

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

**Институт русского языка**

---

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

## **РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ**

**Введение в анализ больших данных**

---

(наименование дисциплины/модуля)

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

**45.03.01 Филология**

---

(код и наименование направления подготовки/специальности)

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

**Прикладная цифровая филология**

---

(наименование (профиль/специализация) ОП ВО)

**2022 г.**

## 1. ЦЕЛЬ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Целью освоения дисциплины «Введение в анализ больших данных» является изучение происхождения больших данных и их возможностей использования. В рамках дисциплины также будет предоставлен обзор некоторых ключевых характеристик больших данных и краткое изложение процесса науки о данных, нацеленного на получения пользы от больших данных.

## 2. ТРЕБОВАНИЯ К РЕЗУЛЬТАТАМ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Освоение дисциплины «Введение в анализ больших данных» 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)
ПК-8.2	Применяет основные законы естественнонаучных дисциплин в профессиональной в том числе педагогической, деятельности, использует методы математическо-статистического анализа, логики и моделирования, теоретического и экспериментального исследования в гуманитарной сфере, в том числе во взаимосвязи с основными филологическими законами и методами	Знать: <ul style="list-style-type: none"><li>– правила, нормы поведения в процессе использования цифровых технологии и коммуникации в цифровых средах.</li><li>– правила информационной безопасности</li><li>– специализированные цифровые инструменты и технологии для профессиональной деятельности</li></ul> Уметь: <ul style="list-style-type: none"><li>– применять цифровые инструменты и технологии в профессиональной деятельности, включая совместную работу и для совместного производства ресурсов и знаний</li><li>– использовать цифровые инструменты для создания знаний и инноваций</li><li>– управлять информацией с применением специализированных цифровых</li></ul>

Шифр	Компетенция	Индикаторы достижения компетенции (в рамках данной дисциплины)
ПК-9.1	Решает стандартные задачи в профессиональной, в том числе педагогической, деятельности с использованием специализированных цифровых инструментов и технологий с учетом требований информационной безопасности	<p>инструментов и технологий</p> <p>Знать:</p> <ul style="list-style-type: none"> <li>– основы современных систем управления базами данных, теорию баз данных, основы программирования</li> </ul> <p>Уметь:</p> <ul style="list-style-type: none"> <li>– использовать цифровые технологии и элементарные языки программирования для разработки алгоритмов и программ в области интеллектуального анализа данных</li> </ul> <p>Владеть:</p> <ul style="list-style-type: none"> <li>– навыками программирования на высокоуровневых языках программирования общего назначения</li> </ul>
ПК-9.2	Использует цифровые технологии и высокоуровневые языки программирования общего назначения для разработки алгоритмов и программ в области интеллектуального анализа данных	
ПК-10.1	Знает современные цифровые технологии и средства управления информацией для применения в профессиональной, в том числе педагогической, деятельности	

<b>Шифр</b>	<b>Компетенция</b>	<b>Индикаторы достижения компетенции (в рамках данной дисциплины)</b>
ПК-10.2	Применяет, комбинирует и адаптирует существующие информационно-коммуникационные технологии для решения задач в профессиональной, в том числе педагогической, деятельности с учетом требований информационной безопасности	

### **3. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОП ВО**

Дисциплина «Введение в анализ больших данных» относится к части, формируемой участниками образовательных отношений Б1.В.ДВ.10.02

В рамках ОП ВО обучающиеся также осваивают другие дисциплины и/или практики, способствующие достижению запланированных результатов освоения дисциплины «Введение в анализ больших данных».

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

<b>Шифр</b>	<b>Наименование компетенции</b>	<b>Предшествующие дисциплины/ модули, практики*</b>	<b>Последующие дисциплины/ модули, практики*</b>
УК-1.3	Осуществляет поиск информации для решения поставленной задачи по различным типам запросов	Информационные системы анализа данных в гуманитарной сфере	Преддипломная практика
УК-6.3	Находит и использует источники получения дополнительной информации для повышения уровня общих и профессиональных знаний с учетом условий, средств, личностных возможностей, этапов карьерного роста, временной перспективы развития деятельности и требований рынка труда		
УК-12.1	Ищет нужные источники информации и данные, воспринимает, анализирует, запоминает и передает 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		

