

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ОБРАЗОВАНИЯ

«Национальный исследовательский ядерный университет «МИФИ»

Саровский физико-технический институт -

филиал федерального государственного автономного образовательного учреждения высшего
образования «Национальный исследовательский ядерный университет «МИФИ»

(СарФТИ НИЯУ МИФИ)

ФИЗИКО-ТЕХНИЧЕСКИЙ ФАКУЛЬТЕТ

Кафедра «Цифровых технологий»

УТВЕРЖДАЮ

Декан ФТФ, член корр. РАН, д.ф-м.н.

_____ **А.К. Чернышев**

«___» _____ **2023 г.**

РАБОЧАЯ ПРОГРАММА УЧЕБНОЙ ДИСЦИПЛИНЫ

Технология больших данных (big data)

наименование дисциплины

Направление подготовки (специальность)	<u>09.03.02 Информационные системы и технологии</u>
Наименование образовательной программы	<u>Инновационные технологии комплексной автоматизации и сквозного управления жизненным циклом</u>
Квалификация (степень) выпускника	<u>бакалавр</u>
Форма обучения	<u>очная</u>
Программа одобрена на заседании кафедры	Зав. кафедрой ЦТ
протокол № от 2023 г.	_____ О.В. Кривошеев «___» _____ 2023 г.

г. Саров, 2023 г.

Программа переутверждена на 202____/202____ учебный год с изменениями в соответствии с семестровыми учебными планами академических групп ФИТЭ, ФТФ на 202____/202____ учебный год.

Заведующий кафедрой ЦТ

О.В. Кривошеев

Программа переутверждена на 202____/202____ учебный год с изменениями в соответствии с семестровыми учебными планами академических групп ФИТЭ, ФТФ на 202____/202____ учебный год.

Заведующий кафедрой ЦТ

О.В. Кривошеев

Программа переутверждена на 202____/202____ учебный год с изменениями в соответствии с семестровыми учебными планами академических групп ФИТЭ, ФТФ на 202____/202____ учебный год.

Заведующий кафедрой ЦТ

О.В. Кривошеев

Программа переутверждена на 202____/202____ учебный год с изменениями в соответствии с семестровыми учебными планами академических групп ФИТЭ, ФТФ на 202____/202____ учебный год.

Заведующий кафедрой ЦТ

О.В. Кривошеев

Семестр	в форме прак- тической подго- товки	Трудоемкость, кред.	Общий объем курса, час.	Лекции, час.	Практич. заня- тия, час.	Лаборат. работы, час.	СРС, час.	КР/ КП	Форма(ы) кон- троля, экз./зач./ЗсО/	Интерактивные часы
4	16	3	108	16	16	-	76	-	Зачёт	12
ИТОГО	16	3	108	16	16	-	76	-	-	12

АННОТАЦИЯ

Курс «Технология больших данных» содержит обзорную информацию о теоретических и практических аспектах технологий, в основе которых лежат принципы получения, преобразования, распределенного хранения и обработки, а также анализа больших объемов данных. Курс направлен на подготовку квалифицированных выпускников, умеющих обоснованно и результативно использовать теоретические знания на практике.

1. ЦЕЛИ И ЗАДАЧИ ОСВОЕНИЯ УЧЕБНОЙ ДИСЦИПЛИНЫ

Цель изучения дисциплины «Технология больших данных» – заключается в формировании представления о больших данных, определении, для каких целей их можно использовать в конкретной организации. Так же подготовка специалистов подразумевает уяснение сути как технических, так и бизнес-требований к применению технологии больших данных.

Задачи дисциплины:

- сформировать чёткое представление о технологии больших данных;
- студент должен знать предпосылки, нормативную базу, а также зарубежный опыт в этой области;
- должен знать перечень и сущность ключевых технологий, составляющих основу технологии больших данных;
- сформировать представление о типах больших данных, новом использовании старых решений, компоненты технологии больших данных;
- должен знать общие моменты по управлению большими данными, ключевые понятия в области больших данных;
- должен уметь анализировать возможность и целесообразность применения технологий больших данных в тех или иных областях.

2. МЕСТО УЧЕБНОЙ ДИСЦИПЛИНЫ В СТРУКТУРЕ ООП ВО

«Технология больших данных» является дисциплиной профиля «Инновационные технологии комплексной автоматизации и сквозного управления жизненным циклом» ОС НИЯУ МИФИ по направлению 09.03.02 «Информационные системы и технологии».

