

Документ подписан простой электронной подписью

Информация о владельце:

ФИО: Ястребов Олег Александрович

Должность: Ректор

Дата подписания: 29.05.2024 11:28:12

Уникальный программный ключ:

ca953a01204891083f939673078ef1a989dae18a

**Федеральное государственное автономное образовательное учреждение высшего образования**

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

**Факультет искусственного интеллекта**

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

## **РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ**

### **МЕТОДЫ К РЕШЕНИЮ ЗАДАЧ INFORMATION EXTRACTION**

(наименование дисциплины/модуля)

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

### **02.04.02 ФУНДАМЕНТАЛЬНАЯ ИНФОРМАТИКА И ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ**

(код и наименование направления подготовки/специальности)

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

### **УПРАВЛЕНИЕ ДАННЫМИ И ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ**

(наименование (профиль/специализация) ОП ВО)

**2024 г.**

## 1. ЦЕЛЬ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Дисциплина «Методы к решению задач Information Extraction» входит в программу магистратуры «Управление данными и искусственный интеллект» по направлению 02.04.02 «Фундаментальная информатика и информационные технологии» и изучается в 3 семестре 2 курса. Дисциплину реализует NOT DEFINED. Дисциплина состоит из 4 разделов и 12 тем и направлена на изучение методов и технологий извлечения информации из текстовых данных.

Целью освоения дисциплины является овладение навыками применения различных методов извлечения информации для решения практических задач обработки текстов.

## 2. ТРЕБОВАНИЯ К РЕЗУЛЬТАТАМ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Освоение дисциплины «Методы к решению задач Information Extraction» 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 поиск, критический анализ проблемных ситуаций на основе системного подхода, выработать стратегию действий	УК-1.3 Имеет практический опыт работы с информационными источниками, опыт научного поиска, создания научных текстов.;
ОПК-1	Способен находить, формулировать и решать актуальные проблемы прикладной математики, фундаментальной информатики и информационных технологий	ОПК-1.2 Умеет осуществлять первичный сбор и анализ материала, интерпретировать различные математические объекты; ОПК-1.3 Имеет практический опыт работы с решением математических задач и применяет его в профессиональной деятельности;
ОПК-3	Способен проводить анализ математических моделей, создавать инновационные методы решения прикладных задач профессиональной деятельности в области информатики и математического моделирования	ОПК-3.2 Умеет соотносить знания в области программирования, интерпретацию прочитанного, определять и создавать информационные ресурсы глобальных сетей, образовательного контента, средств тестирования систем; ОПК-3.3 Имеет практический опыт применения и разработки программного обеспечения, тестирования программных продуктов;
ПК-1	Способен разрабатывать и применять алгоритмы интеллектуальной обработки данных для решения задач профессиональной деятельности	ПК-1.3 Имеет навыки применения существующих и модифицированных алгоритмов интеллектуальной обработки данных для решения задач профессиональной деятельности;

## 3. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОП ВО

Дисциплина «Методы к решению задач Information Extraction» относится к обязательной части блока 1 «Дисциплины (модули)» 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 к решению задач Information Extraction».

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-1	Способен осуществлять поиск, критический анализ проблемных ситуаций на основе системного подхода, выработать стратегию действий	Информационные базы данных; Алгоритмы и структуры данных; Прикладная статистика и анализ данных; Морфологический и семантический анализ; Введение в обработку естественного языка;	Преддипломная практика;
ОПК-1	Способен находить, формулировать и решать актуальные проблемы прикладной математики, фундаментальной информатики и информационных технологий	Обучение с подкреплением; Машинный перевод; Прикладная статистика и анализ данных; Введение в обработку естественного языка; Алгоритмы и структуры данных; Глубинное обучение;	
ОПК-3	Способен проводить анализ математических моделей, создавать инновационные методы решения прикладных задач профессиональной деятельности в области информатики и математического моделирования	Методы глубокого обучения для обработки естественного языка; Морфологический и семантический анализ; Прикладная статистика и анализ данных; Практикум по программированию;	Преддипломная практика;
ПК-1	Способен разрабатывать и применять алгоритмы интеллектуальной обработки данных для решения задач профессиональной деятельности	Алгоритмы и структуры данных; Прикладная статистика и анализ данных; Морфологический и семантический анализ; Методы глубокого обучения для обработки естественного языка; <i>Искусственный интеллект в задачах распознавания образов**;</i> <i>Искусственный интеллект в задачах обработки естественного языка**;</i> <i>Современные устройства центров обработки больших данных**;</i> <i>Современные инструментальные средства разработки ПО для искусственного интеллекта**;</i> Технологическая (проектно-технологическая) практика (учебная);	

