

Документ подписан простой электронной подписью  
Информация о владельце:  
ФИО: Косенок Сергей Михайлович  
Должность: ректор  
Дата подписания: 24.06.2026 06:56:31  
Уникальный программный ключ:  
e3a68f3eaa1e62674b54f4998099d3d6bfdcf836

Центр по управлению качеством образования  
Ханты-Мансийского автономного округа-Югры  
"Сургутский государственный университет"

УТВЕРЖДАЮ  
Проректор по УМР

\_\_\_\_\_ Е.В. Коновалова

11 июня 2026 г., протокол УМС №5

## Большие данные рабочая программа дисциплины (модуля)

Закреплена за кафедрой **Информатики и вычислительной техники**

Учебный план b090302-БезопИнфСист-26-4.plx  
09.03.02 ИНФОРМАЦИОННЫЕ СИСТЕМЫ И ТЕХНОЛОГИИ  
Направленность (профиль): Безопасность информационных систем и технологий

Квалификация **Бакалавр**

Форма обучения **очная**

Общая трудоемкость **2 ЗЕТ**

Часов по учебному плану 72  
в том числе:  
аудиторные занятия 32  
самостоятельная работа 40

Виды контроля в семестрах:  
зачет 7  
контрольная работа 7

### Распределение часов дисциплины по семестрам

Семестр (<Курс>.<Семестр на курсе>)	7 (4.1)		Итого	
	уп	рп		
Неделя	17	2/6		
Вид занятий	уп	рп	уп	рп
Лекции	16	16	16	16
Практические	16	16	16	16
Итого ауд.	32	32	32	32
Контактная работа	32	32	32	32
Сам. работа	40	40	40	40
Итого	72	72	72	72

Программу составил(и):

*преподаватель, Кучин Иван Александрович*

Рабочая программа дисциплины

**Большие данные**

разработана в соответствии с ФГОС:

Федеральный государственный образовательный стандарт высшего образования - бакалавриат по направлению подготовки 09.03.02 Информационные системы и технологии (приказ Минобрнауки России от 19.09.2017 г. № 926)

составлена на основании учебного плана:

09.03.02 ИНФОРМАЦИОННЫЕ СИСТЕМЫ И ТЕХНОЛОГИИ

Направленность (профиль): Безопасность информационных систем и технологий

утвержденного учебно-методическим советом вуза от 11.06.2026 протокол № 5.

Рабочая программа одобрена на заседании кафедры

**Информатики и вычислительной техники**

Зав. кафедрой к. ф-м. н. Лысенкова Светлана Александровна

**1. ЦЕЛИ ОСВОЕНИЯ ДИСЦИПЛИНЫ**

1.1	Сформировать у студентов системные знания, практические умения и навыки работы с большими данными, включая сбор, хранение, обработку и анализ.
-----	--

**2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ООП**

Цикл (раздел) ООП:	ФТД
<b>2.1</b>	<b>Требования к предварительной подготовке обучающегося:</b>
2.1.1	Алгоритмы и языки программирования
2.1.2	Базы данных
2.1.3	Статистические методы и модели управления
2.1.4	Информационные технологии
<b>2.2</b>	<b>Дисциплины и практики, для которых освоение данной дисциплины (модуля) необходимо как предшествующее:</b>
2.2.1	Информационная безопасность и защита информации
2.2.2	Инструментальные средства информационных систем
2.2.3	Безопасность информационных систем
2.2.4	Разработка мобильных приложений

**3. КОМПЕТЕНЦИИ ОБУЧАЮЩЕГОСЯ, ФОРМИРУЕМЫЕ В РЕЗУЛЬТАТЕ ОСВОЕНИЯ ДИСЦИПЛИНЫ (МОДУЛЯ)**

**ПК-1.1: Демонстрирует знания основных методов, моделей и алгоритмов исследования информационных систем и технологий**

**ПК-1.2: Осуществляет выбор методов, моделей исследования информационных систем**

**ПК-2.1: Демонстрирует знания методов, алгоритмов и технологий интеграции программных модулей и компонент**

**ПК-2.2: Применяет на практике методы, алгоритмы и технологии интеграции программных модулей и компонент**

**ПК-2.3: Владеет технологиями интеграции программных модулей и компонент**

**ПК-4.1: Демонстрирует знания методов и технологий обеспечения функционирования баз данных**

**ПК-4.2: Разрабатывает алгоритмы предотвращения потерь и повреждений данных**

**ПК-4.3: Обеспечивает информационную безопасность**

**В результате освоения дисциплины обучающийся должен**

<b>3.1</b>	<b>Знать:</b>
3.1.1	- основные понятия, характеристики и принципы работы с большими данными;
3.1.2	- признаки больших данных: объем, скорость поступления, разнообразие, достоверность и ценность данных;
3.1.3	- основные источники больших данных и форматы их представления;
3.1.4	- методы сбора, хранения, обработки и анализа больших данных;
3.1.5	- основы работы с реляционными и нереляционными базами данных;
3.1.6	- принципы построения распределенных систем хранения и обработки данных;
3.1.7	- назначение технологий и инструментов обработки больших данных, включая Hadoop, Spark, NoSQL-СУБД;
3.1.8	- основные требования к обеспечению безопасности данных при их хранении, обработке и передаче;
3.1.9	- принципы интеграции программных модулей и компонентов при разработке информационных систем, работающих с большими данными.

