

Документ подписан простой электронной подписью  
Информация о владельце:  
ФИО: Ястребов Олег Александрович  
Должность: Ректор  
Дата подписания: 10.06.2022 10:31:24  
Уникальный программный ключ:  
ca953a0120d891083f939673078ef1a989dae18a

Федеральное государственное автономное образовательное учреждение высшего образования  
**«Российский университет дружбы народов»**  
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:**

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 инфокоммуникациями и интеллектуальные системы

(наименование (профиль/специализация) ОП ВО)

2022 г.

## 1. ЦЕЛЬ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Целью освоения дисциплины является ознакомление студентов с методами интеллектуального анализа текстов, освоение принципов морфологического, синтаксического и семантического анализа, извлечения информации из текстов, классификации текстов.

Основными задачами освоения дисциплины являются 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): УК-1; УК-7; ОПК-2; ОПК-3; ОПК-4; ОПК-5; ПК-1; ПК-2

*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, выработать 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 Знает принципы применения цифровых технологий для сбора, отбора и обобщения информации
		УК-7.2 Умеет применять цифровые технологии для поиска, обработки, анализа, хранения и представления информации в области фундаментальной информатики и информационных технологий
		УК-7.3 Владеет навыками применения цифровых технологий и методов поиска, обработки, анализа, хранения и представления информации в области фундаментальной информатики и информационных технологий
ОПК-2	Способен применять компьютерные/суперкомпьютерные методы, современное программное	ОПК-2.1. Знает основные положения и концепции в области программирования, архитектуру языков программирования, теории коммуникации, знает основную терминологию,

Шифр	Компетенция	Индикаторы достижения компетенции (в рамках данной дисциплины)
	обеспечение (в том числе отечественного производства) для решения задач профессиональной деятельности	знаком с перечнем ПО, включенного в Единый Реестр Российских программ ОПК-2.2. Умеет анализировать типовые языки программирования, составлять программы ОПК-2.3. Имеет практический опыт решения задач анализа, интеграции различных типов программного обеспечения, анализа типов коммуникации
ОПК-3	Способен проводить анализ математических моделей, создавать инновационные методы решения прикладных задач профессиональной деятельности в области информатики и математического моделирования	ОПК-3.1. Знает методы теории алгоритмов, методы системного и прикладного программирования, основные положения и концепции в области математических, информационных и имитационных моделей 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
ОПК-4	Способен оптимальным образом комбинировать существующие информационно-коммуникационные технологии для решения задач в области профессиональной деятельности с учетом требований информационной безопасности	ОПК-4.1. Знает принципы сбора и анализа информации, создания информационных систем на стадиях жизненного цикла ОПК-4.2. Умеет осуществлять управление проектами информационных систем ОПК-4.3. Имеет практический опыт анализа и интерпретации информационных систем
ОПК-5	Способен устанавливать и сопровождать программное обеспечение информационных систем, осуществлять эффективное управление разработкой программных средств и проектов	ОПК-5.1 Знает методику установки и администрирования информационных систем и баз данных. Знаком с перечнем ПО, входящим в Единый реестр российских программ ОПК-5.2 Умеет реализовывать техническое сопровождение информационных систем и баз данных ОПК-5.3 Имеет практические навыки установки и инсталляции программных комплексов
ПК-1	Проведение работ по обработке и анализу научно-технической информации и результатов исследований	ПК-1.3. Умеет применять полученные знания в области фундаментальных научных основ 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

Шифр	Компетенция	Индикаторы достижения компетенции (в рамках данной дисциплины)
		исследований в соответствии с выбранной методикой
ПК-2	Организационное и технологическое обеспечение проектирования и дизайна ИС	<p>ПК-2.1. Знает этапы жизненного цикла разработки программных систем, понятие архитектуры и виды архитектур, методологии разработки программных систем, современные CASE-средства; умеет самостоятельно выбирать подходящее CASE-средство для решения задач на каждом этапе жизненного цикла разработки программных систем; владеет навыками использования CASE-инструментов для разработки программных систем</p> <p>ПК-2.2. Знает основные понятия и определения, относящиеся к концепции построения интерактивных систем; основные методы сбора информации о пользователях и модели поведения пользователя; методы проектирования интерактивных систем; умеет строить модели поведения пользователей; производить постановку задачи по разработке интерактивной системы, разрабатывать спецификации интерактивных систем; производить оценку качества дизайна интерактивных систем; владеет методами оценки эффективности интерактивных систем</p> <p>ПК-2.3 Знает основы проектирования и элементы архитектурных решений информационных систем; умеет применять в практической деятельности профессиональные стандарты в области информационных технологий; имеет практический опыт составления технического задания на разработку информационной системы</p> <p>ПК-2.4 Умеет проводит анализ требований к информационной системе; разрабатывать варианты реализации информационной системы; проводить оценку качества, надежности и эффективности информационной системы</p> <p>ПК-2.5 Знает основы программирования; современные методики тестирования разрабатываемых информационных систем; современные инструменты и методы верификации программного кода.</p> <p>ПК-2.7 Методики описания и моделирования бизнес-процессов; средства моделирования бизнес-процессов; основы реинжиниринга бизнес-процессов организации.</p> <p>ПК-2.8 Знает устройство и функционирование современных информационных систем; современные стандарты взаимодействия</p>

Шифр	Компетенция	Индикаторы достижения компетенции (в рамках данной дисциплины)
		информационных систем; программные средства и платформы инфраструктуры информационных технологий организации; современные подходы и стандарты автоматизации организации (например CRM, ERP, ITIL)

### 3. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОП ВО

Дисциплина «Методы интеллектуального анализа текстов» относится к базовой части блока Б1 ОП ВО.

