

**МИНИСТЕРСТВО ТРАНСПОРТА РОССИЙСКОЙ ФЕДЕРАЦИИ**  
**ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ**  
**УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ**  
**«РОССИЙСКИЙ УНИВЕРСИТЕТ ТРАНСПОРТА»**  
**(РУТ (МИИТ))**



Рабочая программа дисциплины (модуля),  
как компонент образовательной программы  
высшего образования - программы бакалавриата  
по направлению подготовки  
09.03.03 Прикладная информатика,  
утвержденной первым проректором РУТ (МИИТ)  
Тимониным В.С.

**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)**

**Методы и системы обработки больших данных**

Направление подготовки: 09.03.03 Прикладная информатика

Направленность (профиль): Прикладная информатика в экономике и бизнесе

Форма обучения: Очная

Рабочая программа дисциплины (модуля) в виде электронного документа выгружена из единой корпоративной информационной системы управления университетом и соответствует оригиналу

Простая электронная подпись, выданная РУТ (МИИТ)  
ID подписи: 564169  
Подписал: заведующий кафедрой Каргина Лариса Андреевна  
Дата: 27.02.2023

## 1. Общие сведения о дисциплине (модуле).

Целями освоения дисциплины являются:

-развитие общематематической культуры, логического и алгоритмического мышления студентов;

-выработка умения моделировать реальные финансово-экономические процессы;

-освоение приемов исследования и решения математически формализованных задач;

Задачи освоения дисциплины:

-получение необходимого математического аппарата для изучения дисциплин профессионального цикла и применения этого аппарата в будущей профессиональной деятельности;

-освоение основных методов обработки больших данных для целей обоснования принятия управленческих решений.

## 2. Планируемые результаты обучения по дисциплине (модулю).

Перечень формируемых результатов освоения образовательной программы (компетенций) в результате обучения по дисциплине (модулю):

**ПК-1** - Способен разрабатывать и адаптировать прикладное программное обеспечение;

**ПК-2** - Способен проектировать ИС по видам обеспечения.

Обучение по дисциплине (модулю) предполагает, что по его результатам обучающийся будет:

**Уметь:**

-разрабатывать концептуальную модель прикладной области;

-выбирать методы анализа данных, а также адаптировать базовые методы к решению прикладных задач;

-формулировать содержательные выводы и рекомендации.

**Знать:**

-принципы обработки больших данных, область их практического применения;

-принципы работы реляционных и нереляционных баз данных.

**Владеть:**

-навыками разработки концептуальной модели;

- навыками обработки больших данных.

### 3. Объем дисциплины (модуля).

#### 3.1. Общая трудоемкость дисциплины (модуля).

Общая трудоемкость дисциплины (модуля) составляет 4 з.е. (144 академических часа(ов)).

3.2. Объем дисциплины (модуля) в форме контактной работы обучающихся с педагогическими работниками и (или) лицами, привлекаемыми к реализации образовательной программы на иных условиях, при проведении учебных занятий:

Тип учебных занятий	Количество часов	
	Всего	Семестр №7
Контактная работа при проведении учебных занятий (всего):	64	64
В том числе:		
Занятия лекционного типа	32	32
Занятия семинарского типа	32	32

3.3. Объем дисциплины (модуля) в форме самостоятельной работы обучающихся, а также в форме контактной работы обучающихся с педагогическими работниками и (или) лицами, привлекаемыми к реализации образовательной программы на иных условиях, при проведении промежуточной аттестации составляет 80 академических часа (ов).

3.4. При обучении по индивидуальному учебному плану, в том числе при ускоренном обучении, объем дисциплины (модуля) может быть реализован полностью в форме самостоятельной работы обучающихся, а также в форме контактной работы обучающихся с педагогическими работниками и (или) лицами, привлекаемыми к реализации образовательной программы на иных условиях, при проведении промежуточной аттестации.

### 4. Содержание дисциплины (модуля).

#### 4.1. Занятия лекционного типа.

№ п/п	Тематика лекционных занятий / краткое содержание
1	Введение в «большие данные» Рассматриваемые вопросы: - Понятие «большие данные». - Этапы работы с Big Data.
2	Введение в «большие данные»

№ п/п	Тематика лекционных занятий / краткое содержание
	<p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> <li>-Классические подходы. Подходы, ориентированные на данные.</li> <li>-Принципы работы с большими данными.</li> </ul>
3	<p>Прикладные инструменты для работы с данными.</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> <li>-Лямбда-архитектура.</li> <li>-Каппа-архитектура.</li> </ul>
4	<p>Прикладные инструменты для работы с данными.</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> <li>-Инструментарий Map/Reduce.</li> <li>-Преимущества и недостатки Map/Reduce.</li> </ul>
5	<p>Прикладные инструменты для работы с данными.</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> <li>-Apache Spark.</li> <li>-Cassandra.</li> </ul>
6	<p>Прикладные инструменты для работы с данными.</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> <li>-Машинное обучение.</li> <li>-Принцип работы машинного обучения.</li> </ul>
7	<p>Работа с большими данными</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> <li>-Уровни в системах обработки данных.</li> <li>-Сбор данных.</li> </ul>
8	<p>Работа с большими данными</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> <li>-Преобразование данных.</li> <li>- Контроль данных.</li> </ul>
9	<p>Работа с большими данными</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> <li>-Анализ данных.</li> <li>-Представление результатов.</li> </ul>
10	<p>Работа с большими данными</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> <li>-Распределённые файловые системы.</li> <li>- Архитектура распределённых файловых систем</li> </ul>
11	<p>Работа с большими данными</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> <li>-Файловая система Hadoop.</li> <li>- Архитектура HDFS</li> </ul>
12	<p>Работа с большими данными</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> <li>-Реляционные хранилища данных.</li> <li>-Особенности реляционных баз данных.</li> </ul>
13	<p>Работа с большими данными</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> <li>-NoSQL- хранилища данных.</li> <li>- Четыре вида нереляционных баз данных</li> </ul>
14	<p>Работа с большими данными</p> <p>Рассматриваемые вопросы:</p>