<b>3.2</b>	<b>Уметь:</b>
3.2.1	- определять источники данных для решения прикладных задач в области информационных систем;
3.2.2	- выполнять сбор, предварительную обработку и структурирование данных;
3.2.3	- выбирать подходящие методы и инструменты хранения и обработки больших данных;
3.2.4	- применять средства анализа данных для выявления закономерностей и получения полезной информации;
3.2.5	- использовать базы данных и программные средства для обработки больших массивов информации;
3.2.6	- проектировать простые схемы обработки данных в информационных системах;
3.2.7	- выполнять интеграцию программных компонентов, связанных с обработкой и хранением данных;
3.2.8	- оценивать качество данных и выявлять ошибки, дублирование и неполноту информации;

#### 4. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

Код занятия	Наименование разделов и тем /вид занятия/	Семестр / Курс	Часов	Компетенции	Литература	Примечание
	<b>Раздел 1. Раздел 1. Введение в Большие данные. Анализ Больших данных</b>					
1.1	Введение в большие данные. Модель 5V/7V /Лек/	7	2	ПК-4.1 ПК-4.2 ПК-4.3 ПК-1.1	Л1.1 Л1.3Л2.1Л3.1	
1.2	Знакомство с инструментами. Установка и настройка окружения (Anaconda + Jupyter + PySpark). Работа с первыми датасетами /Пр/	7	4	ПК-4.2	Л1.2 Л1.3Л2.1Л3.1	
1.3	Экосистема Hadoop. Архитектура HDFS, принципы распределённого хранения, отказоустойчивость. /Лек/	7	2	ПК-4.1 ПК-4.2 ПК-4.3 ПК-1.1	Л1.2 Л1.3Л2.1Л3.1	
1.4	Работа с HDFS: загрузка, копирование, удаление данных. /Пр/	7	4	ПК-4.1 ПК-4.2 ПК-4.3	Л1.2Л3.1	
1.5	Модель программирования MapReduce. Примеры задач, ограничения и причины появления Spark. /Лек/	7	2	ПК-4.3 ПК-1.1 ПК-1.2 ПК-2.1	Л1.2 Л1.3Л2.1Л3.1	
1.6	Реализация алгоритма WordCount с помощью MapReduce /Пр/	7	2	ПК-4.1 ПК-4.2 ПК-4.3 ПК-1.1	Л1.2Л2.1Л3.1	
1.7	Работа с RDD: создание, трансформации, действия. Кэширование. /Пр/	7	2	ПК-4.1 ПК-4.2 ПК-4.3 ПК-1.1	Л1.2Л2.1Л3.1	
1.8	Apache Spark как универсальный движок. RDD API, DAG, планировщик задач /Лек/	7	2	ПК-4.1 ПК-4.2 ПК-4.3 ПК-1.1	Л1.2Л3.1	
	<b>Раздел 2. Раздел 2. Хранение больших данных и распределённая обработка</b>					
2.1	Структурированная обработка в Spark. DataFrame, SQL, оптимизатор Catalyst, Tungsten /Лек/	7	2	ПК-1.2 ПК-2.1 ПК-2.2	Л1.2Л2.1Л3.1	
2.2	Работа с DataFrame: загрузка данных (CSV, Parquet, JSON), очистка, фильтрация, агрегация /Пр/	7	2	ПК-1.1 ПК-1.2 ПК-2.1 ПК-2.2	Л1.2Л2.1Л3.1	
2.3	Потоковая обработка Spark. Интеграция с Kafka /Лек/	7	2	ПК-1.1 ПК-1.2 ПК-2.1 ПК-2.2	Л1.2Л2.1Л3.1	
2.4	Настройка Kafka + Spark Streaming. Обработка потока логов в реальном времени. /Пр/	7	2	ПК-4.2 ПК-4.3 ПК-1.1 ПК-1.2 ПК-2.1	Л1.2Л2.1Л3.1	
	<b>Раздел 3. Раздел3. Машинное обучение на больших данных</b>					