В рамках ОП ВО 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 <sup>1</sup>
УК-1	Способен осуществлять критический анализ проблемных ситуаций на основе системного подхода, выработать стратегию действий	Анализ и оптимизация проектной деятельности	Технологическая (проектно-технологическая) практика Преддипломная практика
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;	Анализ и оптимизация проектной деятельности Дизайн интерактивных систем	Технологическая (проектно-технологическая) практика Преддипломная практика

<sup>1</sup> - заполняется в соответствии с матрицей компетенций и СУП ОП ВО

Шифр	Наименование компетенции	Предшествующие дисциплины/модули, практики	Последующие дисциплины/модули, практики
	проводить оценку информации, ее достоверность, строить логические умозаключения на основании поступающих информации и данных		
ОПК-2	Способен применять компьютерные/суперкомпьютерные методы, современное программное обеспечение (в том числе отечественного производства) для решения задач профессиональной деятельности	Дизайн интерактивных систем Моделирование вычислительных систем	Технологическая (проектно-технологическая) практика Преддипломная практика
ОПК-3	Способен проводить анализ математических моделей, создавать инновационные методы решения прикладных задач профессиональной деятельности в области информатики и математического моделирования	Локальная организация интеллектуальных систем	Технологическая (проектно-технологическая) практика Преддипломная практика
ОПК-4	Способен оптимальным образом комбинировать существующие информационно-коммуникационные технологии для решения задач в области профессиональной деятельности с учетом требований информационной безопасности	Математические основы распознавания образов	Технологическая (проектно-технологическая) практика Преддипломная практика
ОПК-5	Способен устанавливать и сопровождать программное обеспечение	-	Технологическая (проектно-технологическая) практика Преддипломная практика

Шифр	Наименование компетенции	Предшествующие дисциплины/модули, практики	Последующие дисциплины/модули, практики
	информационных систем, осуществлять эффективное управление разработкой программных средств и проектов		
ПК-1	Проведение работ по обработке и анализу научно-технической информации и результатов исследований	Анализ и оптимизация проектной деятельности Локальная организация интеллектуальных систем Математические основы распознавания образов	Технологическая (проектно-технологическая) практика Преддипломная практика
ПК-2	Организационное и технологическое обеспечение проектирования и дизайна ИС	Дизайн интерактивных систем Моделирование вычислительных систем	Технологическая (проектно-технологическая) практика Преддипломная практика

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

Общая трудоемкость дисциплины «Методы интеллектуального анализа текстов» составляет 4 зачетные единицы.

Таблица 4.1. Виды учебной работы по периодам освоения ОП ВО

Вид учебной работы	ВСЕГО, ак.ч.	Семестр(-ы)
		3
<i>Контактная работа, ак.ч.</i>	54	54
Лекции (ЛК)	18	18
Лабораторные работы (ЛР)	-	-
Практические/семинарские занятия (СЗ)	36	36
<i>Самостоятельная работа обучающихся, ак.ч.</i>	63	63
<i>Контроль (экзамен/зачет с оценкой), ак.ч.</i>	27	27
<b>Общая трудоемкость дисциплины</b>	ак.ч.	<b>144</b>
	зач.ед.	<b>4</b>

## 5. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

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

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

## 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
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	-

2 - заполняется только по ОЧНОЙ форме обучения: ЛК – лекции; ЛР – лабораторные работы; СЗ – семинарские занятия.



Тип аудитории	Оснащение аудитории	Специализированное учебное/лабораторное оборудование, ПО и материалы для освоения дисциплины
	и промежуточной аттестации, оснащенная комплектом специализированной мебели и техническими средствами мультимедиа презентаций.	
Для самостоятельной работы обучающихся	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? Dev-C++ (лицензия GNU GPL), свободно распространяемые анализаторы естественных языков. Дополнительное ПО: офисный пакет MS Office или LibreOffice.

## 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:
  - 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 «Лань» <http://e.lanbook.com/>
  - 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/>




Учебно-методические материалы для самостоятельной работы обучающихся при освоении дисциплины<sup>3</sup>:

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

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

Оценочные материалы и балльно-рейтинговая система<sup>4</sup> оценивания уровня сформированности компетенций по итогам освоения дисциплины «Методы интеллектуального анализа текстов» представлены в Приложении (Фонд оценочных средств) к настоящей Рабочей программе дисциплины.

### РАЗРАБОТЧИКИ:

Доцент кафедры информационных технологий		И.В. Смирнов
Должность, БУП	Подпись	Фамилия И.О.
<b>РУКОВОДИТЕЛЬ БУП:</b>		
Зав. кафедрой информационных технологий		Ю.Н. Орлов
Наименование БУП	Подпись	Фамилия И.О.
<b>РУКОВОДИТЕЛЬ ОП ВО:</b>		
Зав. кафедрой прикладной информатики и теории вероятностей		К.Е. Самуйлов
Должность, БУП	Подпись	Фамилия И.О.

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

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