

Документ подписан простой электронной подписью  
Информация о владельце:

ФИО: Исаев Игорь Магомедович

Должность: Проректор по учебной работе

Дата подписания: 28.08.2023 14:45:54

Уникальный идентификатор документа:

d7a26b9e8ca85e98ec3de2eb454b4659d061f249

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

**Федеральное государственное автономное образовательное учреждение  
высшего образования**

**«Национальный исследовательский технологический университет «МИСИС»**

## Рабочая программа дисциплины (модуля)

# Основы Data engineering

Закреплена за подразделением

Кафедра инфокоммуникационных технологий

Направление подготовки

09.03.02 ИНФОРМАЦИОННЫЕ СИСТЕМЫ И ТЕХНОЛОГИИ

Профиль

Квалификация

**Бакалавр**

Форма обучения

**очная**

Общая трудоемкость

**3 ЗЕТ**

Часов по учебному плану

108

Формы контроля в семестрах:

в том числе:

зачет с оценкой 5

аудиторные занятия

51

самостоятельная работа

57

### Распределение часов дисциплины по семестрам

Семестр (<Курс>.<Семестр на курсе>)	5 (3.1)		Итого	
	Неделя			
Вид занятий	УП	РП	УП	РП
Лекции	17	17	17	17
Практические	34	34	34	34
Итого ауд.	51	51	51	51
Контактная работа	51	51	51	51
Сам. работа	57	57	57	57
Итого	108	108	108	108

Программу составил(и):

*ст.преп., Аюпова Анастасия Алексеевна*

Рабочая программа

**Основы Data engineering**

Разработана в соответствии с ОС ВО:

Самостоятельно устанавливаемый образовательный стандарт высшего образования - бакалавриат Федеральное государственное автономное образовательное учреждение высшего образования «Национальный исследовательский технологический университет «МИСиС» по направлению подготовки 09.03.02 ИНФОРМАЦИОННЫЕ СИСТЕМЫ И ТЕХНОЛОГИИ (приказ от 02.04.2021 г. № 119 о.в.)

Составлена на основании учебного плана:

09.03.02 Информационные системы и технологии, 09.03.02-БИСТ-23.plx , утвержденного Ученым советом НИТУ МИСИС в составе соответствующей ОПОП ВО 22.06.2023, протокол № 5-23

Утверждена в составе ОПОП ВО:

09.03.02 Информационные системы и технологии, , утвержденной Ученым советом НИТУ МИСИС 22.06.2023, протокол № 5-23

Рабочая программа одобрена на заседании

**Кафедра инфокоммуникационных технологий**

Протокол от 12.04.2023 г., №9

Руководитель подразделения Кузнецова Ксения Александровна

**1. ЦЕЛИ ОСВОЕНИЯ**

1.1	Цель освоения дисциплины – формирование компетенций в соответствии с учебным планом, а также изучение студентами современных достижений в области проектирования и создания систем, которые собирают, управляют, преобразуют и сохраняют данные в пригодном для использования состоянии
-----	---

**2. МЕСТО В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ**

Блок ОП:		Б1.О
<b>2.1</b>	<b>Требования к предварительной подготовке обучающегося:</b>	
2.1.1	Информационная безопасность	
2.1.2	Сетевые технологии	
2.1.3	Базы данных	
<b>2.2</b>	<b>Дисциплины (модули) и практики, для которых освоение данной дисциплины (модуля) необходимо как предшествующее:</b>	
2.2.1	НИР. Научно-исследовательская работа в области интеллектуальных встраиваемых систем	
2.2.2	Производственная практика по получению профессиональных умений и опыта профессиональной деятельности	
2.2.3	Проектирование информационных систем	
2.2.4	Инфокоммуникационные системы и сети	
2.2.5	Подготовка к процедуре защиты и защита выпускной квалификационной работы	
2.2.6	Преддипломная практика для выполнения выпускной квалификационной работы	
2.2.7	НИР. Научно-исследовательская работа в области программного обеспечения корпоративных информационных систем	
2.2.8	Производственная практика по получению профессиональных умений и опыта профессиональной деятельности	
2.2.9	Подготовка к процедуре защиты и защита выпускной квалификационной работы	
2.2.10	Преддипломная практика для выполнения выпускной квалификационной работы	