3.1	Масштабируемое машинное обучение. MLlib, алгоритмы, распределённое обучение, оценка моделей /Лек/	7	2	ПК-4.1 ПК-4.2 ПК-4.3 ПК-1.1	Л1.3Л2.1Л3.1	
3.2	Построение и оценка моделей MLlib на большом датасете. Pipeline API /Пр/	7	0	ПК-1.2 ПК-2.1 ПК-2.2 ПК-2.3	Л1.2 Л1.3Л2.1Л3.1	
3.3	Контрольная работа /Контр.раб./	7	2	ПК-4.1 ПК-4.2 ПК-4.3 ПК-1.1	Л1.3Л3.1	
3.4	Современный стек: Lakehouse-архитектура, Delta Lake, Iceberg. Безопасность, мониторинг, российское законодательство /Лек/	7	2	ПК-1.2 ПК-2.1 ПК-2.2 ПК-2.3	Л1.2 Л1.3Л2.1Л3.1	
3.5	Итоговый мини-проект: разработка полного пайплайна обработки больших данных. Защита проекта /Ср/ зачет	7	40	ПК-4.1 ПК-4.2 ПК-4.3 ПК-1.1 ПК-1.2 ПК-2.1 ПК-2.2 ПК-2.3	Л1.2 Л1.3Л2.1Л3.1	

### 5. ОЦЕНОЧНЫЕ СРЕДСТВА

#### 5.1. Оценочные материалы для текущего контроля и промежуточной аттестации

Представлены отдельным документом

#### 5.2. Оценочные материалы для диагностического тестирования

Представлены отдельным документом

### 6. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

#### 6.1. Рекомендуемая литература

##### 6.1.1. Основная литература

	Авторы, составители	Заглавие	Издательство, год	Колич-во
Л1.1	Тюрин Ю. Н., Макаров А. А.	Анализ данных на компьютере: учебное пособие по направлениям "Математика", "Математика. Прикладная математика"	М.: Форум, 2008	3
Л1.2	Макшанов А. В., Журавлев А. Е., Тындыкарь Л. Н.	Большие данные. Big Data: учебник для вузов	Санкт-Петербург: Лань, 2022, электронный ресурс	1
Л1.3	Макшанов А. В., Журавлев А. Е., Тындыкарь Л. Н.	Большие данные. Big Data: учебник для вузов	Санкт-Петербург: Лань, 2023, электронный ресурс	1

##### 6.1.2. Дополнительная литература

	Авторы, составители	Заглавие	Издательство, год	Колич-во
Л2.1	Ланских Ю. В., Ланских В. Г., Родионов К. В.	Введение в большие данные: учеб. пособие	Киров: ВятГУ, 2023, электронный ресурс	1

##### 6.1.3. Методические разработки

	Авторы, составители	Заглавие	Издательство, год	Колич-во
Л3.1	Кучин И. А.	Технологии обработки и анализа больших данных: учебно-методическое пособие	Сургут: Издательский центр СурГУ, 2025, электронный ресурс	1

#### 6.2. Перечень ресурсов информационно-телекоммуникационной сети "Интернет"

Э1	Документация Apache hadoop <a href="http://hadoop.apache.org">hadoop.apache.org</a>
Э2	Документация Apache spark <a href="http://spark.apache.org">spark.apache.org</a>

ЭЗ	Документация Jupyter <a href="https://docs.jupyter.org">docs.jupyter.org</a>
<b>6.3.1 Перечень программного обеспечения</b>	
6.3.1.1	операционные системы Microsoft, пакет прикладных программ Microsoft Office
6.3.1.2	Операционная система РЕД ОС
6.3.1.3	LibreOffice- Свободно распространяемое ПО, Мой Офис
<b>6.3.2 Перечень информационных справочных систем</b>	
6.3.2.1	Информационно-правовой портал Гарант.ру <a href="http://www.garant.ru">http://www.garant.ru</a>
6.3.2.2	Справочно-правовая система Консультант Плюс <a href="http://www.consultant.ru/">http://www.consultant.ru/</a>
6.3.2.3	Национальная электронная библиотека (НЭБ)
6.3.2.4	Евразийская патентная информационная система (ЕАПАТИС)
6.3.2.5	Научная электронная библиотека eLIBRARY.RU
6.3.2.6	Гарант
6.3.2.7	Техэксперт

## 7. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

7.1	Учебные аудитория для проведения занятий лекционного типа, занятий семинарского типа (лабораторных занятий), групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации оснащена: комплект специализированной учебной мебели, маркерная (меловая) доска, комплект переносного мультимедийного оборудования - компьютер, проектор, проекционный экран, компьютеры с возможностью выхода в Интернет и доступом в электронную информационно-образовательную среду. Обеспечен доступ к сети Интернет и в электронную информационную среду организации.
-----	---