Дисциплина направлена на формирование у студентов профессиональных компетенций в области применения технологии больших данных в современном мире. В ходе обучения у студентов формируются профессионально-ориентированные навыки, позволяющие применять на практике полученные знания в области практического решения задач, связанных с применением больших данных.

3. ФОРМИРУЕМЫЕ КОМПЕТЕНЦИИ И ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОБУЧЕНИЯ

Ожидается, что в результате освоения дисциплины студент приобретет следующие компетенции:

Профессиональные компетенции (ПК)

в соответствии с задачами и объектами (областями знаний) профессиональной деятельности:

Задача профессиональной деятельности (ЗПД)	Объект или область знания	Код и наименование профессиональной компетенции	Код и наименование индикатора достижения профессиональной компетенции
Тип задачи профессиональной деятельности: производственно-технологический			
разработка и внедрение технологий разработки объектов профессиональной деятельности в различных областях и сферах деятельности	информационные процессы, технологии, системы и сети, их инструментальное (программное, техническое, организационное) обеспечение, способы и методы проектирования, отладки, производства и эксплуатации информационных технологий и систем в различных областях и сферах деятельности.	ПК-5 Способен осуществлять моделирование процессов и систем на основе системного анализа предметной области <i>Основание:</i> Профессиональный стандарт «06.015. Специалист по информационным системам»	З-ПК-5 Знать: основные принципы системного подхода; методы моделирования процессов и систем У-ПК-5 Уметь: проводить анализ предметной области и осуществлять ее формальное представление в виде модели В-ПК-5 Владеть: инструментальными средствами моделирования
разработка и внедрение технологий разработки объектов профессиональной деятельности в различных областях и сферах деятельности	информационные процессы, технологии, системы и сети, их инструментальное (программное, техническое, организационное) обеспечение, способы и методы проектирования, отладки, производства и эксплуатации информационных технологий и систем в различных областях и сферах деятельности.	ПК-2.1 Способен обеспечения эффективной работы баз данных, включая развертывание, сопровождение, оптимизация функционирования баз данных, являющихся частью различных информационных систем <i>Основание:</i> Профессиональный стандарт «06.011. Администратор баз данных»	З-ПК-2.1 знать компоненты программно-аппаратного обеспечения БД и подбор средств для их мониторинга У-ПК-2.1 уметь выбирать способ действия в изменяющихся условиях рабочей ситуации; контролировать, оценивать и корректировать свои действия В-ПК-2.1 владеть Навыками сопровождения и оптимизации функционирования баз данных

4. СТРУКТУРА И СОДЕРЖАНИЕ УЧЕБНОЙ ДИСЦИПЛИНЫ*

№ п/п	Наименование раздела /темы дисциплины	№ недели	Виды учебной работы						
			Лекции	Практ. заня- тия/ семинары	Лаб. работы	СРС	Текущий кон- троль (форма)*	Максималь- ный балл (см. п. 5.3)	
			16	16	-	76			
Семестр 4									
1.1.	Раздел 1-2. Введение в большие данные. Техно- логические основы боль- ших данных.	1-2	2	2	-	10	-	-	
1.2	Раздел 3. Управление большими данными.	3-4	2	2	-	10	УО	5	
1.3	Раздел 4. Аналитика и большие данные.	5-6	2	2	-	10	УО	5	
Рубежный контроль		7	КР						10
2.1	Раздел 5. Применение больших данных. Инте- грация источников дан- ных. Работа с потоками данных в режиме реаль- ного времени и обработ- ка сложных событий.	7, 8,	2	2	-	10	УО	5	
2.2	Раздел 5. Применение больших данных. Ис- пользование больших данных в операционной деятельности. Примене- ние больших данных в вашей организации. Без- опасность и управление средами больших дан- ных.	9-11	2	2	-	10	УО	5	

№ п/п	Наименование раздела /темы дисциплины	№ недели	Виды учебной работы					
			Лекции	Практ. заня- тия/ семинары	Лаб. работы	СРС	Текущий кон- троль (форма)*	Максималь- ный балл (см. п. 5.3)
			16	16	-	76		
2.3	Раздел 6. Решения на ос- нове больших данных.	12-14	2	2	-	10	УО	5
2.4	Раздел 7. Лучшие спосо- бы работы с большими данными. Рекомендации.	15-16	4	4	-	16	УО	5
Рубежный контроль		16						Тест
Промежуточная аттестация			зачет					0 - 50
Посещаемость								5
Итого:			16	16	-	76		100

*Сокращение наименований форм текущего, рубежного и промежуточного контроля:

УО – устный опрос

Контр. – контрольная работа

Тест – тестирование (письменный опрос)

ДЗ – домашнее задание

РГР – расчетно-графическая работа

Э/Зач/ЗсО – экзамен/зачет/зачет с оценкой и др.

