

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РФ
Федеральное государственное бюджетное образовательное
учреждение высшего образования
«Кубанский государственный университет»
(ФГБОУ ВО «КубГУ»)

Факультет компьютерных технологий и прикладной математики
Кафедра вычислительных технологий

УТВЕРЖДАЮ:
Проректор по учебной работе,
качеству образования – первый
проректор
_____ Хагуров Т.А.
05 2022 г.



РАБОЧАЯ ПРОГРАММА УЧЕБНОЙ ДИСЦИПЛИНЫ
Б1.В.03 «СПЕЦСЕМИНАР»

Направление подготовки 02.04.02 Фундаментальная информатика и
информационные технологии

(код и наименование направления подготовки/специальности)

Направленность (профиль) "Интеллектуальные системы и технологии"
(наименование направленности (профиля) специализации)

Программа подготовки академическая
(академическая /прикладная)

Форма обучения очная
(очная, очно-заочная, заочная)

Квалификация (степень) выпускника магистр
(бакалавр, магистр, специалист)

Краснодар 2022

Рабочая программа дисциплины Б1.В.03 – Спецсеминар составлена в соответствии с федеральным государственным образовательным стандартом высшего образования (ФГОС ВО) по направлению подготовки 02.04.02 «Фундаментальная информатика и информационные технологии»

Программу составил(и):

Вишняков Ю.М., зав. кафедрой вычислительных технологий, д.т.н., профессор



подпись

Рабочая программа дисциплины Б1.В.03 – Спецсеминар утверждена на заседании кафедры вычислительных технологий, протокол №9 от «18» мая 2022 г.

Заведующий кафедрой (разработчик) Вишняков Ю.М.

фамилия, инициалы



подпись

Утверждена на заседании учебно-методической комиссии факультета компьютерных технологий и прикладной математики, протокол №6 от «25» мая 2022 г.

Председатель УМК факультета



А.В. Коваленко

Рецензенты:

Гаркуша О.В., доцент кафедры информационных технологий ФБГОУ ВО «Кубанский государственный университет», кандидат физико-математических наук.

Схаляхо Ч.А., доцент КВВУ им. С.М. Штеменко, кандидат физико-математических наук, доцент

1. ЦЕЛИ И ЗАДАЧИ ОСВОЕНИЯ ДИСЦИПЛИНЫ³

1.1. Цель освоения дисциплины

Цель дисциплины Б1.В.03 – Спецсеминар – дать студентам знания по автоматической обработке естественных языков и типовым программным приложениям, научить решать комплексные задачи в области обработки нечисловой информации.

1.2. Задачи дисциплины

В результате освоения дисциплины должны быть решены следующие основные задачи. Студент должен:

- знать основные подходы по моделям и автоматической обработке естественных языков, а также их приложениям в обработке текстовой информации на естественном языке, приобрести навыки решения комплексных задач в области проектирования систем обработки нечисловой информации.

- уметь применять знания по основным подходам построения моделей и автоматической обработке естественных языков, а также их приложениям в обработке текстовой информации на естественном языке в области проектирования систем обработки нечисловой информации.

- владеть восприятием, анализом и обобщением информации в профессиональной области и выбором путей решения профессиональных задач на основе знаний и умений дисциплины «Спецсеминар».

1.3. Место дисциплины (модуля) в структуре образовательной программы

Дисциплина Б1.В.03 Спецсеминар относится к Блоку 1 "Дисциплины (модули)" учебного плана. В соответствии с рабочим учебным планом дисциплина изучается на 1 курсе в 1 и 2 семестрах. Вид промежуточной аттестации: зчет.

Преддисциплинами для дисциплины Б1.В.03 Спецсеминар являются дисциплины бакалавриата, обеспечивающие знание языков и методов программирования, дискретной математики, нечеткой логики, теории графов и комбинаторики, формальных грамматики и языков, программирования и его методов и технологий.

Дисциплина Б1.В.03 Спецсеминар представляет собой преддисциплину для таких дисциплин как «Интеллектуальные информационные системы и технологии», и других дисциплин, связанных с разработкой программного обеспечения систем различного назначения, научно-исследовательской работы, практик, выпускной квалификационной работы.

1.4. Перечень планируемых результатов обучения по дисциплине (модулю), соотнесенных с планируемыми результатами освоения образовательной программы.

Процесс освоения дисциплины направлен на формирование следующих компетенций:

Код и наименование индикатора*	Результаты обучения по дисциплине (знает, умеет, владеет (навыки и/или опыт деятельности))
УК-1. Способен осуществлять критический анализ проблемных ситуаций на основе системного подхода, вырабатывать стратегию действий	
УК-1.1: Знает принципы сбора, отбора и обобщения информации..	Знает принципы сбора, отбора и обобщения информации в рамках спецсеминара.
УК-1.2: Умеет соотносить разнородные явления и систематизировать их в рамках избранных видов профессиональной деятельности.	Умеет соотносить разнородные явления и систематизировать их в рамках избранных видов профессиональной деятельности в рамках спецсеминара.
УК-1.3: Имеет практический опыт работы с информационными объектами и сетью Интернет, опыт научного поиска, опыт библиографического разыскания, создания научных текстов.	Имеет практический опыт работы с информационными объектами и сетью Интернет, опыт научного поиска, опыт библиографического разыскания, создания научных текстов в рамках спецсеминара.
УК-4. Способен применять современные коммуникативные технологии, в том числе на иностранном(ых) языке(ах), для академического и профессионального взаимодействия	