<b>Шифр</b>	<b>Наименование компетенции</b>	<b>Предшествующие дисциплины/ модули, практики*</b>	<b>Последующие дисциплины/ модули, практики*</b>
УК-12.2	Проводит оценку информации, ее достоверность, строит логические умозаключения на основании поступающих информации и данных		
ОПК-2.2	Обладает навыками филологического анализа и интерпретации языкового материала, различных типов текстов		
ОПК-3.3	Корректно осуществляет библиографический поиск в традиционных и электронных источниках и составляет корректное описание на основе полученных результатов		
ОПК-4.1	Владет методикой сбора и филологического анализа языковых и литературных фактов и текстов разной степени сложности		
ОПК-7.1	Знает основные цифровые технологии, принципы их работы и методы для изучения и моделирования объектов профессиональной, в том числе педагогической, деятельности, анализа данных, представления информации и пр.		
ОПК-7.2	Применяет современные цифровые методы и технологии анализа, обработки и представления информации в профессиональной, в том числе педагогической, сфере деятельности		
ПК-9.1	Решает стандартные задачи в профессиональной, в том числе педагогической, деятельности с использованием специализированных цифровых инструментов и технологий с учетом требований информационной безопасности		
ПК-9.2	Использует цифровые технологии и высокоуровневые языки программирования общего назначения для разработки алгоритмов и программ в области интеллектуального анализа данных		
ПК-8.2	Применяет основные законы естественнонаучных дисциплин в профессиональной в том числе педагогической, деятельности, использует методы математическо-статистического анализа, логики и моделирования, теоретического и экспериментального исследования в гуманитарной сфере, в том числе во взаимосвязи с основными филологическими законами и методами		
ПК-10.2	Применяет, комбинирует и адаптирует существующие информационно-коммуникационные технологии для решения задач в профессиональной, в том числе педагогической, деятельности с учетом требований информационной безопасности		
ОПК-2.2	Обладает навыками филологического анализа и интерпретации языкового материала, различных типов текстов	Инструменты искусственного интеллекта для анализа и обработки текста	
ОПК-3.2.	Соотносит знания в области теории литературы с конкретным литературным материалом и дает историко-литературную интерпретацию прочитанного и определяет жанровую специфику фольклорного и литературного явления		

<b>Шифр</b>	<b>Наименование компетенции</b>	<b>Предшествующие дисциплины/ модули, практики*</b>	<b>Последующие дисциплины/ модули, практики*</b>
ОПК-3.3	Корректно осуществляет библиографический поиск в традиционных и электронных источниках и составляет корректное описание на основе полученных результатов		
ОПК-4.2.	Интерпретирует тексты разных типов и жанров на основе существующих методик		
ОПК-7.1.	Знает основные цифровые технологии, принципы их работы и методы для изучения и моделирования объектов профессиональной, в том числе педагогической, деятельности, анализа данных, представления информации и пр.		
ОПК-7.2.	Применяет современные цифровые методы и технологии анализа, обработки и представления информации в профессиональной, в том числе педагогической, сфере деятельности		
ПК-8.2.	Применяет основные законы естественнонаучных дисциплин в профессиональной в том числе педагогической, деятельности, использует методы математическо-статистического анализа, логики и моделирования, теоретического и экспериментального исследования в гуманитарной сфере, в том числе во взаимосвязи с основными филологическими законами и методами		
ПК-9.2.	Использует цифровые технологии и высокоуровневые языки программирования общего назначения для разработки алгоритмов и программ в области интеллектуального анализа данных		
ПК-10.2	Применяет, комбинирует и адаптирует существующие информационно-коммуникационные технологии для решения задач в профессиональной, в том числе педагогической, деятельности с учетом требований информационной безопасности		
ОПК-2.2	Обладает навыками филологического анализа и интерпретации языкового материала, различных типов текстов		
ОПК-3.2.	Соотносит знания в области теории литературы с конкретным литературным материалом и дает историко-литературную интерпретацию прочитанного и определяет жанровую специфику фольклорного и литературного явления		
ОПК-3.3	Корректно осуществляет библиографический поиск в традиционных и электронных источниках и составляет корректное описание на основе полученных результатов		
ОПК-4.2.	Интерпретирует тексты разных типов и жанров на основе существующих методик		
ОПК-7.1.	Знает основные цифровые технологии, принципы их работы и методы для изучения и моделирования объектов профессиональной, в том числе педагогической, деятельности, анализа данных, представления информации и пр.		
ОПК-7.2.	Применяет современные цифровые методы и технологии анализа, обработки и представления		