4.2. Содержание дисциплины, структурированное по разделам (темам)

Лекционный курс

№	Наименование раздела /темы дисциплины	Содержание
Раздел 1. Лекционный материал		
1.1.	Раздел 1-2. Введение в большие данные. Технологические основы больших данных.	Эволюция управления данными. Определение больших данных и их технологических компонентов. Типы больших данных. Структурированные и неструктурированные данные. Данные в реальном времени и неоперативные данные. Распределенные вычисления при работе с большими данными. Элементы стека больших данных. Интеграция аналитика и приложений с большими данными. Понятие виртуализации. Влияние виртуализации на большие данные. Абстракция. Облачные вычисления и их роль в больших данных.
1.2	Раздел 3. Управление большими данными.	Различия между базами данных в мире больших данных. Использование MapReduce для анализа больших данных. Основы Hadoop. Повышение эффективности распределенной файловой системы Hadoop с использованием языков программирования и программных инструментов. Разработка приложений больших данных для аналитики. Использование программно-аппаратных комплексов при управлении большими данными.
1.3	Раздел 4. Аналитика и большие данные.	Большие данные для составления отчетов, визуализации и прогнозного анализа. Примеры структурной аналитики. Неструктурированные данные. Текстовая аналитика. Интеграция неструктурированных и структурированных данных. Модели использования больших данных.
1.4	Раздел 5. Применение больших данных.	Определение и интеграция нужных данных для реализации больших данных. Оценка точности результатов. Поточные данные. Обработка сложных событий. Ввод больших данных в эксплуатацию. План действия по внедрению больших данных.
1.5	Раздел 6. Решения на осно-	Обзор моделей использования больших данных. Примене-

	ве больших данных.	ние больших данных для обнаружения мошенничества. Использование отзывов клиентов в социальных сетях при принятии корпоративных решений. Совершенствование медицинской диагностики с помощью больших данных.
1.6	Раздел 7. Лучшие способы работы с большими данными. Рекомендации.	Осознание ваших целей. Определение плана действий. Оценка имеющихся данных и потребности в данных. Проверка полезности данных. Четкий начальный план. Взаимодействие с бизнесом. Интеграция больших данных в процессы. Создание и выполнение планов.
Раздел 2. Практические работы		
2.1	Занятие 1.	Разработка успешных архитектур управления большими данными.
2.2	Занятие 2.	История распределенных вычислений. Основы распределенных вычислений.
2.3	Занятие 3.	Облачные сервисы и большие данные.
2.4	Занятие 4.	Значение реляционных СУБД в управлении больших данных.
2.5	Занятие 5..	Основные принципы MapReduce.
2.6	Занятие 6.	Распределенная файловая система Hadoop (HDFS).
2.7	Занятие 7.	Основные источники больших данных.
2.8	Занятие 8.	Применение больших данных. Интеграция, работа с потоками и пр.
План лабораторных занятий		
	Лабораторные занятия не предусмотрены	

4.3. Перечень учебно-методического обеспечения для самостоятельной работы студентов

1. Гурвиц, Джудит. Просто о больших данных / Гурвиц Джудит, Ньюджент Алан, Халпер Ферн, Кауфман Марсия: [перевод с английского]. - Москва: Эксмо, 2015. - 400 с. - (Библиотека Сбербанка. Т. 58). ISBN 978-5-699-85806-4

2. Силен Дэви, Мейсман Арно, Али Махомед. Основы Data Science и Big Data. Python и наука о данных. – СПб.: Питер, 2017. – 336 с.: ил. – (Серия «Библиотека программиста»). ISBN 978-5-496-02517-1

5. ОЦЕНОЧНЫЕ СРЕДСТВА ДЛЯ ТЕКУЩЕГО КОНТРОЛЯ УСПЕВАЕМОСТИ, ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ ПО ИТОГАМ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Фонд оценочных средств по дисциплине обеспечивает проверку освоения планируемых результатов обучения (компетенций и их индикаторов) посредством мероприятий текущего, рубежного и промежуточного контроля по дисциплине.