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

\*\* - элективные дисциплины /практики

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

Общая трудоемкость дисциплины «Методы к решению задач Information Extraction» составляет «4» зачетные единицы.

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>	36		36
Лекции (ЛК)	0		0
Лабораторные работы (ЛР)	0		0
Практические/семинарские занятия (СЗ)	36		36
<i>Самостоятельная работа обучающихся, ак.ч.</i>	90		90
<i>Контроль (экзамен/зачет с оценкой), ак.ч.</i>	18		18
<b>Общая трудоемкость дисциплины</b>	<b>ак.ч.</b>	<b>144</b>	<b>144</b>
	<b>зач.ед.</b>	<b>4</b>	<b>4</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 Information Extraction	1.1	Определение предметной области Information Extraction и основные понятия	СЗ
		1.2	Типы информации и задачи извлечения информации	СЗ
		1.3	Методы оценки качества извлечения информации	СЗ
Раздел 2	Методы извлечения именованных сущностей	2.1	Токенизация и разметка текста	СЗ
		2.2	Методы извлечения именованных сущностей	СЗ
		2.3	Оценка качества извлечения именованных сущностей	СЗ
Раздел 3	Методы извлечения отношений между сущностями	3.1	Роль зависимостей и синтаксических структур в извлечении отношений	СЗ
		3.2	Методы извлечения отношений между сущностями	СЗ
		3.3	Оценка качества извлечения отношений	СЗ
Раздел 4	Методы извлечения фактов	4.1	Методы извлечения структурированных фактов из текста	СЗ
		4.2	Применение извлеченных фактов в реальных задачах	СЗ
		4.3	Оценка качества извлечения фактов	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)
Компьютерный класс	Компьютерный класс для проведения занятий, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, оснащенная персональными компьютерами (в количестве 15 шт.), доской (экраном) и техническими средствами мультимедиа презентаций.	
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.	
Для	Аудитория для самостоятельной работы	

самостоятельной работы	обучающихся (может использоваться для проведения семинарских занятий и консультаций), оснащенная комплектом специализированной мебели и компьютерами с доступом в ЭИОС.	
------------------------	---	--

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

## 7. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

*Основная литература:*

1. Барсегян А. Технологии анализа данных: Data Mining, Text Mining, Visual Mining, OLAP. 2 изд. – БХВ-Петербург, 2008.

2. Ульман Д., Лесковец Ю., Раджараман А. Анализ больших наборов данных. – Litres, 2022.

*Дополнительная литература:*

1. Рашид Т. Создаем нейронную сеть. – Litres, 2020.

*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 к решению задач Information Extraction».

\* - 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 к решению задач Information Extraction» представлены в Приложении к настоящей Рабочей программе дисциплины.

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

**РАЗРАБОТЧИК:**

_____	_____	Воробьева Александра
<i>Должность, БУП</i>	<i>Подпись</i>	Андреевна
		<i>Фамилия И.О.</i>

**РУКОВОДИТЕЛЬ БУП:**

Заведующий кафедрой	_____	_____
<i>Должность БУП</i>	<i>Подпись</i>	<i>Фамилия И.О.</i>

**РУКОВОДИТЕЛЬ ОП ВО:**

_____	_____	Воробьева Александра
<i>Должность, БУП</i>	<i>Подпись</i>	Андреевна
		<i>Фамилия И.О.</i>