<b>Шифр</b>	<b>Наименование компетенции</b>	<b>Предшествующие дисциплины/ модули, практики*</b>	<b>Последующие дисциплины/ модули, практики*</b>
	информации в профессиональной, в том числе педагогической, сфере деятельности		
ПК-8.2.	Применяет основные законы естественнонаучных дисциплин в профессиональной в том числе педагогической, деятельности, использует методы математическо-статистического анализа, логики и моделирования, теоретического и экспериментального исследования в гуманитарной сфере, в том числе во взаимосвязи с основными филологическими законами и методами		
ПК-9.2.	Использует цифровые технологии и высокоуровневые языки программирования общего назначения для разработки алгоритмов и программ в области интеллектуального анализа данных		
ПК-10.2	Применяет, комбинирует и адаптирует существующие информационно-коммуникационные технологии для решения задач в профессиональной, в том числе педагогической, деятельности с учетом требований информационной безопасности		
УК-6.3	Находит и использует источники получения дополнительной информации для повышения уровня общих и профессиональных знаний с учетом условий, средств, личностных возможностей, этапов карьерного роста, временной перспективы развития деятельности и требований рынка труда		
УК-12.1	Ищет нужные источники информации и данные, воспринимает, анализирует, запоминает и передает 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		
ОПК-7.1	Знает основные цифровые технологии, принципы их работы и методы для изучения и моделирования объектов профессиональной, в том числе педагогической, деятельности, анализа данных, представления информации и пр.	Программирование на Python	
ОПК-7.2	Применяет современные цифровые методы и технологии анализа, обработки и представления информации в профессиональной, в том числе педагогической, сфере деятельности		
ПК-8.2	Применяет основные законы естественнонаучных дисциплин в профессиональной в том числе педагогической, деятельности, использует методы математическо-статистического анализа, логики и моделирования, теоретического и экспериментального исследования в гуманитарной сфере, в том числе во взаимосвязи с основными филологическими законами и методами		
ПК9.2	Использует цифровые технологии и высокоуровневые языки программирования общего		

Шифр	Наименование компетенции	Предшествующие дисциплины/ модули, практики*	Последующие дисциплины/ модули, практики*
	назначения для разработки алгоритмов и программ в области интеллектуального анализа данных		
ПК-10.1	Знает современные цифровые технологии и средства управления информацией для применения в профессиональной, в том числе педагогической, деятельности		
ПК-10.2	Применяет, комбинирует и адаптирует существующие информационно-коммуникационные технологии для решения задач в профессиональной, в том числе педагогической, деятельности с учетом требований информационной безопасности		

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

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

Общая трудоемкость дисциплины «Введение в анализ больших данных» составляет 5 зачетных единиц.

Таблица 4.1. Виды учебной работы по периодам освоения ОП ВО для ОЧНОЙ формы обучения

Вид учебной работы	ВСЕГО, ак.ч.	Семестр(-ы)			
		7	8		
<i>Контактная работа, ак.ч.</i>	78	30	48		
Лекции (ЛК)	39	15	24		
Лабораторные работы (ЛР)					
Практические/семинарские занятия (СЗ)	39	15	24		
<i>Самостоятельная работа обучающихся, ак.ч.</i>	66	33	33		
<i>Контроль (зачет), ак.ч.</i>	36	9	27		
<b>Общая трудоемкость дисциплины</b>	ак.ч.	<b>180</b>	<b>72</b>	<b>108</b>	
	зач.ед.	<b>5</b>	<b>2</b>	<b>3</b>	

Таблица 4.2. Виды учебной работы по периодам освоения ОП ВО для ЗАОЧНОЙ формы обучения\*

Вид учебной работы	ВСЕГО, ак.ч.	Семестр(-ы)			
		10			
<i>Контактная работа, ак.ч.</i>	36	36			
Лекции (ЛК)	18	18			
Лабораторные работы (ЛР)					
Практические/семинарские занятия (СЗ)	18	18			
<i>Самостоятельная работа обучающихся, ак.ч.</i>	135	135			
<i>Контроль (зачет с оценкой), ак.ч.</i>	9	9			
<b>Общая трудоемкость дисциплины</b>	ак.ч.	<b>180</b>	<b>180</b>		
	зач.ед.	<b>5</b>	<b>5</b>		

\* - заполняется в случае реализации программы в заочной форме

## 5. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

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

Наименование раздела дисциплины	Содержание раздела (темы)	Вид учебной работы*
Раздел 1. Понимание ландшафта больших данных	Тема 1.1. Что такое большие данные?	ЛК
	Тема 1.2. Что делает большие данные ценными?	
	Тема 1.3. Откуда берутся большие данные?	
	Тема 1.4. Как меняются данные?	
	Тема 1.5. Сдвиг в обработке из-за больших данных	
Раздел 2. Примеры использования больших данных	Тема 2.1. Правительство	ЛК
	Тема 2.2. Коммерческая деятельность	
	Тема 2.3. Защита	
	Тема 2.4. Здравоохранение и науки о жизни	
	Тема 2.5. Глубокое обучение и социальные сети	
	Тема 2.6. Исследовательская экосистема	
	Тема 2.7. Астрономия и физика	
	Тема 2.8. Окружающая среда, Земля и полярные науки	
	Тема 2.9. Энергия	
Раздел 3. Характеристики больших данных	Тема 3.1. Объем	ЛК
	Тема 3.2. Разнообразие	
	Тема 3.3. Скорость	СЗ
	Тема 3.4. Достоверность	
	Тема 3.5. Ценность	
	Тема 3.6. Другие	
Раздел 4. Начало работы с аналитикой больших данных	Тема 4.1. Как происходит наука о больших данных?	ЛК
	Тема 4.2. Пять компонентов науки о данных	
	Тема 4.3. Шаги в процессе обработки данных	СЗ
	Тема 4.4. Роль аналитика данных	
	Тема 4.5. Смешивание данных из нескольких источников	
Раздел 5. Анализ больших данных в контексте	Тема 5.1. Сосредоточиться на контексте, а не только на интеграции	ЛК
	Тема 5.2. Объединение больших данных с пространственными данными	