5.1. Паспорт фонда оценочных средств по дисциплине

Связь между формируемыми компетенциями и формами контроля их освоения представлена в следующей таблице:

Раздел	Темы занятий	Компетенция	Индикаторы освоения	Текущий контроль, неделя
Семестр 4				
Раздел 1	Раздел 1-2. Введение в большие данные. Технологические основы больших данных.	ПК-5, ПК-2.1	З - ПК-5, У- ПК-5, В- ПК-5 З-ПК-2.1; У-ПК-2.1; В-ПК-2.1	УО
	Раздел 3. Управление большими данными.			
	Раздел 4. Аналитика и большие данные.			
Рубежный контроль		ПК-5, ПК-2.1	З - ПК-5, У- ПК-5, В- ПК-5 З-ПК-2.1; У-ПК-2.1; В-ПК-2.1	КР
Раздел 2	Раздел 5. Применение больших данных.	ПК-5, ПК-2.1	З - ПК-5, У- ПК-5, В- ПК-5 З-ПК-2.1; У-ПК-2.1; В-ПК-2.1	УО
	Раздел 6. Решения на основе больших данных.			
	Раздел 7. Лучшие способы работы с большими данными. Рекомендации.			

Рубежный контроль	ПК-5, ПК-2.1	З - ПК-5, У- ПК-5, В- ПК-5 З-ПК-2.1; У-ПК-2.1; В-ПК-2.1	Тест
Промежуточная аттестация	ПК-5, ПК-2.1	З - ПК-5, У- ПК-5, В- ПК-5 З-ПК-2.1; У-ПК-2.1; В-ПК-2.1	Экзамен

5.2. Примерные контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков (или) опыта деятельности, характеризующие этапы формирования компетенций в процессе освоения образовательной программы

5.2.1. Оценочные средства для текущего контроля

5.2.1.1. Типовые задания к контрольной работе

Вариант 1

1. Эволюция управления данными, этапы развития.
2. Разработка успешных архитектур управления большими данными.
3. Обзор истории распределенных вычислений.

Вариант 2

1. Основные источники больших данных.
2. Большие данные как инструмент бизнес-планирования. Сколько этапов, их суть.
3. Характеристика типов больших данных.

Вариант 3

1. Оптимизация бизнес-процессов с помощью анализа больших данных, практические примеры.
2. Десять лучших способов работы с большими данными.
3. Десять рекомендаций: что следует и чего не следует делать в работе с большими данными.

Вариант 4

1. Программно-аппаратные комплексы и хранилища больших данных.
2. Что такое аналитика больших данных.
3. Облачные сервисы и большие данные.

5.2.1.2. Темы для подготовки рефератов

1. Предиктивная аналитика в задачах цифрового бизнеса.
2. Системы поддержки принятия решений на основе когнитивных технологий.
3. Классы задач Data Mining.
4. Нейросети и их применение в прогнозировании.
5. Прогнозирование средствами Data Mining.
6. Data Mining в анализе клиентов.
7. Решение задачи классификации на основе моделей нейронных сетей.
8. Решение задачи классификации с помощью деревьев решений.
9. Применение карт Кохонена для сегментации клиентов интернет-магазина.
10. Обогащение данных о клиенте на основе Social Mining.
11. Таргетированная реклама на основе Social Mining.
12. Поиск документов средствами Text Mining.
13. Анализ использования веб-ресурсов с помощью Web Mining.
14. Рынок BI-платформ.
15. Современные тенденции развития систем BI.
16. OLAP как многомерный анализ данных.
17. Проблемы анализа неструктурированных данных.
18. Технология вычислений in-memory.
19. Большие данные: источники, характеристики.
20. Технология больших данных Apache Hadoop.

5.2.3.1 Примерные вопросы к зачёту

1. Этапы развития управления данными.
2. Что такое большие данные.
3. Что такое структурированные данные.
4. Что такое неструктурированные данные.
5. Объединение больших данных.
6. Что такое DARPA. Достоинства согласованной модели.
7. Стек больших данных, количество уровней стека, краткая характеристика каждого уровня.
8. Уровень избыточной инфраструктуры стека больших данных.
9. Уровень инфраструктуры безопасности стека больших данных.
10. Уровень операционных баз данных стека больших данных.
11. Уровень служб и инструментов для организации данных стека больших данных.
12. Уровень аналитического хранилища данных стека больших данных.