Код и наименование индикатора*	Результаты обучения по дисциплине (<i>знает, умеет, владеет (навыки и/или опыт деятельности)</i>)
УК-4.1: Знает литературную форму государственного языка, основы устной и письменной коммуникации на иностранном языке, функциональные стили родного языка, требования к деловой коммуникации.	Знает литературную форму государственного языка, основы устной и письменной коммуникации на иностранном языке, функциональные стили родного языка, требования к деловой коммуникации в рамках спецсеминара.
УК-4.2: Умеет выражать свои мысли на государственном, родном и иностранном языке в ситуации деловой коммуникации.	Умеет выражать свои мысли на государственном, родном и иностранном языке в ситуации деловой коммуникации в рамках спецсеминара.
УК-4.3: Имеет практический опыт составления текстов разной функциональной принадлежности и разных жанров на государственном и родном языках, опыт перевода текстов с иностранного языка на родной, опыт говорения на государственном и иностранном языках.	Имеет практический опыт составления текстов разной функциональной принадлежности и разных жанров на государственном и родном языках, опыт перевода текстов с иностранного языка на родной, опыт говорения на государственном и иностранном языках в рамках спецсеминара.
ПК-4. Владеть навыками участия в научных дискуссиях, выступления с сообщениями и докладами, устно, письменного и виртуального (размещение в информационных сетях) характера; представления материалов собственных исследований; проводить корректуру, редактирование, реферирование работ	
ПК-4.1. Знает основы ведения научной дискуссии и формы устного научного высказывания.	Знает основы ведения научной дискуссии и формы устного научного высказывания в рамках спецсеминара.
ПК-4.2. Умеет вести корректную дискуссию в области информационных технологий, задавать вопросы и отвечать на поставленные вопросы по теме научной работы.	Умеет вести корректную дискуссию в области информационных технологий, задавать вопросы и отвечать на поставленные вопросы по теме научной работы в рамках спецсеминара.
ПК-4.3. Имеет практический опыт участия в научных студенческих конференциях, очных, виртуальных, заочных обсуждениях научных проблем в области информационных технологий	Имеет практический опыт участия в научных студенческих конференциях, очных, виртуальных, заочных обсуждениях научных проблем в области информационных технологий в рамках спецсеминара.
ПК-5. Способность приобретать и использовать организационно- управленческие навыки в конкретной профессиональной и социальной деятельности; разрабатывать, реализовывать и управлять процессами жизненного цикла программных продуктов.	
ПК-5.1. Знает основы разработки и реализации процессов жизненного цикла программного обеспечения.	ПК-5.1. Знает основы разработки и реализации процессов жизненного цикла программного обеспечения в рамках спецсеминара.
ПК-5.2. Умеет приобретать и использовать организационно- управленческие навыки в профессиональной и социальной деятельности.	ПК-5.2. Умеет приобретать и использовать организационно- управленческие навыки в профессиональной и социальной деятельности в рамках спецсеминара.
ПК-5.3. Имеет практический опыт управления процессами жизненного цикла программных продуктов.	ПК-5.3. Имеет практический опыт управления процессами жизненного цикла программных продуктов в рамках спецсеминара..

Результаты обучения по дисциплине достигаются в рамках осуществления всех видов контактной и самостоятельной работы обучающихся в соответствии с утвержденным учебным планом.

Индикаторы достижения компетенций считаются сформированными при достижении соответствующих им результатов обучения.

2. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

2.1. Распределение трудоёмкости дисциплины по видам работ

Общая трудоёмкость дисциплины составляет 5 зач.ед. (180 часов), их распределение по видам работ представлено в таблице (для студентов ОФО)

Вид работы	Всего	Форма обучения
------------	-------	----------------

	часов	Очная		очно-заочная	заочная
		1 семестр (часы)	2 семестр (часы)	X семестр (часы)	X курс (часы)
Контактная работа в том числе:	78,4	36,2	42,2		
Аудиторные занятия (всего):	64	36	28		
В том числе:					
Занятия лекционного типа					
Занятия семинарского типа (семинары, практ. занятия)					
Лабораторные занятия	64	36	28		
Иная контрольная работа	14,4	0,2	14,2		
Контроль самостоятельной работы	14	0	14		
Промежуточная аттестация (ИКР)	0,4	0,2	0,2		
Самостоятельная работа, в том числе	101,6	71,8	29,8		
В том числе:					
Курсовая работа	20		20		
<i>Проработка учебного (теоретического) материала</i>	56	50	6		
<i>Выполнение индивидуальных заданий (подготовка сообщений, презентаций)</i>	20	20			
<i>Реферат</i>					
<i>Подготовка к текущему контролю</i>	4,6	1,8	3,8		
Контроль: зачет					
Общая трудоемкость	в час	180	108	72	
	в т.ч. контактная работа	78,4	36,2	42,2	

2.2. Содержание дисциплины:

Распределение видов учебной работы и их трудоемкости по разделам дисциплины.
Разделы дисциплины, изучаемые в 1 семестре (очная форма)

№ раздела	Наименование разделов	Количество часов				
		Всего	Аудиторная работа			Внеаудиторная работа
			Л	ПЗ	ЛР	
1	2	3	4	5	6	7
1.	Естественный язык. Устная и письменная речь. Речевые категории, структура речи..	9	3		4	2
2.	Грамматическая предложений. Синтаксические конструкции	10	3		4	3
3.	Словно-зависимые представления предложений. Дерево зависимостей.	12			4	8

№ раздела	Наименование разделов	Количество часов				
		Всего	Аудиторная работа			Внеауди- торная работа СРС
			Л	ПЗ	ЛР	
1	2	3	4	5	6	7
4.	Отношение подчинения. Проективные предложения. Основные характеристики и свойства.	12			4	8
5.	Отрезочное представление предложений. Система составляющих.	12			4	8
6.	Свойства и построение системы составляющих. Связь системы составляющих и проективных предложений.	12			4	8
7.	Модели семантики естественных языков. Частотная модель. Вероятностная модель. Грамматическая модель.	12			4	8
8.	Семантические меры близости фрагментов текстов. Частотная (TF IDF). Вероятностная. Грамматическая.	12			4	8
9.	Семантическая близость текстов и документов на основе различных мер близостей. Релевантность в информационном поиске.	11,8			4	7,8
	ИТОГО по разделам дисциплины	107,8			36	71,8
	Контроль самостоятельной работы (КСР)					
	Промежуточная аттестация (ИКР)	0,2				
	Подготовка к текущему контролю					
	Общая трудоемкость по дисциплине	108				

Разделы дисциплины, изучаемые в 2 семестре (очная форма)

№ раздела	Наименование разделов	Количество часов				
		Всего	Аудиторная работа			Внеауди- торная работа СРС
			Л	ПЗ	ЛР	
1	2	3	4	5	6	7
1.	Вычислительная интерпретационная модель понимания. Вычислительное представление смысла. Семантические схемы.	9			3	3
2.	Парсинг. Морфологический, синтаксический. Яндекс-парсер.	10			3	3
3.	Частотный анализ текстовой информации. Статистическая меры IT IDF для оценки важности слова в контексте документа. Определение уникальности текста на основе морфологического разбора слов.				3	3
4.	Средства языка Python для работы с текстами. Библиотека NLTK. Установка, настройка. Токенизация. Морфологический анализ. Синтаксический анализ.				3	3

№ раздела	Наименование разделов	Количество часов				
		Всего	Аудиторная работа			Внеаудитор ная работа
			Л	ПЗ	ЛР	
1	2	3	4	5	6	7
5.	Технологии разговорного искусственного интеллекта. Библиотека DeepPavlov. Настройка, и работа.	10			3	3
6.	Работа в DeepPavlov с индивидуальным dataset. Анализ выхода.	10			3	3
7.	Проект от Google нейросеть BERT. Принцип работы, до обучение. Кластеризация текстов.	13			3	3,8
8.	Машинный перевод. Векторизация. Encoding,	13			3	4
9.	Encoding, обучающие выборки и обучение. Характеристики, трудоемкость	13			4	4
	ИТОГО по разделам дисциплины	57,8			28	29,8
	Контроль самостоятельной работы (КСР)	14				
	Промежуточная аттестация (ИКР)	0,2				
	Подготовка к текущему контролю					
	Общая трудоемкость по дисциплине	72				

Примечание: Л – лекции, КСР – контрольные и самостоятельные работы, ЛР – лабораторные занятия, СРС – самостоятельная работа студента, Д-доклад, РГЗ – расчетно-графическое задание.

2.3. Содержание разделов дисциплины:

2.3.1. Занятия лекционного типа

Учебным планом не предусмотрены

2.3.2. Занятия семинарского типа (практические / семинарские занятия/ лабораторные работы)

1-й семестр

№	Наименование лабораторных работ	Форма текущего контроля
1	2	3
1.	Естественный язык. Устная и письменная речь. Речевые категории, структура речи.	Защита ЛР
2.	Грамматическая предложений. Синтаксические конструкции	Защита ЛР
3.	Словно-зависимые представления предложений. Дерево зависимостей.	Защита ЛР
4.	Отношение подчинения. Проективные предложения. Основные характеристики и свойства.	Защита ЛР
5.	Отрезочное представление предложений. Система составляющих.	Защита ЛР
6.	Свойства и построение системы составляющих. Связь системы составляющих и проективных предложений.	Защита ЛР
7.	Модели семантики естественных языков. Частотная модель. Вероятностная модель. Грамматическая модель.	Защита ЛР
8.	Семантические меры близости фрагментов текстов. Частотная (TF IDF). Вероятностная. Грамматическая.	Защита ЛР

9.	Семантическая близость текстов и документов на основе различных мер близостей. Релевантность в информационном поиске.	
----	---	--

2-й семестр

№	Наименование лабораторных работ	Форма текущего контроля
1	2	3
1.	Вычислительная интерпретационная модель понимания. Вычислительное представление смысла. Семантические схемы.	Защита ЛР
2.	Парсинг. Морфологический, синтаксический. Яндекс–парсер.	Защита ЛР
3.	Частотный анализ текстовой информации. Статистические меры IT IDF для оценки важности слова в контексте документа. Определение уникальности текста на основе морфологического разбора слов.	Защита ЛР
4.	Средства языка Python для работы с текстами. Библиотека NLTK. Установка, настройка. Токенизация. Морфологический анализ. Синтаксический анализ.	Защита ЛР
5.	Технологии разговорного искусственного интеллекта. Библиотека DeepPavlov. Настройка, и работа.	Защита ЛР
6.	Работа в DeepPavlov с индивидуальным dataset. Анализ выхода.	Защита ЛР
7.	Проект от Google нейросеть BERT. Принцип работы, до обучение. Кластеризация текстов.	Защита ЛР
8.	Машинный перевод. Векторизация. Encoding,	Защита ЛР
9.	Encoding, обучающие выборки и обучение. Характеристики, трудоемкость	

Защита лабораторной работы (ЛР), выполнение курсового проекта (КП), курсовой работы (КР), расчетно-графического задания (РГЗ), написание реферата (Р), эссе (Э), коллоквиум (К), тестирование (Т) и т.д.

2.3.3. Примерная тематика курсовых работ (проектов)

В рамках дисциплины предусмотрено выполнение курсовых работ. Темы курсовых работ выдаются в начале 2 семестра и обучающийся в течении изучения дисциплины работает над темой. Темы связаны либо с дисциплиной, либо в контексте дисциплины связаны с проводимыми обучающимся научными исследованиями, сформулированными научными руководителями.

2.4. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине (модулю)

№	Вид СРС	Перечень учебно-методического обеспечения дисциплины по выполнению самостоятельной работы
1	2	3
1.	Работа с лекционным материалом	1
2.	Изучение теоретического материала к лабораторным занятиям	1. Немченко, В.Н. Введение в языкознание [Текст] : учебник для вузов / В. Н. Немченко. - 2-е изд., перераб. и доп. - Москва : Юрайт, 2013. - 679 с. 2. Боярский К. К. Введение в компьютерную лингвистику. Учебное пособие. – СПб: НИУ ИТМО, 2013. – 72 с. 3. 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.
3.	Подготовка к зачету	Методические рекомендации по организации самостоятельной работы студентов утвержденные кафедрой вычислительных технологий, протокол № 14 от 14.06.2017 г.

Учебно-методические материалы для самостоятельной работы обучающихся из числа инвалидов и лиц с ограниченными возможностями здоровья (ОВЗ) предоставляются в формах, адаптированных к ограничениям их здоровья и восприятия информации:

Для лиц с нарушениями зрения:

- в печатной форме увеличенным шрифтом,
- в форме электронного документа,

Для лиц с нарушениями слуха:

- в печатной форме,
- в форме электронного документа.

Для лиц с нарушениями опорно-двигательного аппарата:

- в печатной форме,
- в форме электронного документа,

Данный перечень может быть конкретизирован в зависимости от контингента обучающихся.

3. ОБРАЗОВАТЕЛЬНЫЕ ТЕХНОЛОГИИ, ПРИМЕНЯЕМЫЕ ПРИ ОСВОЕНИИ ДИСЦИПЛИНЫ (МОДУЛЯ)

Сем естр	Вид занятия	Используемые интерактивные образовательные технологии	Количество часов
1	ЛР	Отрезочное представление предложений. Система составляющих.	2
		Свойства и построение системы составляющих. Связь системы составляющих и проективных предложений.	2
		Модели семантики естественных языков. Частотная модель. Вероятностная модель. Грамматическая модель.	2
		Семантические меры близости фрагментов текстов. Частотная (TF IDF). Вероятностная. Грамматическая.	2
		Семантическая близость текстов и документов на основе различных мер близостей. Релевантность в информационном поиске.	2
2	ЛР	Средства языка Python для работы с текстами. Библиотека NLTK. Установка, настройка. Токенизация. Морфологический анализ. Синтаксический анализ.	2
		Технологии разговорного искусственного интеллекта. Библиотека DeepPavlov. Настройка, и работа.	2
		Работа в DeepPavlov с индивидуальным dataset. Анализ выхода.	2

	Проект от Google нейросеть BERT. Принцип работы, до обучение. Кластеризация текстов.	2
<i>Итого:</i>		16

4. ОЦЕНОЧНЫЕ СРЕДСТВА ДЛЯ ТЕКУЩЕГО КОНТРОЛЯ УСПЕВАЕМОСТИ И ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ

Оценочные средства предназначены для контроля и оценки образовательных достижений обучающихся, освоивших программу учебной дисциплины «Спецсеминар».

Оценочные средства включает контрольные материалы для проведения **текущего контроля** в форме тестовых заданий, доклада-презентации по проблемным темам разделов дисциплины, разно уровневых заданий и промежуточной аттестации в форме вопросов и заданий к экзамену.

Структура оценочных средств для текущей и промежуточной аттестации

№ п/п	Код и наименование индикатора	Результаты обучения	Наименование оценочного средства	
			Текущий контроль	Промежуточная аттестация
	УК-1.1: Знает принципы сбора, отбора и обобщения информации..	Знает принципы сбора, отбора и обобщения информации в рамках спецсеминара.	опрос по теме, лабораторная работа	Вопросы на зачете 1-30
	УК-1.2: Умеет соотносить разнородные явления и систематизировать их в рамках избранных видов профессиональной деятельности.	Умеет соотносить разнородные явления и систематизировать их в рамках избранных видов профессиональной деятельности в рамках спецсеминара.	опрос по теме, лабораторная работа	Вопросы на зачете 1-30
	УК-1.3: Имеет практический опыт работы с информационными объектами и сетью Интернет, опыт научного поиска, опыт библиографического разыскания, создания научных текстов.	Имеет практический опыт работы с информационными объектами и сетью Интернет, опыт научного поиска, опыт библиографического разыскания, создания научных текстов в рамках спецсеминара.	опрос по теме, лабораторная работа	Вопросы на зачете 1-30
	УК-4.1: Знает литературную форму государственного языка, основы устной и письменной коммуникации на иностранном языке, функциональные стили родного языка, требования к деловой коммуникации.	Знает литературную форму государственного языка, основы устной и письменной коммуникации на иностранном языке, функциональные стили родного языка, требования к деловой коммуникации в рамках спецсеминара.	опрос по теме, лабораторная работа	Вопросы на зачете 1-30
	УК-4.2: Умеет выражать свои мысли на государственном, родном и иностранном языке в ситуации деловой коммуникации.	Умеет выражать свои мысли на государственном, родном и иностранном языке в ситуации деловой коммуникации в рамках спецсеминара.	опрос по теме, лабораторная работа	Вопросы на зачете 1-30
	УК-4.3: Имеет практический опыт составления текстов разной функциональности принадлежности и разных жанров на государственном и родном языках, опыт перевода текстов с иностранного языка на родной, опыт говорения на государственном и иностранном языках.	Имеет практический опыт составления текстов разной функциональности принадлежности и разных жанров на государственном и родном языках, опыт перевода текстов с иностранного языка на родной, опыт говорения на государственном и иностранном языках в рамках спецсеминара.	опрос по теме, лабораторная работа	Вопросы на зачете 1-30
	ПК-4.1. Знает основы ведения научной дискуссии и формы устного научного высказывания.	Знает основы ведения научной дискуссии и формы устного научного высказывания в рамках спецсеминара.	опрос по теме, лабораторная работа	Вопросы на зачете 1-30

	ПК-4.2. Умеет вести корректную дискуссию в области информационных технологий, задавать вопросы и отвечать на поставленные вопросы по теме научной работы.	Умеет вести корректную дискуссию в области информационных технологий, задавать вопросы и отвечать на поставленные вопросы по теме научной работы в рамках спецсеминара.	опрос по теме, лабораторная работа	Вопросы на зачете 1-30
	ПК-4.3. Имеет практический опыт участия в научных студенческих конференциях, очных, виртуальных, заочных обсуждениях научных проблем в области информационных технологий	Имеет практический опыт участия в научных студенческих конференциях, очных, виртуальных, заочных обсуждениях научных проблем в области информационных технологий в рамках спецсеминара.	опрос по теме, лабораторная работа	Вопросы на зачете 1-30
0	ПК-5.1. Знает основы разработки и реализации процессов жизненного цикла программного обеспечения.	ПК-5.1. Знает основы разработки и реализации процессов жизненного цикла программного обеспечения в рамках спецсеминара.	опрос по теме, лабораторная работа	Вопросы на зачете 1-30
1	ПК-5.2. Умеет приобретать и использовать организационно-управленческие навыки в профессиональной и социальной деятельности.	ПК-5.2. Умеет приобретать и использовать организационно-управленческие навыки в профессиональной и социальной деятельности в рамках спецсеминара.	опрос по теме, лабораторная работа	Вопросы на зачете 1-30
2	ПК-5.3. Имеет практический опыт управления процессами жизненного цикла программных продуктов.	ПК-5.3. Имеет практический опыт управления процессами жизненного цикла программных продуктов в рамках спецсеминара..	опрос по теме, лабораторная работа	Вопросы на зачете 1-30

Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций в процессе освоения образовательной программы

Темы выступлений к круглому столу

1. *Модели и критерии семантической близости;*
2. *Инструменты Python для работы с естественно-языковой информацией. Библиотека NLTK;*
3. *Нейросеть DeepPfvlov;*
4. *Нейросеть BERT;*
5. *Разноязыковой автоматический перевод.*

Зачетно-экзаменационные материалы для промежуточной аттестации (зачет)

Перечень вопросов, которые выносятся на зачет.

1-й семестр

1. Алфавит, цепочка, полная и усеченная итерации алфавита.
2. Что такое естественный язык, основные функции, характеристики и свойства;
3. Что понимается под структурой письменной и устной речи?
4. Какие основные речевые категории Вы знаете?
5. Морфология слов. Морфологический разбор
6. Грамматическая структура предложений.
7. Синтаксические конструкции. Основная задача разбора предложений.
8. Слово-зависимые представления предложений. Основная роль словосочетания.
9. Отношение непосредственного подчинения. Транзитивное замыкание отношения непосредственного подчинения.
10. Дерево зависимостей и его свойства.
11. Проективные предложения. Основные особенности, свойства и характеристики.
12. Деревья зависимостей проективных предложений

13. Отрезочное представление предложений.
14. Словная шкала и ее свойства.
15. Система составляющих.
16. Разбиения словной шкалы. Оценочные характеристики.
17. Множественность отрезочных представлений.
18. Свойства и построение системы составляющих
19. Связь системы составляющих с проективными предложениями.
20. Как соотносится семантика и выражающий ее естественный язык.
21. Семантическое значение слова. Характеристика и свойства.
22. Как образуется семантическое значение предложения, фрагмента текста, документа.
23. Что такое интерпретация семантики в компьютерной обработке текстов естественных языков?
24. Частотная модель интерпретации семантики.
25. Вероятностная модель интерпретации семантики.
26. Грамматическая модель интерпретации семантики
27. Что такое семантическая близость и ее использование в задачах информационного поиска, каталогизации, классификации текстов и документов.
28. Частотная модель семантической близости.
29. Вероятностная модель семантической близости.
30. Грамматическая модель семантической близости.
31. Семантическая близость слов.
32. Семантическая близость предложений.
33. Семантическая близость документов.
34. Релевантность в информационном поиске.

2-семестр

1. Вычислительное представление смысла, функционал смысловыразительности.
2. Представление функционала смысловыразительности в частотной модели смысла.
3. Представление функционала смысловыразительности в грамматической модели смысла.
4. Операции контекстного уточнения смысла.
5. Формульное представление функционала смысловыразительности.
6. Преобразование функционала смысловыразительности к ОПЗ подобному виду.
7. Вычислительное представление функционала, семантическая схема.
8. Средства языка Python для работы с текстами.
9. Использование библиотеки NLTK в задачах токенизации.
10. Использование библиотеки NLTK в задачах стэмминга.
11. Использование библиотеки NLTK в задачах семантических рассуждений.
12. Использование библиотеки NLTK в задачах семантического наполнения слов.
13. Частотный анализ текстовой информации. Критерии анализа.
14. Частотные характеристики текстовых сообщений.
15. Метод опорных векторов в классификации
16. Задача распознавания спама.
17. Определение уникальности текста на основе морфологического разбора слов.
18. Обработка неструктурированных текстов.
19. Парсеры. Назначение и функции.
20. Что такое морфологический парсер?
21. Что такое синтаксический парсер.
22. Как работает Яндекс–парсер.
23. Обработка неструктурированных текстов шаблонным методом.

Критерии оценивания результатов обучения

Оценка “зачтено” - практические задания выполнены в срок в объеме не менее 80%. Студент демонстрирует правильные, уверенные действия по применению полученных знаний на практике, грамотное и логически стройное изложение материала при аргументации ответов на вопросы при защите заданий.

Оценка «не зачтено» - практические задания не выполнены либо предоставлены не в срок в объеме менее 60%, Студент демонстрирует наличие грубых ошибок в ответе, непонимание сущности излагаемого вопроса, неумение применять знания на практике, неуверенность и неточность ответов на дополнительные и наводящие вопросы.

Оценочные средства для инвалидов и лиц с ограниченными возможностями здоровья выбираются с учетом их индивидуальных психофизических особенностей.

– при необходимости инвалидам и лицам с ограниченными возможностями здоровья предоставляется дополнительное время для подготовки ответа на экзамене;

– при проведении процедуры оценивания результатов обучения инвалидов и лиц с ограниченными возможностями здоровья предусматривается использование технических средств, необходимых им в связи с их индивидуальными особенностями;

– при необходимости для обучающихся с ограниченными возможностями здоровья и инвалидов процедура оценивания результатов обучения по дисциплине может проводиться в несколько этапов.

Процедура оценивания результатов обучения инвалидов и лиц с ограниченными возможностями здоровья по дисциплине (модулю) предусматривает предоставление информации в формах, адаптированных к ограничениям их здоровья и восприятия информации:

Для лиц с нарушениями зрения:

- в печатной форме увеличенным шрифтом,
- в форме электронного документа.

Для лиц с нарушениями слуха:

- в печатной форме,
- в форме электронного документа.

Для лиц с нарушениями опорно-двигательного аппарата:

- в печатной форме,
- в форме электронного документа.

Данный перечень может быть конкретизирован в зависимости от контингента обучающихся.

5. ПЕРЕЧЕНЬ УЧЕБНОЙ ЛИТЕРАТУРЫ, ИНФОРМАЦИОННЫХ РЕСУРСОВ И ТЕХНОЛОГИЙ

5.1. Учебная литература

5.1.1. Основная литература:

1. Боярский К. К. Введение в компьютерную лингвистику. Учебное пособие. – СПб: НИУ ИТМО, 2013. – 72 с.
2. 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.
3. Боярский К. К. Введение в компьютерную лингвистику. Учебное пособие. – СПб: НИУ ИТМО, 2013. – 72 с.
4. 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.
5. Yuri M. Vishnyakov, Renat Y. Vishnyakov Computational theory of semantics representation in scientific and technical texts // AMCSM_2018 IOP Publishing IOP, Conf. Series: Journal of Physics: Conf. Series 1202 (2019) 012008 doi:10.1088/1742-6596/1202/1/012008
6. Y M Vishnyakov and R Y Vishnyakov Measurement of semantic proximity within computational theory of semantic interpretation // AMCSM_2018 IOP Publishing Journal of Physics: Conf. Series 1203 (2019) 012050 doi:10.1088/1742-6596/1203/1/012050
7. Yury M. Vishnyakov*, Renat Yu. Vishnyakov/ The Linguistic Proximity in Information Retrieval and Document Classification. 14th IEEE International Symposium on Computational Intelligence and Informatics to be held on November 19-21, 2013 in Budapest, Hungary. p. 131-134.
8. Yury Vishnyakov, Renat Vishnyakov // Representation of semantically cohesive sentence fragments in scientific and technical texts // 2014 IEEE 12th International Symposium on Applied Machine Intelligence and Informatics, page(s). 295-298, DOI: 10.1109/SAMI.2014. 6822425

9. Вишняков Ю.М., Вишняков Р.Ю. Вычислительная семантическая интерпретация текстов научно-технического стиля // Современные наукоемкие технологии, 2016, № 12-2. – С.236-242.

5.1.2. Дополнительная литература:

1. Алексеев, В.Е. Структуры данных. Модели вычислений / В.Е. Алексеев, В.А. Таланов. - 2-е изд., исправ. - Москва : Национальный Открытый Университет «ИНТУИТ», 2016. - 248 с. : схем., ил. - (Основы информационных технологий). - Библиогр. в кн. - ISBN 5-9556-0066-3 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=428782>

5.1.3. Учебно-методическая литература

1. Методы распознавания символьных последовательностей: учеб. пособие /Ю.М. Вишняков, Р.Ю. Вишняков. – Краснодар: Кубанский гос. ун-т, 2019. – 106 с. – 500 экз. ISBN 978-5-8209-1720-2.

5.2. Периодическая литература

1. Автоматика и вычислительная техника.
2. Реферативный журнал ВИНТИ
3. Базы данных компании «Ист Вью» <http://dlib.eastview.com>
4. Электронная библиотека GREBENNIKON.RU <https://grebennikon.ru/>

5.3. Интернет-ресурсы, в том числе современные профессиональные базы данных и информационные справочные системы

Электронно-библиотечные системы (ЭБС):

1. ЭБС «ЮРАЙТ» <https://urait.ru/>
2. ЭБС «УНИВЕРСИТЕТСКАЯ БИБЛИОТЕКА ОНЛАЙН» www.biblioclub.ru
3. ЭБС «BOOK.ru» <https://www.book.ru>
4. ЭБС «ZNANIUM.COM» www.znanium.com
5. ЭБС «ЛАНЬ» <https://e.lanbook.com>

Профессиональные базы данных:

1. Web of Science (WoS) <http://webofscience.com/>
2. Scopus <http://www.scopus.com/>
3. ScienceDirect www.sciencedirect.com
4. Журналы издательства Wiley <https://onlinelibrary.wiley.com/>
5. Научная электронная библиотека (НЭБ) <http://www.elibrary.ru/>
6. Полнотекстовые архивы ведущих западных научных журналов на Российской платформе научных журналов НЭИКОН <http://archive.neicon.ru>
7. Национальная электронная библиотека (доступ к Электронной библиотеке диссертаций Российской государственной библиотеки (РГБ) <https://rusneb.ru/>
8. Президентская библиотека им. Б.Н. Ельцина <https://www.prilib.ru/>
9. Электронная коллекция Оксфордского Российского Фонда <https://ebookcentral.proquest.com/lib/kubanstate/home.action>
10. Springer Journals <https://link.springer.com/>
11. Nature Journals <https://www.nature.com/siteindex/index.html>
12. Springer Nature Protocols and Methods <https://experiments.springernature.com/sources/springer-protocols>
13. Springer Materials <http://materials.springer.com/>
14. zbMath <https://zbmath.org/>
15. Nano Database <https://nano.nature.com/>
16. Springer eBooks: <https://link.springer.com/>
17. "Лекториум ТВ" <http://www.lektorium.tv/>
18. Университетская информационная система РОССИЯ <http://uisrussia.msu.ru>

Информационные справочные системы:

1. Консультант Плюс - справочная правовая система (доступ по локальной сети с компьютеров библиотеки)

Ресурсы свободного доступа:

1. Американская патентная база данных <http://www.uspto.gov/patft/>
2. Полные тексты канадских диссертаций <http://www.nlc-bnc.ca/thesescanada/>
3. КиберЛенинка (<http://cyberleninka.ru/>);
4. Министерство науки и высшего образования Российской Федерации <https://www.minobrnauki.gov.ru/>;
5. Федеральный портал "Российское образование" <http://www.edu.ru/>;
6. Информационная система "Единое окно доступа к образовательным ресурсам" <http://window.edu.ru/>;
7. Единая коллекция цифровых образовательных ресурсов <http://school-collection.edu.ru/> .
8. Федеральный центр информационно-образовательных ресурсов (<http://fcior.edu.ru/>);
9. Проект Государственного института русского языка имени А.С. Пушкина "Образование на русском" <https://pushkininstitute.ru/>;
10. Справочно-информационный портал "Русский язык" <http://gramota.ru/>;
11. Служба тематических толковых словарей <http://www.glossary.ru/>;
12. Словари и энциклопедии <http://dic.academic.ru/>;
13. Образовательный портал "Учеба" <http://www.ucheba.com/>;
14. Законопроект "Об образовании в Российской Федерации". Вопросы и ответы http://xn--273--84d1f.xn--plai/voprosy_i_otvety
15. Philology.ru [Электронный ресурс]: [филологический портал]. - Режим доступа:– <http://www.philology.ru/>, свободный (дата обращения: 2.02.2017) (библиотека филологических текстов (монографий, статей, методических пособий).
16. Языкознание.ру [Электронный ресурс] : [образовательный портал]. – Режим доступа:– <http://yazykoznanie.ru>, свободный (дата обращения: 2.02.2017) (ресурс для изучающих различные лингвистические дисциплины).
17. Linguists [Электронный ресурс]: [образовательный портал]. – Режим доступа: <http://linguists.narod.ru>, свободный (дата обращения: 12.02.2017) (Ресурсы для переводчиков и лингвистов, содержит список других сетевых ресурсов).
18. Лингвистика для школьников [Электронный ресурс]: [образовательный сайт]. – Режим доступа: –<http://lingling.ru/>, свободный (дата обращения: 2.02.2017).
19. COGNITIV [Электронный ресурс]: [образовательный портал]. – Режим доступа: <http://cognitiv.narod.ru>, свободный (дата обращения: 5.01.2017) (Сайт для ученых-языковедов всех специальностей (обмен новейшей информацией в области лингвистики; обсуждение фундаментальных и прикладных проблем языкознания, а также вопросов взаимоотношения языка, культуры и общества).
20. Лингвистический энциклопедический словарь [Электронный ресурс]: [он-лайн-словарь]. – Режим доступа: <http://lingvisticheskiy-slovar.ru/>, свободный (дата обращения: 17.01.2017).
21. Linguistics Dictionary Glossary Terms Lexicon Online [Электронный ресурс]: [образовательный ресурс]. – Режим доступа: <http://www.glossary.sil.org/>, свободный (дата обращения: 12.02.2017) (глоссарий, содержащий более 950 лингвистических терминов с перекрестными ссылками и списком источников (SIL International).

Собственные электронные образовательные и информационные ресурсы КубГУ:

1. Среда модульного динамического обучения <http://moodle.kubsu.ru>
2. База учебных планов, учебно-методических комплексов, публикаций и конференций <http://mschool.kubsu.ru/>
3. Библиотека информационных ресурсов кафедры информационных образовательных технологий <http://mschool.kubsu.ru/>;
4. Электронный архив документов КубГУ <http://docspace.kubsu.ru/>
5. Электронные образовательные ресурсы кафедры информационных систем и технологий в образовании КубГУ и научно-методического журнала "ШКОЛЬНЫЕ ГОДЫ" <http://icdau.kubsu.ru/>

6. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ (модуля)

Для освоения учебного материала студенту необходимо ознакомиться со структурой курса и методикой овладения материалом. Весь курс построен от простого к сложному и каждая его тема основана на материалах предыдущих тем. В это связи студенту необходимо не терять логику курса и строго ей следовать. В лекционном материале даются, как правило, теоретические сведения, которые раскрываются на практических примерах. Для закрепления теоретических знаний студент получает индивидуальное задание к циклу лабораторных работ, который охватывает весь теоретический материал. Каждая лабораторная работы защищается по мере выполнения. Таким образом, выполняя весь цикл лабораторных работ, студент получает и осваивает знания в соответствии с компетенциями курса. По выступлениям на круглом столе с преподавателем согласовывается тема выступления и готовится само выступление. Во время текущей аттестации могут проводиться контрольные опросы по начитанному теоретическому и практическому материалу.

В освоении дисциплины инвалидами и лицами с ограниченными возможностями здоровья большое значение имеет индивидуальная учебная работа (консультации) – дополнительное разъяснение учебного материала.

Индивидуальные консультации по предмету являются важным фактором, способствующим индивидуализации обучения и установлению воспитательного контакта между преподавателем и обучающимся инвалидом или лицом с ограниченными возможностями здоровья.

7. МАТЕРИАЛЬНО-ТЕХНИЧЕСКАЯ БАЗА, НЕОБХОДИМАЯ ДЛЯ ОСУЩЕСТВЛЕНИЯ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА ПО ДИСЦИПЛИНЕ

Наименование специальных помещений	Оснащенность специальных помещений	Перечень лицензионного программного обеспечения
Учебные аудитории для проведения занятий лекционного типа (ауд. 129, 131, А305).	Мебель: учебная мебель Технические средства обучения: проектор, экран, компьютер/ноутбук) и соответствующим программным обеспечением (ПО)	PowerPoint, доступ к Microsoft Teams
Учебные аудитории для проведения занятий семинарского типа, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации ауд. 129, 131, А305	Мебель: учебная мебель Технические средства обучения: экран, компьютер Оборудование: кондиционер	PowerPoint, доступ к Microsoft Teams
Учебные аудитории для проведения лабораторных работ. Лаборатория (ауд. 102-106, А301-303).	Мебель: учебная мебель Технические средства обучения: экран, проектор, компьютер Оборудование:	системы программирования на языках высокого уровня, сетевой доступ к ресурсам, в частности С++, Object Pascal и пр. с возможностью многопользовательской работы

Для самостоятельной работы обучающихся предусмотрены помещения, укомплектованные специализированной мебелью, оснащенные компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду университета.

Наименование помещений для самостоятельной работы обучающихся	Оснащенность помещений для самостоятельной работы обучающихся	Перечень лицензионного программного обеспечения
Помещение для самостоятельной работы	Мебель: учебная мебель Комплект	Доступ печатным и электронным

Наименование помещений для самостоятельной работы обучающихся	Оснащенность помещений для самостоятельной работы обучающихся	Перечень лицензионного программного обеспечения
обучающихся (читальный зал Научной библиотеки)	специализированной мебели: компьютерные столы Оборудование: компьютерная техника с подключением к информационно-коммуникационной сети «Интернет» и доступом в электронную информационно-образовательную среду образовательной организации, веб-камеры, коммуникационное оборудование, обеспечивающее доступ к сети интернет (проводное соединение и беспроводное соединение по технологии Wi-Fi)	информационным ресурсам
Помещение для самостоятельной работы обучающихся (ауд. 146)	Мебель: учебная мебель Комплект специализированной мебели: компьютерные столы Оборудование: компьютерная техника с подключением к информационно-коммуникационной сети «Интернет» и доступом в электронную информационно-образовательную среду образовательной организации, веб-камеры, коммуникационное оборудование, обеспечивающее доступ к сети интернет (проводное соединение и беспроводное соединение по технологии Wi-Fi)	системы программирования на языках C++ и Object Pascal с возможностью многопользовательской работы