**3. РЕЗУЛЬТАТЫ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ, СООТНЕСЕННЫЕ С ФОРМИРУЕМЫМИ КОМПЕТЕНЦИЯМИ**

<b>УК-2: Способен собирать и интерпретировать данные и определять круг задач в рамках поставленной цели, выбирать оптимальные способы решения, исходя из действующих правовых норм, имеющихся ресурсов и ограничений, умение обосновывать принятые решения</b>
<b>Знать:</b>
УК-2-31 принципы, методы и средства анализа и структурирования профессиональной информации
<b>УК-1: Способен осуществлять поиск, критический анализ и синтез информации, умение анализировать процессы и системы с использованием соответствующих аналитических, вычислительных и экспериментальных методов, применять системный подход для решения поставленных задач</b>
<b>Уметь:</b>
УК-1-У1 анализировать профессиональную информацию, выделять в ней главное, структурировать, оформлять и представлять в виде аналитических обзоров
<b>УК-2: Способен собирать и интерпретировать данные и определять круг задач в рамках поставленной цели, выбирать оптимальные способы решения, исходя из действующих правовых норм, имеющихся ресурсов и ограничений, умение обосновывать принятые решения</b>
<b>Владеть:</b>
УК-2-В1 навыками разработки программных комплексов для решения прикладных задач

**4. СТРУКТУРА И СОДЕРЖАНИЕ**

Код занятия	Наименование разделов и тем /вид занятия/	Семестр / Курс	Часов	Формируемые индикаторы компетенций	Литература и эл. ресурсы	Примечание	КМ	Выполняемые работы
	<b>Раздел 1. Основные понятия и концепции Data Engineering</b>							
1.1	Основная терминология Data Engineering /Лек/	5	1	УК-2-31	Л1.2			
1.2	Концепция потоков данных. ELT и ETL. /Лек/	5	2	УК-2-31	Э1			

1.3	Визуализация данных, дашборды и отчетность - Business Intelligence /Лек/	5	2	УК-2-31				
1.4	Пакетная и стриминговая загрузка данных /Лек/	5	2	УК-2-31	Л2.3			
1.5	Визуализация данных в Excel. /Пр/	5	2	УК-1-У1	Э2			
1.6	Визуализация данных с помощью Yandex DataLens /Пр/	5	2	УК-1-У1	Л3.3			
1.7	Извлечение данных из текстового файла /Пр/	5	2	УК-1-У1	Э4			
1.8	Группировка данных /Пр/	5	2	УК-1-У1	Л2.2			
1.9	Инструментальные средства интеллектуального анализа данных /Пр/	5	2	УК-1-У1				
1.10	Подготовка к практическим занятиям /Ср/	5	22	УК-1-У1 УК-2-В1	Л3.2 Э5			
<b>Раздел 2. Интеграция данных и создание потоков данных (data pipelines)</b>								
2.1	Файловые хранилища, форматы файлов, концепция колоночного хранения данных в файлах. Метаданные. /Лек/	5	2	УК-2-31	Э3			
2.2	Аналитические базы данных и их отличие от реляционных. Организация колоночного хранения данных. Партиционирование и индексы. /Лек/	5	2	УК-2-31				
2.3	Основные инструменты создания потоков данных. Оркестраторы. /Лек/	5	2	УК-2-31	Л3.3			
2.4	Состав и назначение Spark-кластера /Пр/	5	2	УК-1-У1	Э1			
2.5	Создание пользовательских функций /Пр/	5	2	УК-1-У1				
2.6	Настройка количества операций отображения /Пр/	5	2	УК-1-У1	Л2.3 Э2			
2.7	Импорт данных HDFS при помощи Sqoop /Пр/	5	2	УК-1-У1				
2.8	Работа с программным интерфейсом DStream /Пр/	5	2	УК-1-У1	Л2.2 Э4			
2.9	Подготовка к практическим занятиям /Ср/	5	20	УК-1-У1 УК-2-В1	Э5			
<b>Раздел 3. Облачные вычисления (Cloud Computing)</b>								
3.1	Основные платформы облачных вычислений. Бессерверные вычисления. /Лек/	5	2	УК-2-31	Л1.1			
3.2	Службы облачных вычислений. IaaS, PaaS, SaaS. /Лек/	5	2	УК-2-31	Э2			
3.3	Создание кластера /Пр/	5	2	УК-1-У1				
3.4	Масштабирование контейнеров /Пр/	5	2	УК-1-У1				