13. Облачные модели развертывания и взаимодействия.
14. Что такое виртуализация. В чем заключается суть виртуализации.
15. Значение реляционных СУБД в среде больших данных.
16. Что такое MapReduce, основные принципы.
17. Типы операционных баз данных.
18. Что такое Hadoop.
19. Интеграция больших данных и традиционных хранилищ.
20. Специальные подходы к анализу больших данных.
21. Области применения больших данных.

5.3. Шкалы оценки образовательных достижений

Рейтинговая оценка знаний является интегральным показателем качества теоретических и практических знаний и навыков студентов по дисциплине и складывается из оценок, полученных в ходе текущего контроля и промежуточной аттестации.

Результаты текущего контроля и промежуточной аттестации подводятся по шкале балльно-рейтинговой системы.

Шкала каждого контрольного мероприятия лежит в пределах от 0 до установленного максимального балла включительно. Итоговая аттестация по дисциплине оценивается по 100-балльной шкале и представляет собой сумму баллов, заработанных студентом при выполнении заданий в рамках текущего и промежуточного контроля.

Итоговая оценка выставляется в соответствии со следующей шкалой:

Сумма баллов	Оценка по 4-ех балльной шкале	Оценка ECTS	Требования к уровню освоению учебной дисциплины
90-100	5 – «отлично»	A	Оценка «отлично» выставляется студенту, если он глубоко и прочно усвоил программный материал, исчерпывающе, последовательно, четко и логически стройно его излагает, умеет тесно увязывать теорию с практикой, использует в ответе материал монографической литературы.
85-89	4 – «хорошо»	B	Оценка «хорошо» выставляется студенту, если он твёрдо знает материал, грамотно и по существу излагает его, не допуская существенных неточностей в ответе на вопрос.
75-84		C	
70-74		D	
65-69	3 – «удовлетворительно»	E	Оценка «удовлетворительно» выставляется студенту, если он имеет знания только основного материала, но не усвоил его деталей, допускает неточности, недостаточно правильные формулировки, нарушения логической последовательности в изложении программного материала.
60-64			
Ниже 60	2 – «неудовлетворительно»	F	Оценка «неудовлетворительно» выстав-

	но»		ляется студенту, который не знает значительной части программного материала, допускает существенные ошибки. Как правило, оценка «неудовлетворительно» ставится студентам, которые не могут продолжить обучение без дополнительных занятий по соответствующей дисциплине.
--	-----	--	--

6. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ УЧЕБНОЙ ДИСЦИПЛИНЫ

ОСНОВНАЯ ЛИТЕРАТУРА:

1. Гурвиц, Джудит. Просто о больших данных / Гурвиц Джудит, Ньюджент Алан, Халпер Ферн, Кауфман Марсия: [перевод с английского]. - Москва: Эксмо, 2015. - 400 с. - (Библиотека Сбербанка. Т. 58). ISBN 978-5-699-85806-4
2. Силен Дэви, Мейсман Арно, Али Махомед. Основы Data Science и Big Data. Python и наука о данных. – СПб.: Питер, 2017. – 336 с.: ил. – (Серия «Библиотека программиста»). ISBN 978-5-496-02517-1

7 МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ УЧЕБНОЙ ДИСЦИПЛИНЫ

Лекционные занятия проходят в аудитории, оснащенной компьютером и проектором. Каждая лекция сопровождается презентацией, содержащей теоретический материал.

8. ОБРАЗОВАТЕЛЬНЫЕ ТЕХНОЛОГИИ

При чтении лекционного материала используется электронное сопровождение курса: справочно-иллюстративный материал воспроизводится и озвучивается в аудитории с использованием проектора и переносного компьютера в реальном времени.

На сайте кафедры также находится методический и справочный материал.

9. МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ СТУДЕНТАМ ПО ОРГАНИЗАЦИИ ИЗУЧЕНИЯ ДИСЦИПЛИНЫ

Примерным учебным планом на изучение дисциплины отводится один семестр. В конце предусмотрен зачет.

Подготовка семинарским занятиям является разделом самостоятельной работы и учитывается в балльно-рейтинговой системе.

Программа составлена в соответствии с требованиями ОС ВО НИЯУ МИФИ к обязательному минимуму содержания основной образовательной программы по направлению подготовки 09.03.02 Информационные системы и технологии.

Программу составил:

Рецензент: