

Документ подписан простой электронной подписью

Информация о владельце:

ФИО: Ястребов Олег Александрович

Должность: Ректор

Дата подписания: 28.10.2024 14:29:36

Уникальный программный ключ:

sa953a01204891083f939673078ef1a989aae18a

**Федеральное государственное автономное образовательное учреждение высшего образования**

**«Российский университет дружбы народов имени Патриса Лумумбы»**

**Факультет физико-математических и естественных наук**

(наименование основного учебного подразделения (ОУП)-разработчика ОП ВО)

## **РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ**

### **МЕТОДЫ ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА ТЕКСТОВ**

(наименование дисциплины/модуля)

**Рекомендована МССН для направления подготовки/специальности:**

#### **09.04.03 ПРИКЛАДНАЯ ИНФОРМАТИКА**

(код и наименование направления подготовки/специальности)

**Освоение дисциплины ведется в рамках реализации основной профессиональной образовательной программы высшего образования (ОП ВО):**

#### **ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ И АНАЛИЗ ДАННЫХ**

(наименование (профиль/специализация) ОП ВО)

**2023 г.**

## 1. ЦЕЛЬ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Дисциплина «Методы интеллектуального анализа текстов» входит в программу магистратуры «Искусственный интеллект и анализ данных» по направлению 09.04.03 «Прикладная информатика» и изучается в 3 семестре 2 курса. Дисциплину реализует Кафедра математического моделирования и искусственного интеллекта. Дисциплина состоит из 3 разделов и 7 тем и направлена на изучение методов интеллектуального анализа текстов, освоение принципов морфологического, синтаксического и семантического анализа, извлечения информации из текстов, классификации текстов.

Целью освоения дисциплины является получение навыков работы с библиотеками обработки текстов, применение инструментов обработки текстов к решению практических задач.

## 2. ТРЕБОВАНИЯ К РЕЗУЛЬТАТАМ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Освоение дисциплины «Методы интеллектуального анализа текстов» направлено на формирование у обучающихся следующих компетенций (части компетенций):

*Таблица 2.1. Перечень компетенций, формируемых у обучающихся при освоении дисциплины (результаты освоения дисциплины)*

Шифр	Компетенция	Индикаторы достижения компетенции (в рамках данной дисциплины)
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; УК-1.2 Умеет соотносить разнородные явления и систематизировать их в рамках избранных видов профессиональной деятельности; УК-1.3 Имеет практический опыт работы с информационными источниками, опыт научного поиска, создания научных текстов;
УК-7	Способен: искать нужные источники информации и данные, воспринимать, анализировать, запоминать и передавать информацию с использованием цифровых средств, а также с помощью алгоритмов при работе с полученными из различных источников данными с целью эффективного использования полученной информации для решения задач; проводить оценку информации, ее достоверность, строить логические умозаключения на основании поступающих информации и данных	УК-7.1 Способен: искать нужные источники информации и данные, воспринимать, анализировать, запоминать и передавать информацию с использованием цифровых средств, а также с помощью алгоритмов при работе с полученными из различных источников данными с целью эффективного использования полученной информации для решения задач; проводить оценку информации, ее достоверность, строить логические умозаключения на основании поступающих информации и данных;
ПК-1	Проведение работ по обработке и анализу научно-технической информации и результатов исследований	ПК-1.1 Знает основы научно-исследовательской деятельности в области информационных и коммуникационных технологий, принципы построения научной работы, методы сбора и анализа полученного материала; владеет навыками подготовки научных обзоров, публикаций, рефератов и библиографий по тематике проводимых исследований на русском и иностранном языках; способен осуществлять подготовку к публикации материалов в научно-технических изданиях; ПК-1.2 Умеет применять полученные знания в области

Шифр	Компетенция	Индикаторы достижения компетенции (в рамках данной дисциплины)
		математики и информатики, а также решать стандартные задачи собственной научно-исследовательской деятельности; умеет решать научные задачи с пониманием существующих подходов к верификации моделей по тематике исследований в соответствии с выбранной методикой; ПК-1.3 Знает основы ведения научной дискуссии и формы устного научного высказывания; умеет вести корректную дискуссию в области информационных технологий, задавать вопросы и отвечать на поставленные вопросы по теме научного исследования; владеет навыками выступлений и научной аргументации при анализе объекта научной и профессиональной деятельности; способен принимать участие в работе научных семинаров, научно-технических конференций;
ПК-2	Организационное и технологическое обеспечение проектирования и дизайна ИС	ПК-2.1 Знает этапы жизненного цикла разработки программных систем, виды архитектур, методологии разработки программных систем, основные понятия и определения, относящиеся к концепции построения информационных систем, методы проектирования информационных систем; ПК-2.2 Умеет применять в практической деятельности профессиональные стандарты в области информационных технологий, проводить анализ требований к информационной системе, разрабатывать варианты реализации информационной системы, производить оценку качества, надежности и эффективности информационных систем;

### 3. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОП ВО

Дисциплина «Методы интеллектуального анализа текстов» относится к блоку по выбору блока образовательной программы высшего образования.

В рамках образовательной программы высшего образования обучающиеся также осваивают другие дисциплины и/или практики, способствующие достижению запланированных результатов освоения дисциплины «Методы интеллектуального анализа текстов».

*Таблица 3.1. Перечень компонентов ОП ВО, способствующих достижению запланированных результатов освоения дисциплины*

Шифр	Наименование компетенции	Предшествующие дисциплины/модули, практики*	Последующие дисциплины/модули, практики*
УК-7	Способен: искать нужные источники информации и данные, воспринимать, анализировать, запоминать и передавать информацию с использованием цифровых средств, а также с помощью алгоритмов при работе с полученными из различных источников данными с целью эффективного использования полученной информации для решения задач; проводить оценку	Интеллектуальные системы и их применение; Методы машинного обучения; Глубокое обучение и обучение с подкреплением;	Преддипломная практика;

Шифр	Наименование компетенции	Предшествующие дисциплины/модули, практики*	Последующие дисциплины/модули, практики*
	информации, ее достоверность, строить логические умозаключения на основании поступающих информации и данных		
УК-1	Способен осуществлять критический анализ проблемных ситуаций на основе системного подхода, выработать стратегию действий	<p>Моделирование беспроводных сетей;</p> <p>Методы машинного обучения;</p> <p>Введение в компьютерные науки и искусственный интеллект;</p> <p>Распознавание образов и обработка изображений;</p> <p>Математическая теория телетрафика;</p> <p>Построение и анализ моделей беспроводных сетей 5G;</p> <p>Прикладные методы компьютерной лингвистики;</p> <p>Глубокое обучение и обучение с подкреплением;</p> <p>Показатели эффективности беспроводных сетей 5G;</p> <p>Модели мультисервисных сетей;</p> <p>Нотации моделирования и анализ бизнес-процессов;</p> <p>Интеллектуальные системы и их применение;</p> <p>Основы компьютерной лингвистики;</p> <p>Объектные и распределенные базы данных;</p>	Технологическая (проектно-технологическая) практика; Преддипломная практика;
ПК-1	Проведение работ по обработке и анализу научно-технической информации и результатов исследований	<p>Ознакомительная практика;</p> <p>Введение в компьютерные науки и искусственный интеллект;</p> <p>Прикладные методы компьютерной лингвистики;</p> <p>Глубокое обучение и обучение с подкреплением;</p> <p>Коллективное поведение интеллектуальных систем;</p> <p>Показатели эффективности беспроводных сетей 5G;</p> <p>Основы компьютерной лингвистики;</p> <p>Иностранный язык в профессиональной деятельности;</p> <p>Модели мультисервисных сетей;</p> <p>Построение и анализ моделей беспроводных сетей 5G;</p> <p>Математическая теория телетрафика;</p>	Преддипломная практика;
ПК-2	Организационное и технологическое обеспечение проектирования и дизайна ИС	<p>Введение в компьютерные науки и искусственный интеллект;</p> <p>Распознавание образов и обработка изображений;</p> <p>Языки программирования для задач искусственного интеллекта;</p>	Технологическая (проектно-технологическая) практика;

Шифр	Наименование компетенции	Предшествующие дисциплины/модули, практики*	Последующие дисциплины/модули, практики*
		Прикладные методы компьютерной лингвистики; Глубокое обучение и обучение с подкреплением; Коллективное поведение интеллектуальных систем; Интеллектуальные системы и их применение; Основы компьютерной лингвистики; Показатели эффективности беспроводных сетей 5G; Модели мультисервисных сетей; Построение и анализ моделей беспроводных сетей 5G; Объектные и распределенные базы данных; Нотации моделирования и анализ бизнес-процессов; Ознакомительная практика;	

\* - заполняется в соответствии с матрицей компетенций и СУП ОП ВО

\*\* - элективные дисциплины /практики

#### 4. ОБЪЕМ ДИСЦИПЛИНЫ И ВИДЫ УЧЕБНОЙ РАБОТЫ

Общая трудоемкость дисциплины «Методы интеллектуального анализа текстов» составляет «5» зачетных единиц.

Таблица 4.1. Виды учебной работы по периодам освоения образовательной программы высшего образования для очной формы обучения.

Вид учебной работы	ВСЕГО, ак.ч.		Семестр(-ы)
			3
<i>Контактная работа, ак.ч.</i>	54		54
Лекции (ЛК)	18		18
Лабораторные работы (ЛР)	0		0
Практические/семинарские занятия (СЗ)	36		36
<i>Самостоятельная работа обучающихся, ак.ч.</i>	99		99
<i>Контроль (экзамен/зачет с оценкой), ак.ч.</i>	27		27
<b>Общая трудоемкость дисциплины</b>	<b>ак.ч.</b>	<b>180</b>	<b>180</b>
	<b>зач.ед.</b>	<b>5</b>	<b>5</b>

## 5. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

Таблица 5.1. Содержание дисциплины (модуля) по видам учебной работы

Номер раздела	Наименование раздела дисциплины	Содержание раздела (темы)		Вид учебной работы*
Раздел 1	Введение в интеллектуальный анализ текстов	1.1	Основные определения. Постановки задач обработки текстов и подходы к их решению	ЛК, СЗ
		1.2	Естественный язык. Типология естественных языков. Уровни анализа. Лингвистическое исследование.	ЛК, СЗ
Раздел 2	Методы лингвистического анализа текстов	2.1	Методы морфологического анализа. Проблема морфологической многозначности. Методы снятия морфологической многозначности.	ЛК, СЗ
		2.2	Синтаксис естественных языков. Способы описания синтаксической структуры предложения. Методы синтаксического анализа.	ЛК, СЗ
		2.3	Способы формализации семантики. Методы семантического анализа. Реляционно-ситуационный анализ текста. Дистрибутивная семантика	ЛК, СЗ
Раздел 3	Прикладные задачи анализа текстов	3.1	Извлечение информации из текстов. Классификация и кластеризация текстов.	ЛК, СЗ
		3.2	Методы машинного обучения для решения задач обработки естественного языка	ЛК, СЗ

\* - заполняется только по **ОЧНОЙ** форме обучения: ЛК – лекции; ЛР – лабораторные работы; СЗ – практические/семинарские занятия.

## 6. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Таблица 6.1. Материально-техническое обеспечение дисциплины

Тип аудитории	Оснащение аудитории	Специализированное учебное/лабораторное оборудование, ПО и материалы для освоения дисциплины (при необходимости)
Лекционная	Аудитория для проведения занятий лекционного типа, оснащенная комплектом специализированной мебели; доской (экраном) и техническими средствами мультимедиа презентаций.	Компьютер/ноутбук с доступом сети Интернет и электронно-образовательной среде Университета, браузер, ПО для просмотра PDF, MS Teams. Дополнительное ПО: офисный пакет MS Office или LibreOffice.
Семинарская	Аудитория для проведения занятий семинарского типа, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, оснащенная комплектом специализированной мебели и техническими средствами мультимедиа презентаций.	Компьютер/ноутбук с доступом сети Интернет и электронно-образовательной среде Университета, браузер, ПО для просмотра PDF, MS Teams, ОС Linux, офисный пакет MS Office.

Тип аудитории	Оснащение аудитории	Специализированное учебное/лабораторное оборудование, ПО и материалы для освоения дисциплины (при необходимости)
Для самостоятельной работы	Аудитория для самостоятельной работы обучающихся (может использоваться для проведения семинарских занятий и консультаций), оснащенная комплектом специализированной мебели и компьютерами с доступом в ЭИОС.	Компьютер/ноутбук с доступом сети Интернет и электронно-образовательной среде Университета, браузер, ПО для просмотра PDF, MS Teams, ОС Linux, офисный пакет MS Office.

\* - аудитория для самостоятельной работы обучающихся указывается **ОБЯЗАТЕЛЬНО!**

## 7. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

*Основная литература:*

1. Осипов Г.С. Лекции по искусственному интеллекту. Изд. 2-е, испр. и доп. М.: Книжный дом «ЛИБРОКОМ», 2009. – 272 с.

2. Смирнов И.В. Введение в анализ естественных языков: учебно-методическое пособие. Москва: Изд-во РУДН, 2014. – 85 с.

*Дополнительная литература:*

1. Боярский К.К. Введение в компьютерную лингвистику. Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики. 2013

2. Jurafsky D., Martin J. H. Speech and language processing. – Pearson, 2014.

*Ресурсы информационно-телекоммуникационной сети «Интернет»:*

1. ЭБС РУДН и сторонние ЭБС, к которым студенты университета имеют доступ на основании заключенных договоров

- Электронно-библиотечная система РУДН – ЭБС РУДН

<http://lib.rudn.ru/MegaPro/Web>

- ЭБС «Университетская библиотека онлайн» <http://www.biblioclub.ru>

- ЭБС Юрайт <http://www.biblio-online.ru>

- ЭБС «Консультант студента» [www.studentlibrary.ru](http://www.studentlibrary.ru)

- ЭБС «Троицкий мост»

2. Базы данных и поисковые системы

- электронный фонд правовой и нормативно-технической документации

<http://docs.cntd.ru/>

- поисковая система Яндекс <https://www.yandex.ru/>

- поисковая система Google <https://www.google.ru/>

- реферативная база данных SCOPUS

<http://www.elsevierscience.ru/products/scopus/>

*Учебно-методические материалы для самостоятельной работы обучающихся при освоении дисциплины/модуля\*:*

1. Курс лекций по дисциплине «Методы интеллектуального анализа текстов».

\* - все учебно-методические материалы для самостоятельной работы обучающихся размещаются в соответствии с действующим порядком на странице дисциплины **в ТУИС!**



## **8. ОЦЕНОЧНЫЕ МАТЕРИАЛЫ И БАЛЛЬНО-РЕЙТИНГОВАЯ СИСТЕМА ОЦЕНИВАНИЯ УРОВНЯ СФОРМИРОВАННОСТИ КОМПЕТЕНЦИЙ ПО ДИСЦИПЛИНЕ**

Оценочные материалы и балльно-рейтинговая система\* оценивания уровня сформированности компетенций (части компетенций) по итогам освоения дисциплины «Методы интеллектуального анализа текстов» представлены в Приложении к настоящей Рабочей программе дисциплины.

\* - ОМ и БРС формируются на основании требований соответствующего локального нормативного акта РУДН.

**РАЗРАБОТЧИК:**

Доцент кафедры  
математического  
моделирования и  
искусственного интеллекта

*Должность, БУП*

*Подпись*

**Васильев С.А.**

*Фамилия И.О.*

**РУКОВОДИТЕЛЬ БУП:**

Заведующий кафедрой  
математического  
моделирования и  
искусственного интеллекта

*Должность БУП*

*Подпись*

**Малых М.Д.**

*Фамилия И.О.*

**РУКОВОДИТЕЛЬ ОП ВО:**

Заведующий кафедрой  
математического  
моделирования и  
искусственного интеллекта

*Должность, БУП*

*Подпись*

**Малых М.Д.**

*Фамилия И.О.*