3.5	Маршрутизация входящих соединений /Пр/	5	2	УК-1-У1	Л2.1 Э3			
3.6	Конфигурирование развертывания в uml файлах /Пр/	5	2	УК-1-У1				
3.7	Оркестрация сервисов масштабных приложений /Пр/	5	2	УК-1-У1	Л1.1 Э4			
3.8	Назначение кеша данных /Пр/	5	2	УК-1-У1	Л3.2 Э4			
3.9	Стратегии замещения данных в кеш /Пр/	5	2	УК-1-У1	Э5			
3.10	Подготовка к практическим занятиям /Ср/	5	15	УК-1-У1 УК-2 -В1	Э2			

## 5. ФОНД ОЦЕНОЧНЫХ МАТЕРИАЛОВ

### 5.1. Контрольные мероприятия (контрольная работа, тест, коллоквиум, экзамен и т.п), вопросы для самостоятельной подготовки

Код КМ	Контрольное мероприятие	Проверяемые индикаторы компетенций	Вопросы для подготовки
КМ1	Контрольная работа №1	УК-1-У1	<ol style="list-style-type: none"> <li>1. Какие задачи решает аппарат анализа?</li> <li>2. Что такое диаграмма потоков данных?</li> <li>3. Чем отличается диаграмма потоков данных от блок-схемы алгоритма?</li> <li>4. Какие элементы диаграммы потоков данных вы знаете?</li> <li>5. Как формируется иерархия диаграмм потоков данных?</li> <li>6. Какую задачу решает диаграмма потоков данных высшего (нулевого) уровня? Почему ее называют контекстной моделью?</li> <li>7. Чем нагружены вершины диаграммы потоков данных?</li> <li>8. Чем нагружены дуги диаграммы потоков данных?</li> <li>9. Как организован словарь требований?</li> <li>10. С помощью аппарата структурного анализа создайте модель системы управления летательного аппарата.</li> <li>11. С чем связана необходимость расширения диаграмм потоков данных для систем реального времени?</li> <li>12. Какие средства расширения вы знаете?</li> <li>13. Как решается проблема расширения возможностей управления на базе диаграмм потоков данных?</li> <li>14. Каковы особенности диаграммы управляющих потоков?</li> <li>15. Поясните понятие активатора процесса.</li> <li>16. Поясните понятие условия данных.</li> <li>17. Поясните понятие управляющей спецификации.</li> <li>18. Поясните понятие окна управляющей спецификации.</li> <li>19. Как организована спецификация процесса?</li> </ol>

КМ2	Контрольная работа №2	УК-1-У1	<ol style="list-style-type: none"> <li>1. Определение облака. Обязательные свойства облака.</li> <li>2. Виды облаков (модели развертывания).</li> <li>3. Модели обслуживания (основные и дополнительные).</li> <li>4. Основные достоинства технологии виртуализации.</li> <li>5. Технологии виртуализации. Виртуализация серверов (полная виртуализация)</li> <li>6. Технологии виртуализации. Виртуализация серверов (паравиртуализация)</li> <li>7. Технологии виртуализации. Виртуализация на уровне операционных систем.</li> <li>8. Технологии виртуализации. Виртуализация приложений и представлений.</li> <li>9. High availability cloud. Основные отличия от обычного облака.</li> <li>10. История и предпосылки построения публичных облаков. Особенности предоставления услуг.</li> <li>11. Основные крупные поставщики облачных услуг (Google, Amazon, Яндекс, MS).</li> <li>12. Способы организации публичного облака.</li> <li>13. Расчет затрат на облачные ресурсы на примере MS Azure.</li> <li>14. Описание и спецификация стандартных узлов.</li> <li>15. Распространенные сценарии запуска и работы с публичными облаками</li> <li>16. Особенности предоставления услуг у крупных публичных облаков.</li> <li>17. Особенности работы с MS Azure.</li> <li>18. Особенности работы с Яндекс.Облако.</li> </ol>
-----	-----------------------	---------	--

### 5.2. Перечень работ, выполняемых по дисциплине (Курсовая работа, Курсовой проект, РГР, Реферат, ЛР, ПР и т.п.)

Код работы	Название работы	Проверяемые индикаторы компетенций	Содержание работы
P1	Практические работы	УК-1-У1	Выполнение практических работ по темам практических занятий
P2	Домашнее задание	УК-2-В1;УК-1-У1	<p>Рекомендуемые темы домашних заданий</p> <ol style="list-style-type: none"> <li>1 Технологии виртуализации</li> <li>2. Примеры облачных сервисов Microsoft.</li> <li>3. Примеры облачных сервисов Google.</li> <li>4. Веб-службы в Облаке</li> </ol>

### 5.3. Оценочные материалы, используемые для экзамена (описание билетов, тестов и т.п.)

экзамен не предусмотрен

### 5.4. Методика оценки освоения дисциплины (модуля, практики. НИР)

По курсу предусмотрен зачет с оценкой. Зачет с оценкой проставляется на основе оценок текущего контроля Оценка «отлично» - обучающийся показывает глубокие, исчерпывающие знания в объеме пройденной программы, уверенно действует по применению полученных знаний на практике, грамотно и логически стройно излагает материал при ответе, умеет формулировать выводы из изложенного теоретического материала, знает дополнительно рекомендованную литературу.

Оценка «хорошо» - обучающийся показывает твердые и достаточно полные знания в объеме пройденной программы, допускает незначительные ошибки при освещении заданных вопросов, правильно действует по применению знаний на практике, четко излагает материал.

Оценка «удовлетворительно» - обучающийся показывает знания в объеме пройденной программы, ответы излагает хотя и с ошибками, но уверенно исправляемыми после дополнительных и наводящих вопросов, правильно действует по применению знаний на практике;

Оценка «неудовлетворительно» - обучающийся допускает грубые ошибки в ответе, не понимает сущности излагаемого вопроса, не умеет применять знания на практике, дает неполные ответы на дополнительные и наводящие вопросы.

Оценка «неявка» – обучающийся на зачет с оценкой не явился

## 6. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ

### 6.1. Рекомендуемая литература

#### 6.1.1. Основная литература

Авторы, составители	Заглавие	Библиотека	Издательство, год
---------------------	----------	------------	-------------------

	Авторы, составители	Заглавие	Библиотека	Издательство, год
Л1.1	Губарев В. В., Савульчик С. А., Чистяков Н. А.	Введение в облачные вычисления и технологии: учебное пособие	Электронная библиотека	Новосибирск: Новосибирский государственный технический университет, 2013
Л1.2	Крутиков В. Н., Мешечкин В. В.	Анализ данных: учебное пособие	Электронная библиотека	Кемерово: Кемеровский государственный университет, 2014

#### 6.1.2. Дополнительная литература

	Авторы, составители	Заглавие	Библиотека	Издательство, год
Л2.1	Ушакова Л. Р.	Построение корреляционной модели и статистический анализ данных по термодинамическим свойствам галогенидов щелочных металлов: выпускная квалификационная работа: студенческая научная работа	Электронная библиотека	Екатеринбург, 2016
Л2.2	Агалаков С. А.	Анализ данных в среде R: практикум	Электронная библиотека	Омск: Омский государственный университет им. Ф.М. Достоевского, 2020

#### 6.1.3. Методические разработки

	Авторы, составители	Заглавие	Библиотека	Издательство, год
Л3.1	Истомина А. П.	Анализ данных качественных исследований: практикум	Электронная библиотека	Ставрополь: Северо- Кавказский Федеральный университет (СКФУ), 2016
Л3.2	Жуковский О. И.	Информационные технологии и анализ данных: учебное пособие	Электронная библиотека	Томск: Эль Контент, 2014
Л3.3	Мельниченко А. С.	Математическая статистика и анализ данных (N 3431): учеб. пособие	Библиотека МИСиС	М.: [МИСиС], 2018

#### 6.2. Перечень ресурсов информационно-телекоммуникационной сети «Интернет»

Э1	Электронные информационные ресурсы Российской государственной библиотеки	<a href="https://www.rsl.ru/">https://www.rsl.ru/</a>
Э2	Электронно-библиотечная система "Библиокомплектатор" (IPRbooks)	<a href="http://www.bibliocomplectator.ru/available">http://www.bibliocomplectator.ru/available</a>
Э3	Научная электронная библиотека «КиберЛенинка»	<a href="https://cyberleninka.ru/">https://cyberleninka.ru/</a>
Э4	Портал российских электронных библиотек	<a href="http://www.elbib.ru">www.elbib.ru</a>
Э5	Электронно-библиотечная система «Лань» –	<a href="https://e.lanbook.com/">https://e.lanbook.com/</a>

#### 6.3 Перечень программного обеспечения

П.1	Microsoft Office
П.2	LMS Canvas
П.3	MS Teams
П.4	Консультант Плюс
П.5	WinRAR
П.6	MATLAB

#### 6.4. Перечень информационных справочных систем и профессиональных баз данных

И.1	Полнотекстовые российские научные журналы и статьи:
И.2	Научная электронная библиотека eLIBRARY <a href="https://elibrary.ru/">https://elibrary.ru/</a>
И.3	Полнотекстовые деловые публикации информгентств и прессы по 53 отраслям <a href="https://polpred.com/news">https://polpred.com/news</a>
И.4	Иностранные базы данных (доступ с IP адресов МИСиС):
И.5	аналитическая база (индексы цитирования) Web of Science <a href="https://apps.webofknowledge.com">https://apps.webofknowledge.com</a>
И.6	аналитическая база (индексы цитирования) Scopus <a href="https://www.scopus.com/">https://www.scopus.com/</a>
И.7	наукометрическая система InCites <a href="https://apps.webofknowledge.com">https://apps.webofknowledge.com</a>

И.8	научные журналы издательства Elsevier <a href="https://www.sciencedirect.com/">https://www.sciencedirect.com/</a>
И.9	Электронный ресурс

### 7. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ

Ауд.	Назначение	Оснащение
Любой корпус Компьютерный класс	Учебная аудитория для проведения практических занятий:	экран, проектор, доска, комплект учебной мебели на 30 посадочных мест, персональные компьютеры, доступ к ЭИОС университета LMS Canvas, лицензионные программы MS Teams, MS Office
Любой корпус Мультимедийная	Учебная аудитория для проведения занятий лекционного типа и/или для проведения практических занятий:	комплект учебной мебели до 36 мест для обучающихся, мультимедийное оборудование, магнитно-маркерная доска, рабочее место преподавателя, ПК с доступом к ИТС «Интернет», ЭИОС университета через личный кабинет на платформе LMS Canvas, лицензионные программы MS Office, MS Teams, ESET Antivirus
Читальный зал №3 (Б)		комплект учебной мебели на 44 места для обучающихся, МФУ Xerox VersaLink B7025 с функцией масштабирования текстов и изображений, 8 ПК с доступом к ИТС «Интернет», ЭИОС университета через личный кабинет на платформе LMS Canvas, лицензионные программы MS Office, MS Teams, ESET Antivirus.

### 8. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ

Организация занятий направлена на изучение студентами общих вопросов изучаемого курса.

Предусматриваются домашние задания по различным разделам курса в форме подготовки мультимедийных докладов.

Проведение аудиторных занятий предусматривает использование в учебном курсе активных и интерактивных технологий:

- проведение лекций с использованием интерактивных и мультимедийных технологий (презентация в формате MS PowerPoint);
- использование при проведении занятий специализированной лаборатории с возможностью проведения занятий в интерактивной форме;
- использование при проведении занятий активных форм обучения - учебных видеоматериалов и компьютерных тренажеров.

Дисциплина относится к основополагающим и требует значительного объема самостоятельной работы.

Отдельные учебные вопросы выносятся на самостоятельную проработку и контролируются посредством текущей аттестации.

При этом организуются групповые и индивидуальные консультации.

Качественное освоение дисциплины возможно только при систематической самостоятельной работе, что поддерживается системой текущей и рубежной аттестации.

В связи с использованием во время занятий мультимедийных технологий для проведения практических занятий требуется специализированная мультимедийная аудитория с возможностью показа видеоматериалов с аудиосопровождением и доступом к сети Интернет. Аудитория выбирается в зависимости от количества студентов, изучающих в текущем семестре данную дисциплину, при численности студентов до 30 человек рекомендуется аудитория Л-810, при численности менее 14 человек – Л-813