№ п/п	Тематика лекционных занятий / краткое содержание
	-Виды NoSQL: HBase, MongoDB. - Сравнение реляционных и нереляционных баз данных.

#### 4.2. Занятия семинарского типа.

##### Практические занятия

№ п/п	Тематика практических занятий/краткое содержание
1	Большие данные На практическом занятии студент изучит понятие «большие данные».
2	Большие данные На практическом занятии студент изучит примеры больших данных.
3	Большие данные На практическом занятии студент изучит принципы работы с большими данными.
4	Большие данные На практическом занятии студент изучит получение первичных статистик набора данных.
5	Большие данные На практическом занятии студент изучит возможности MS Excel.
6	Инструменты для работы с данными. В результате практического занятия изучается инструментарий HDMR.
7	Инструменты для работы с данными. В результате практического занятия изучаются примеры применения операторов Map и Reduce на языках Python и R при работе с большими данными.
8	Анализ В результате практического занятия осваиваются элементы кластерного анализа с учителем и без учителя.
9	Анализ В результате практического занятия осваиваются SQL запросы.
10	Анализ В результате практического занятия осваиваются NoSQL запросы.
11	Анализ В результате практического занятия осваиваются визуализация.
12	Анализ В результате практического занятия осваиваются примеры визуализации.
13	Анализ В результате практического занятия осваиваются инструменты визуализации.
14	Анализ В результате практического занятия осваиваются виды визуализации

#### 4.3. Самостоятельная работа обучающихся.

№ п/п	Вид самостоятельной работы
1	Подготовка к практическим занятиям.
2	Самостоятельное изучение темы «База данных MongoDB».
3	Подготовка к промежуточной аттестации.

4	Подготовка к текущему контролю.
---	---------------------------------

5. Перечень изданий, которые рекомендуется использовать при освоении дисциплины (модуля).

№ п/п	Библиографическое описание	Место доступа
1	Постреляционные хранилища данных : учебное пособие для вузов / Ю. П. Парфенов ; под научной редакцией Н. В. Папуловской. — Москва : Издательство Юрайт, 2022. — 121 с. — ISBN 978-5-534-09837-2.	<a href="https://urait.ru/bcode/492609">https://urait.ru/bcode/492609</a> (дата обращения: 13.04.2023).— Текст : электронный
2	Большие данные. Big Data: учебник для вузов - Макшанов А. В., Журавлев А. Е., Тындыкарь Л. Н. - Издательство "Лань" 2- е издание, - 2022г. - 188 с. - ISBN - 978-5-8114-9690-7.	<a href="https://e.lanbook.com/book/198599">https://e.lanbook.com/book/198599</a> (дата обращения: 13.04.2023) - Текст: электронный

6. Перечень современных профессиональных баз данных и информационных справочных систем, которые могут использоваться при освоении дисциплины (модуля).

Официальный сайт РУТ (МИИТ): (<https://www.miit.ru/>);

Научно-техническая библиотека РУТ (МИИТ): (<http://library.miit.ru/>);

Образовательная платформа «Юрайт»: (<https://urait.ru/>);

Электронно-библиотечная система издательства «Лань»: (<http://e.lanbook.com/>);

Федеральная служба государственной статистики: (<https://rosstat.gov.ru/>);

Библиотека естественных наук РАН: (<http://www.benran.ru/>).

7. Перечень лицензионного и свободно распространяемого программного обеспечения, в том числе отечественного производства, необходимого для освоения дисциплины (модуля).

1. Офисный пакет приложений Microsoft Office;

2. PyCharm;

3. СУБД Microsoft SQL Server.

8. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине (модулю).

Для проведения лекционных занятий необходима аудитория с мультимедиа аппаратурой.

9. Форма промежуточной аттестации:

Экзамен в 7 семестре.

10. Оценочные материалы.

Оценочные материалы, применяемые при проведении промежуточной аттестации, разрабатываются в соответствии с локальным нормативным актом РУТ (МИИТ).

Авторы:

доцент, доцент, к.н. кафедры  
«Информационные системы  
цифровой экономики»

Е.А. Сеславина

Согласовано:

Заведующий кафедрой ИСЦЭ  
Председатель учебно-методической  
комиссии

Л.А. Каргина

М.В. Ишханян