иллюстрации, образы, диаграммы, таблицы) для отображения их на слайдах в соответствии с логикой, целью и спецификой материала.

6. Подобрать дизайн и форматировать слайды (количество картинок и текста, их расположение, цвет и размер).

7. Проверить визуальное восприятие презентации.

Практические советы по подготовке презентации - готовьте отдельно:

- печатный текст + слайды + раздаточный материал;
- *слайды* – визуальная подача информации, которая должна содержать минимум текста, максимум изображений, несущих смысловую нагрузку, выглядеть наглядно и просто;
- *текстовое содержание презентации* – устная речь или чтение, которая должна включать аргументы, факты, доказательства и эмоции;
- *рекомендуемое число слайдов* 17-22;
- *обязательная информация для презентации*: тема, фамилия и инициалы выступающего; план сообщения; краткие выводы из всего сказанного; список использованных источников;
- *раздаточный материал* – должен обеспечивать ту же глубину и охват, что и живое выступление: люди больше доверяют тому, что они могут унести с собой, чем исчезающим изображениям, слова и слайды забываются, а раздаточный материал остается постоянным осязаемым напоминанием; раздаточный материал важно раздавать в конце презентации; раздаточный материалы должны отличаться от слайдов, должны быть более информативными.

Критерии оценки индивидуальных заданий (проектов)

– 100-86 баллов выставляется, если содержание и составляющие части соответствуют выданному заданию. Продемонстрировано владение навыками подготовки документа по теме. Фактических ошибок, связанных с пониманием структуры и содержания задания нет.

– 85-76 - баллов выставляется, если при выполнении задания допущено не более одной ошибки. Продемонстрировано владение навыками подготовки документа по теме. Фактических ошибок, связанных с пониманием структуры и содержания задания нет.

– 75-61 балл выставляется, если при выполнении задания допущено не более двух ошибок. Продемонстрировано знание и владение навыками подготовки документа по теме. Допущено не более 2 ошибок, связанных с пониманием структуры и содержания задания.

– 60-50 баллов - если структура и содержание задания не соответствуют требуемым.

Шкала оценивания

Менее 60 баллов	незачтено	неудовлетворительно
От 61 до 75 баллов	зачтено	удовлетворительно
От 76 до 85 баллов	зачтено	хорошо
От 86 до 100 баллов	зачтено	отлично

IV. КОНТРОЛЬ ДОСТИЖЕНИЯ ЦЕЛЕЙ КУРСА

№ п/п	Контролируемые разделы/темы дисциплины	Коды и этапы формирования компетенций	Оценочные средства		
			текущий контроль	промежуточная аттестация	
1	Раздел.1 Компьютерная лингвистика.	ОПК-2 ОПК-7. ПК-4.	знает	Конспект (ПР-7)	Вопросы к зачету 1-17
			умеет владеет	Лабораторные работы (практические занятия) № 1,2 (ПР-6) Доклад УО-3	
2	Раздел.2 Модель лингвистической онтологии для автоматической обработки текстов	ОПК-2 ОПК-7. ПК-4.	знает	Конспект (ПР-7)	Вопросы к зачету 18-26
			умеет владеет	Лабораторные работы № 1,2 (ПР-6) Доклад УО-3	

3	Раздел 3. Примеры систем	ОПК-2 ОПК-7. ПК-4.	знает	Конспект (ПР-7)	Вопросы к зачету 27-35
			умеет владеет	Собеседование (УО-1)	

Типовые задания, методические материалы, определяющие процедуры оценивания знаний, умений и навыков и (или) опыта деятельности, а также критерии и показатели, необходимые для оценки знаний, умений, навыков и характеризующие этапы формирования компетенций в процессе освоения образовательной программы, представлены в разделе VI.

V. СПИСОК УЧЕБНОЙ ЛИТЕРАТУРЫ И ИНФОРМАЦИОННО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Основная литература

(электронные и печатные издания)

1. Введение в прикладную лингвистику: учебное пособие / А. Н. Баранов; Московский государственный университет, Филологический факультет. А. Н. Баранов. - М.: URSS [Либроком], 2013. – 367 с. <https://lib.dvfu.ru:8443/lib/item?id=chamo:779256&theme=FEFU>
2. Автоматическое понимание текстов: Системы, модели, ресурсы: Учебное пособие / Н.Н. Леонтьева. – М.: Академия, 2006г.-304с. <https://lib.dvfu.ru:8443/lib/item?id=chamo:235710&theme=FEFU>
3. Тезаурусы в задачах информационного поиска. / Н.В. Лукашевич - М.: Изд-во Московского университета, 2011. -512с. <https://lib.dvfu.ru:8443/lib/item?id=chamo:403174&theme=FEFU>
4. Онтологии и тезаурусы: модели, инструменты, приложения: Учебное пособие. / Б.В. Добров, В.В. Иванов, Н.В. Лукашевич, В.Д. Соловьев. - М.: Интернет-Университет Информационных Технологий; БИНОМ. Лаборатория знаний, 2009. - 173 с.: ил. - (Серия "Основы информационных технологий"). <http://www.iprbookshop.ru/22417>

5. Онтологическая семантика. Знания. Онтологии. Онтологически ориентированные методы информационного анализа текстов [Электронный ресурс] / Рубашкин В.Ш. - М. : ФИЗМАТЛИТ, 2012. – 348 с. <http://www.studentlibrary.ru/book/ISBN9785922114394.html>
6. Гаврилова, Т. А. Интеллектуальные технологии в менеджменте: инструменты и системы [Электронный ресурс] : Учеб. пособие. 2-е изд. / Т. А. Гаврилова, Д. И. Муромцев; Высшая школа менеджмента СПбГУ : СПб.: Изд-во «Высшая школа менеджмента»; Издат. дом С.-Петербур. гос. ун-та, 2008. . 488 с. - ISBN 978-5-9924-0017-5 <http://znanium.com/bookread2.php?book=504514>

Дополнительная литература

1. Интернет-коммуникация как новая речевая формация / [Агагюлова С. И., Галичкина Е. Н., Горошко Е. И. и др.; научн. ред.: Т.Н.Колокольцева, О.В. Луговина] М.: Флинта Наука, 2012. - 323 с. <https://lib.dvfu.ru:8443/lib/item?id=chamo:675726&theme=FEFU>
2. Информационные технологии в лингвистике / А.В. Зубов, И.И. Зубова. – М.: Академия, 2004.- 207с. <https://lib.dvfu.ru:8443/lib/item?id=chamo:245160&theme=FEFU>
3. Компьютерная семантика русского языка / В. А. Тузов - СПб.: Изд-во СПбГУ, 2003. – 400 с.
4. Компьютерная лингвистика. Учебное пособие / Марчук Ю.Н. - М.:Изд.АСТ, 2007. - 317 с.
5. Корпусная лингвистика / Е.В. Грудева. - М.: ФЛИНТА, 2012. - 165 с.
6. Начала компьютерной лингвистики: Учеб. пособие. / Ю.И. Шемякин. - М.: Изд-во МГОУ, А/О "Росвузнаука", 1992. - 113с. <https://lib.dvfu.ru:8443/lib/item?id=chamo:37257&theme=FEFU>
7. Автоматическая обработка текстов на естественном языке и компьютерная лингвистика / Е.И. Большакова, Э.С.Клышинский, Д.В. Ландэ, А.А.Носков, О.В. Пескова, Е.В.- Ягунова М.: МИЭМ, 2011. - 272с. <http://window.edu.ru/library/pdf2txt/465/78465/59324>
8. Введение в компьютерную лингвистику. Учебное пособие. / К.К. Боярский. - СПб: НИУ ИТМО, 2013. - 72 с. <http://window.edu.ru/resource/387/80387>

Текущие публикации в следующих изданиях:

1. Журнал «Вопросы языкознания».
2. Журнал «Известия Академии Наук, сер. Литературы и языка.»
3. Журнал «ComputationalLinguistics (CL)».
4. Журнал «Computational Linguistics in the Netherlands (CLIN)».
5. Журнал «Linguistic Issues in Language Technology (LiLT)».
6. Журнал «Web journal of Formal, Computational & Cognitive Linguistics»
[электронный ресурс] .-Режим доступа:<http://fccl.ksu.ru>

Перечень ресурсов информационно-телекоммуникационной сети «Интернет»:

1. <http://dxdy.ru/topic4210-60.html> Научный форум dxdy. Моделирование русского языка
2. <http://lingvoforum.net/index.php?topic=29965.0>
Лингвофорум>>Компьютерная лингвистика
3. <http://forum.dialog-21.ru/actualthread.aspx?tid=363> Форум Диалог
4. http://uisrussia.msu.ru/linguist/_B_comput_ling.jsp Научно-образовательный портал "Лингвистика в России: ресурсы для исследователей"
5. <http://mytts.forum2x2.ru/f45-forum> Форум myTTS
6. <http://www.antic-r.ru/bibl5.htm> Книги по линггвистике
7. <http://window.edu.ru/library/pdf2txt/465/78465/59324> Автоматическая обработка текстов на естественном языке и компьютерная лингвистика / Е.И. Большакова, Э.С.Клышинский, Д.В. Ландэ, А.А.Носков, О.В. Пескова, Е.В.-Ягунова М.: МИЭМ, 2011г. -272 с.
8. <http://window.edu.ru/resource/387/80387> Введение в компьютерную лингвистику. Учебное пособие. / К.К, Боярский. - СПб: НИУ ИТМО, 2013. - 72 с.

Перечень информационных технологий и программного обеспечения

1. Операционная система Windows.
2. Microsoft Office.
3. Интернет.
4. Лингвистические программы: ClearTextReader'sEdition, Словогрыз, xReplacer; RussianMorphologicalDictionary, Рабочее Место Лингвиста, Galaktika-ZOOM; SakramentText-to-SpeechEnginev2.0, Аки Росс, Письмовник;

NeurOKSemanticSuite, ARMEngine 4.0, Следопыт 3.0; Гиперсловарь Ариадна, Словарь сокращений русского языка, Обратные и частотные словари русского языка; Гиперсловарь Ариадна, Словарь сокращений русского языка, Обратные и частотные словари русского языка.

При осуществлении образовательного процесса студентами и профессорско-преподавательским составом используются следующие информационно-справочные системы:

1. Научная электронная библиотека eLIBRARY.
2. Электронно-библиотечная система издательства «Лань».
3. Электронная библиотека "Консультант студента".
4. Электронно-библиотечная система IPRbooks.
5. Информационная система "ЕДИНОЕ ОКНО доступа к образовательным ресурсам".
6. Доступ к электронному заказу книг в библиотеке ДВФУ, доступ к нормативным документам ДВФУ, расписанию, рассылке писем.

Лекции проводятся с использованием проектора и мультимедийного комплекса для проведения лекций внутренней системы портала ДВФУ. Лабораторные занятия проводятся в специализированном компьютерном классе.

VI. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ

Дисциплина изучается в следующих организационных формах: лекции, лабораторное занятие; самостоятельное изучение теоретического материала; самостоятельное выполнение индивидуального задания; индивидуальные консультации.

Работа на лекции

Слушание и запись лекций – сложный вид вузовской аудиторной работы. Внимательное слушание и конспектирование лекций предполагает интенсивную умственную деятельность студента. Краткие записи лекций, их конспектирование помогает усвоить учебный материал. Конспект является полезным тогда, когда записано самое существенное, основное и сделано это самим студентом. Не надо стремиться записать дословно всю лекцию. Такое «конспектирование» приносит больше вреда, чем пользы. Запись лекций рекомендуется вести по возможности собственными формулировками. Желательно запись осуществлять на одной странице, а следующую оставлять для проработки учебного материала самостоятельно в домашних условиях.

Конспект лекции лучше подразделять на пункты, параграфы, соблюдая красную строку. Этому в большой степени будут способствовать пункты плана лекции, предложенные преподавателям. Принципиальные места, определения, формулы и другое следует сопровождать замечаниями «важно», «особо важно», «хорошо запомнить» и т.п. Можно делать это и с помощью разноцветных маркеров или ручек. Лучше если они будут собственными, чтобы не приходилось просить их у однокурсников и тем самым не отвлекать их во время лекции. Целесообразно разработать собственную «маркографию» (значки, символы), сокращения слов. Не лишним будет и изучение основ стенографии. Работая над конспектом лекций, всегда необходимо использовать не только учебник, но и ту литературу, которую дополнительно рекомендовал лектор. Именно такая серьезная, кропотливая работа с лекционным материалом позволит глубоко овладеть знаниями.

Лабораторные работы

При выполнении лабораторной работы необходимо следовать методическим рекомендациям по ее выполнению. Результатом лабораторной работы является программа и созданный документ, который демонстрируется преподавателю в конце работы. Студент должен уметь отвечать на вопросы преподавателя, поясняя процесс создания документа и выполнения работы.

Самостоятельная работа студента

Основной формой самостоятельной работы студента является изучение теоретического материала, его дополнение рекомендованной литературой, выполнение индивидуальных заданий, а также активная работа на лабораторных занятиях. Целью выполнения индивидуальных заданий является закрепление практических навыков, полученных в процессе выполнения лабораторной работы.

Контроль за выполнением самостоятельной работы студента производится в виде контроля каждого этапа работы (см. План-график). Студент должен планировать график самостоятельной работы по дисциплине и придерживаться его.

VII. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Лекции проводятся с использованием проектора и внутренней системы портала ДВФУ. Лабораторные занятия проходят в аудиториях, оборудованных компьютерами типа Lenovo C360G-i34164G500UDK с лицензионными программами Microsoft Office 2013 и аудиовизуальными средствами проектор Panasonic DLPPjectorPT-D2110XE, плазма LG FLATRON M4716CCBAM4716CJ. Для выполнения самостоятельной работы студенты в жилых корпусах ДВФУ обеспечены Wi-Fi.

VII. ФОНДЫ ОЦЕНОЧНЫХ СРЕДСТВ

Паспорт ФОС

Общепрофессиональные компетенции выпускников и индикаторы их достижения:

Наименование категории и (группы) компетенции	Код и наименование общепрофессиональной компетенции	Код и наименование индикатора достижения общепрофессиональной компетенции
	ОПК-2. Способен понимать принципы работы современных информационных технологий и программных средств, в том числе отечественного производства, и использовать их при решении задач профессиональной деятельности	ОПК-2.1. Знает типичные программные системы (приложения) в области АОТ и их архитектурные особенности ОПК-2.2. Умеет разрабатывать системы, которые позволяли бы взаимодействовать с ЭВМ в конкретной проблемной области на естественном языке или каком-то его ограниченном варианте ОПК-2.3. Владеет навыками работы со специальными программными средствами автоматизированной обработки текстов

	<p>ОПК-7. Способен применять в практической деятельности основные концепции, принципы, теории и факты, связанные с информатикой</p>	<p>ОПК-7.1. Знает сферу применения и перспективы развития информационно-поисковых тезаурусов, применение тезаурусного метода поиска в Интернет существенные отличия естественных языков от искусственных и особенности компьютерных моделей естественного языка, какие лингвистические технологии следует применять в зависимости от условий задачи, понимать, на какие лингвистические данные и ресурсы может опираться</p> <p>ОПК-7.2. Умеет интерпретировать результаты автоматической обработки лингвистических данных разрабатывать системы, которые позволяли бы взаимодействовать с ЭВМ в конкретной проблемной области на естественном языке или каком-то его ограниченном</p> <p>ОПК-7.3. Владеет методикой использования компьютерных технологий в предметной области приемами работы с прикладным программным обеспечением</p>
--	---	---

Профессиональные компетенции выпускников и индикаторы их достижения:

Задача профессиональной деятельности	Объект или область знания	Код и наименование профессиональной компетенции	Код и наименование индикатора достижения профессиональной компетенции	Основание (ПС, анализ иных требований, предъявляемых к выпускникам
Тип задач профессиональной деятельности: научно-исследовательский				

<p>Участие в научно-исследовательских и опытно-конструкторских работах в области программной инженерии. Анализ и выбор программно-технологических платформ, сервисов и информационных ресурсов программной инженерии; подготовка обзоров, аннотаций, составление рефератов и докладов, публикаций и библиографии по научно-исследовательской работе в области программной инженерии</p>	<p>Прикладные и информационные процессы. Информационные технологии и. Программное обеспечение</p>	<p>ПК-4. Готовность к использованию методов и инструментальных средств исследования объектов профессиональной деятельности</p>	<p>ПК-4.1. Знает технологию автоматизированной обработки текстовой информации, основные принципы представления знаний о предметной области в виде рубрикаторов, тезаурусов, онтологий особенности реализации естественно-языковых систем, назначение лингвистического ПК-4.2. Умеет проектировать процесс обработки языковых данных, работать с данными: понимать, какие именно данные нужны для решения поставленной задачи, как взаимодействуют лингвистические и иные данные, откуда их можно получить; определять основные классы ЕЯ-систем; ПК-4.3. Владеет широким диапазоном различных информационно-коммуникационных технологий; методами поиска информации по работе с естественным языком</p>	<p>06.028 Системный программист 06.022 Системный аналитик</p>
---	---	--	---	--

№ п/п	Контролируемые разделы/темы дисциплины	Коды и этапы формирования компетенций	Оценочные средства		
			текущий контроль	промежуточная аттестация	
1	Раздел.1 Компьютерная лингвистика.	ОПК-2 ОПК-7. ПК-4.	знает	Конспект (ПР-7)	Вопросы к зачету 1-17
			умеет владеет	Лабораторные работы (практические занятия) № 1,2 (ПР-6) Доклад УО-3	
2	Раздел.2 Модель лингвистической онтологии для автоматической обработки текстов	ОПК-2 ОПК-7. ПК-4.	знает	Конспект (ПР-7)	Вопросы к зачету 18-26
			умеет владеет	Лабораторные работы № 1,2 (ПР-6) Доклад УО-3	
3	Раздел 3. Примеры систем	ОПК-2 ОПК-7. ПК-4.	знает	Конспект (ПР-7)	Вопросы к зачету 27-35
			умеет владеет	Собеседование (УО-1)	

Шкала оценивания уровня сформированности компетенций

Код и формулировка компетенции	Этапы формирования компетенции		критерии	показатели
ОПК-2. Способен использовать	знает (пороговый)	типичные программные системы	знание архитектуры	способность перечислить особенности архитектуры

<p>ь современных информационных технологии и программные средства, в том числе отечественного производства, при решении задач профессиональной деятельности;</p>	уровень)	(приложения) в области АОТ и их архитектурные особенности;	систем АОТ;	систем АОТ;
	умеет (продвинутой)	разрабатывать системы, которые позволяли бы взаимодействовать с ЭВМ в конкретной проблемной области на естественном языке или каком-то его ограниченном варианте;	умение разрабатывать системы, работающие на ЕЯ;	способность выбрать подходящие программные средства для выполнения разработок;
	владеет (высокий)	навыками работы со специальными программными средствами автоматизированной обработки текстов ;	владение набором операций, предоставляемых современными автоматизированной обработки текстов;	способность выполнить практические задания по работе с специальными программными средствами автоматизированной обработки текстов, продемонстрировать результаты и пояснить процесс выполнения задания;
<p>ОПК-7. Способен применять в практической деятельности и основные концепции, принципы, теории и факты, связанные с информатикой;</p>	знает (пороговый уровень)	сферу применения и перспективы развития информационно-поисковых тезаурусов, применение тезаурусного метода поиска в Интернет существенно	знание основные принципы разработки, создания и использования традиционных информационно-поисковых тезаурусов;	способность перечислить технические и программные средства, предназначенные для работы с тезаурусами;

		<p>ые отличия естественных языков от искусственных и особенности компьютерных моделей естественного языка, какие лингвистические технологии следует применять в зависимости от условий задачи, понимать, на какие лингвистические данные и ресурсы может опираться;</p>		
	<p>умеет (продвинутой)</p>	<p>интерпретировать результаты автоматической обработки лингвистических данных разрабатывать системы, которые позволяли бы взаимодействовать с ЭВМ в конкретной проблемной области на естественном языке или каком-то его</p>	<p>умение работать с корпусом как репрезентативная выборкой текста;</p>	<p>способность выполнить разметку корпуса;</p>

		ограниченном;		
	владеет (высокий)	методикой использования компьютерных технологий в предметной области приемами работы с прикладными программам обеспечением;	владение методикой работы с технологиями обработки ЕЯ;	способность выполнить практические задания по работе с технологиями обработки ЕЯ;
ПК-4. Готовность к использованию методов и инструментальных средств исследования объектов профессиональной деятельности;	знает (пороговый уровень)	технологии автоматизированной обработки текстовой информации, основные принципы представления знаний о предметной области в виде рубрикаторов, тезаурусов, онтологий особенности реализации естественно-языковых систем, назначение лингвистического;	знание представлений знаний о предметной области. Знание о тезаурусах, онтологиях;	способность перечислить и раскрыть суть технологии автоматизированной обработки текстовой информации. Способность определить подходящие онтологии;
	умеет (продвинутой)	проектировать процесс обработки языковых данных, работать с данными;	умение работать с лингвистическими данными;	способность получать лингвистические данные;

		понимать, какие именно данные нужны для решения поставленной задачи, как взаимодействуют лингвистические и иные данные, откуда их можно получить; определять основные классы ЕЯ-систем;		
	владеет (высокий)	широким диапазоном различных информационных технологий; методами поиска информации по работе с естественным языком;	владение различными информационными технологиями;	способность работать с различными информационно-коммуникационными технологиями;

Методические рекомендации, определяющие процедуры оценивания результатов освоения дисциплины

Текущая аттестация студентов. Текущая аттестация студентов проводится в соответствии с локальными нормативными актами ДВФУ и является обязательной.

Текущая аттестация проводится в форме вопросов к зачету для проверки теоретических знаний, а также в форме защиты проекта, выполняемого в рамках самостоятельной работы параллельно с

лабораторными работами, и осуществляется ведущим преподавателем.

Объектами оценивания выступают:

– степень усвоения теоретических знаний - оценивается в форме вопросов к зачету;

- уровень овладения практическими умениями и навыками оценивается в форме защиты индивидуального задания (проекта) и по результатам выполнения лабораторных работ.

Промежуточная аттестация студентов. Промежуточная аттестация студентов проводится в соответствии с локальными нормативными актами ДВФУ и является обязательной.

По дисциплине предусмотрены зачет, которые проводятся в устной форме, с учетом защиты индивидуального задания (проекта) и результатам выполнения лабораторных работ.

Критерии выставления оценки студенту на зачете/экзамене

Баллы (рейтингов ой оценки)	Оценка зачета (стандартная)	Требования к сформированным компетенциям
86-100	«зачтено»/ «отлично»	Оценка «зачтено» / «отлично» выставляется студенту, если он глубоко и прочно усвоил программный материал, исчерпывающе, последовательно, четко и логически стройно его излагает, умеет тесно увязывать теорию с практикой, свободно справляется с задачами, вопросами и другими видами применения знаний, причем не затрудняется с ответом при видоизменении заданий, правильно обосновывает принятое решение, владеет разносторонними навыками и приемами выполнения практических задач.
76-85	«зачтено»/ «хорошо»	Оценка «зачтено» / «хорошо» выставляется студенту, если он твердо знает материал, грамотно и по существу излагает его, не допуская существенных неточностей в ответе на вопрос, правильно применяет теоретические положения при решении практических вопросов и задач, владеет необходимыми навыками и приемами их выполнения.

61-75	«зачтено»/ «удовлетворительно»	Оценка «зачтено» / «удовлетворительно» выставляется студенту, если он имеет знания только основного материала, но не усвоил его деталей, допускает неточности, недостаточно правильные формулировки, нарушения логической последовательности в изложении программного материала, испытывает затруднения при выполнении практических работ.
0-60	«не зачтено»/ «неудовлетворительно»	Оценка «не зачтено» / «неудовлетворительно» выставляется студенту, который не знает значительной части программного материала, допускает существенные ошибки, неуверенно, с большими затруднениями выполняет практические работы. Как правило, оценка «неудовлетворительно» ставится студентам, которые не могут продолжить обучение без дополнительных занятий по соответствующей дисциплине.

ОЦЕНОЧНЫЕ СРЕДСТВА ДЛЯ ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ ВОПРОСЫ К ЗАЧЕТУ

по дисциплине «МЕТОДЫ РАБОТЫ С ЕСТЕСТВЕННЫМ ЯЗЫКОМ»

Раздел.1 Компьютерная лингвистика.

1. Задачи компьютерной лингвистики.
2. Особенности системы ЕЯ: уровни и связи.
3. Моделирование в компьютерной лингвистике.
4. Автоматическая обработка звучащей речи.
5. Системы распознавания речи.
6. Механизмы распознавания речи.
7. Автоматический синтез речи.
8. Автоматическое индексирование .
9. Создание словарей, конкордансов, частотных словарей.
10. Автоматическое реферирование.

11. Автоматическое порождение текста.
12. Получение компрессированного текста.
13. Основные характеристики текста.
14. Модель "смысл \Leftrightarrow текст".
15. Программные средства лингвистической обработки.
16. Представление лингвистических данных. Корпусная лингвистика.
17. Архитектура инструментальных ЕЯ-систем.

Раздел.2 Модель лингвистической онтологии для автоматической обработки текстов

18. Назначение онтологий.
19. Онтологии предметных областей и прикладные онтологии.
20. Лексические онтологии.
21. Языки описания онтологий.
22. Основные синтаксические структуры: классы, отношения, аксиомы.
23. SemanticWeb.
24. Структура информационно-поискового тезауруса.
25. Построение информационно-поисковых тезаурусов.
26. Тезаурус для автоматического концептуального индексирования.

Раздел.3 Примеры систем

27. Алгоритмы реферирования - метод «нижняя граница», позиция, IF-IDF.
28. Алгоритмы реферирования - IF-IDF + позиция, центрирование, центрирование+позиция.
29. Алгоритмы реферирования - метод с учителем, сингулярное разложение матриц.
30. Информационно-поисковые системы (ИПС).
31. Информационно-поисковые языки (на ограниченном естественном языке).
32. Индексирование. Поисковый образ документа.
33. Поисковый образ запроса и поисковое предписание.

34.Алгоритмы информационного поиска.

35.Программное обеспечение для информационно-поисковых систем.

Критерии оценки:

86-100 баллов - выставляется студенту, если вопрос полностью раскрыт без ошибок и даны ответы на все дополнительные вопросы

76-85 баллов - выставляется студенту, если вопрос полностью раскрыт без ошибок и даны ответы не на все дополнительные вопросы

61-75 баллов - выставляется студенту, если вопрос не полностью раскрыт без ошибок и даны ответы на все дополнительные вопросы

50-60 баллов - выставляется студенту, если вопрос не полностью раскрыт без ошибок и даны ответы не на все дополнительные вопросы

ОЦЕНОЧНЫЕ СРЕДСТВА ДЛЯ ТЕКУЩЕЙ АТТЕСТАЦИИ

Текущий контроль

Текущий контроль предполагает систематическую проверку усвоения учебного материала, сформированности компетенций или их элементов, регулярно осуществляемую на протяжении изучения дисциплины, в соответствии с ее рабочей программой.