Раздел 6. Получение выгоды от прогнозной аналитики и больших данных	Тема 6.1. Зачем нужна предиктивная аналитика больших данных?	ЛК
	Тема 6.2. Вывод прогнозной аналитики на передний план	
	Тема 6.3. Получение реальной ценности для бизнеса от прогнозного анализа	
	Тема 6.4. Построение стратегии больших данных	
Раздел 7. Распределенная файловая система	Тема 7.1. Начало работы с Hadoop	ЛК, СЗ
	Тема 7.2. Экосистема Hadoop	
	Тема 7.3. Распределенная файловая система Hadoop: система хранения больших данных	
	Тема 7.4. YARN: менеджер ресурсов для Hadoop	
Раздел 8. Гуманизация аналитики больших данных	Тема 8.1. Предоставление больших данных тем, кто в них нуждается	ЛК
	Тема 8.2. Принципы гуманизации данных	
	Тема 8.3. Гуманизация рабочего процесса аналитики больших данных	
	Тема 8.4. Рассмотрение потребления аналитики больших данных	

\* - заполняется только по **ОЧНОЙ** форме обучения: ЛК – лекции; ЛР – лабораторные работы; СЗ – 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.	Проектор и/или большой экран/монитор, компьютер, ОС MS Windows 10, MS Office 2016 или MS Office 2019 или MS Office 365, портал Microsoft Azure, доступ к интернету
Лаборатория	Аудитория для проведения лабораторных работ, индивидуальных консультаций, текущего контроля и промежуточной аттестации, оснащенная комплектом специализированной мебели и оборудованием.	Проектор и/или большой экран/монитор, компьютер, ОС MS Windows 10, MS Office 2016 или MS Office 2019 или MS Office 365,

Тип аудитории	Оснащение аудитории	Специализированное учебное/лабораторное оборудование, ПО и материалы для освоения дисциплины (при необходимости)
		портал Microsoft Azure, доступ к интернету
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.	Проектор и/или большой экран/монитор, компьютер, ОС MS Windows 10, MS Office 2016 или MS Office 2019 или MS Office 365, портал Microsoft Azure, доступ к интернету
Компьютерный класс	Компьютерный класс для проведения занятий, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, оснащенная персональными компьютерами (в количестве 12 шт.), доской (экраном) и техническими средствами мультимедиа презентаций.	Проектор и/или большой экран/монитор, компьютер, ОС MS Windows 10, MS Office 2016 или MS Office 2019 или MS Office 365, портал Microsoft Azure, доступ к интернету
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.	Проектор и/или большой экран/монитор, компьютер, ОС MS Windows 10, MS Office 2016 или MS Office 2019 или MS Office 365, портал Microsoft Azure, доступ к интернету

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

## 7. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

*Основная литература:*

1. Holmes DA. Big Data. A very short introduction. Oxford: Oxford University Press; 2017.

2. Mohanty H, Bhuyan P, Chenthati D (eds.). Big Data: A Primer. New Delhi: Springer India, New Delhi; 2015.

*Дополнительная литература:*

1. Mayer-Schonberger V. & Cukier C. (2014) Big Data: A Revolution that will transform how we live, work and think, Eamon Dolan/Mariner Books.

*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 «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/>
- 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 «Введение в анализ больших данных».
2. Презентационные материалы по дисциплине «Введение в анализ больших данных»

\* - 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 «Введение в анализ больших данных» представлены в Приложении к настоящей Рабочей программе дисциплины.

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

**РАЗРАБОТЧИКИ:**

**Центр развития цифровых  
технологий для  
образовательных процессов  
РУДН**

Должность, БУП



Подпись

**Азофейфа Г.Э.Х.**

Фамилия И.О.

**РУКОВОДИТЕЛЬ БУП:**

**Директор  
Института русского языка**

Наименование БУП



Подпись

**Должикова А.В.**

Фамилия И.О.

**РУКОВОДИТЕЛЬ ОП ВО:**

**Заведующий кафедрой  
русского языка № 5  
Института русского языка**

Должность, БУП



Подпись

**Куриленко В.Б.**

Фамилия И.О.