Состоит в проверке правильности выполнения заданий по самостоятельной работе. Задание зачтено, если нет ошибок. По текущим ошибкам даются пояснения.

Тесты предназначены для проверки знаний по компетенциям. Проверка достижения умений и навыков по компетенциям проверяется выполнением практических и лабораторных работ.

Примерные тесты для проверки сформированности компетенций

ОПК-2. Способен понимать принципы работы современных информационных технологий и программных средств, в том числе отечественного производства, и использовать их при решении задач профессиональной деятельности	Знает типичные программные системы (приложения) в области АОТ и их архитектурные особенности
1.Укажите слои ЕЯ-системы	ответы

	а) слой морфологический б) слой взаимодействия в) слой данных г) слой интерпретации
2. Системы обработки ЕЯ-текстов	ответы а) системы на базе разметки б) системы на базе фреймов в) системы на базе аннотаций

ОПК-7. Способен применять в практической деятельности основные концепции, принципы, теории и факты, связанные с информатикой	Знает сферу применения и перспективы развития информационно-поисковых тезаурусов, применение тезаурусного метода поиска в Интернет существенные отличия естественных языков от искусственных и особенности компьютерных моделей естественного языка, какие лингвистические технологии следует применять в зависимости от условий задачи, понимать, на какие лингвистические данные и ресурсы может опираться
1. Основной единицей тезаурусов являются термины, которые разделяются на:	ответы а) дескрипторы б) прескрипторы в) аскрипторы
2. Тезаурус для автоматического индексирования должен включать	ответы а) термины б) специфических терминов в) универсальные термины

ПК-4. Готовность к использованию методов и инструментальных средств исследования объектов профессиональной деятельности	Знает технологию автоматизированной обработки текстовой информации, основные принципы представления знаний о предметной области в виде рубрикаторов, тезаурусов, онтологий особенности реализации естественно-языковых систем, назначение лингвистического
1. Индексирование документа - это процедура отображения текста документа в определенную форму, предназначенную для автоматической обработки (индекс документа). Укажите какого вида индексирование документа не существует	ответы а) автоматическое индексирование б) статистическое индексирование в) ручное индексирование

<p>2.Разработка лингвистического процессора (ЛП) предполагает описание лингвистических свойств обрабатываемого текста ЕЯ, и это описание организуется как</p>	<p>ответы</p> <p>а) модель языка</p> <p>б) синтаксическое дерево</p> <p>в) сеть</p>
---	--

Критерии оценки докладов и собеседования

- 100-86 баллов выставляется, если бакалавр/группа точно определили содержание и составляющие части задания, умеют аргументировано отвечать на вопросы, связанные с заданием. Продемонстрировано знание и владение навыками самостоятельной исследовательской работы по теме. Фактических ошибок, связанных с пониманием проблемы, нет.

- 85-76 - баллов - работа бакалавра/группы характеризуется смысловой цельностью, связностью и последовательностью изложения; допущено не более 1 ошибки при объяснении смысла или содержания проблемы. Продемонстрированы исследовательские умения и навыки. Фактических ошибок, связанных с пониманием проблемы, нет.

- 75-61 балл – проведен достаточно самостоятельный анализ основных этапов и смысловых составляющих проблемы; понимание базовых основ и теоретического обоснования выбранной темы. Привлечены основные источники по рассматриваемой теме. Допущено не более 2 ошибок в смысле или содержании проблемы

- 60-50 баллов - если работа представляет собой пересказанный или полностью переписанный исходный текст без каких бы то ни было комментариев, анализа. Не раскрыта структура и теоретическая составляющая темы. Допущено три или более трех ошибок смыслового содержания раскрываемой проблемы

Шкала оценивания докладов и собеседования

Менее 60 баллов	Не зачтено
От 61 до 75 баллов	зачтено
От 76 до 85 баллов	зачтено
От 86 до 100 баллов	зачтено