

Министерство науки и высшего образования РФ  
Федеральное государственное автономное образовательное учреждение  
высшего образования  
«СИБИРСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»

## **РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)**

### **Б1.В.14 ТЕХНОЛОГИИ BIG DATA**

Направление подготовки (специальность) 09.03.03 Прикладная информатика

Профиль подготовки (специализация) 09.03.03.04 Прикладная информатика в  
государственном и муниципальном управлении

Форма обучения очная

Год набора 2021

## **РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)**

Программу составили

Доцент, к.ф.-м.н. \_\_\_\_\_ Буреева М. А.

## 1 Цели и задачи изучения дисциплины

### 1.1 Цель преподавания дисциплины:

В соответствии с требованиями ФГОС по направлению подготовки «Прикладная информатика» целью преподавания дисциплины «Технологии Big Data» является изучение современных высокопроизводительных распределенных систем, а также ознакомление с типовыми методами и алгоритмами параллельной обработки сверхбольших массивов данных с использованием стека технологий Big Data.

### 1.2 Задачи изучения дисциплины:

Студенты должны

знать: теоретические основы организации параллельной распределенной обработки данных на программном уровне;

уметь: реализовывать параллельные алгоритмы обработки данных на высокоуровневых языках программирования с использованием библиотек;

владеть: навыками работы со средствами выполнения и отладки прикладного ПО для распределенных систем.

1.3 Перечень планируемых результатов обучения по дисциплине (модулю), соотнесенных с планируемыми результатами освоения образовательной программы высшего образования:

| Код и наименование индикатора достижения компетенции  | Запланированные результаты обучения по дисциплине   |
|---|---|
| УК-1 Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач | Знать принципы сбора, отбора и обобщения информации.<br>Уметь анализировать и систематизировать разнородные данные.<br>Владеть навыками систематизации наблюдаемых данных.  |
| ПК-9 Способность осуществлять ведение базы данных и поддержку информационного обеспечения решения прикладных задач                  | Знать теоретические основы организации распределенных вычислений.<br>Уметь реализовывать параллельные алгоритмы обработки данных на высокоуровневых языках программирования с использованием библиотек.<br>Владеть навыками использования программных средств, поддерживающих технологии Big Data, для решения практических задач в предметной области. |

### 1.4 Особенности реализации дисциплины.

URL-адрес и название электронного обучающего курса

<https://e.sfu-kras.ru/course/view.php?id=34135>

Дисциплина реализуется с применением ЭО и ДОТ

## 2 Объем дисциплины (модуля)

| Вид учебной работы                          | Всего,<br>зачетных<br>единиц<br>(акад.час) | Семестр  |
|---|--|----------|
|   |  | 7        |
| <b>Общая трудоемкость дисциплины</b>        | 5 (180)                                    | 5 (180)  |
| <b>Контактная работа с преподавателем:</b>  | 1 (36)                                     | 1 (36)   |
| занятия лекционного типа                    | 0,5 (18)                                   | 0,5 (18) |
| лабораторные работы                         | 0,5 (18)                                   | 0,5 (18) |
| <b>Самостоятельная работа обучающихся</b>   | 4 (144)                                    | 4 (144)  |
| <b>Вид промежуточной аттестации (Зачет)</b> |  | Зачёт    |

### 3 Содержание дисциплины (модуля)

| № п/п   | Вид работ | Темы занятия   | Объем часов | Семестр /курс | Часы в эл. формате | РО        | Мероприятия текущего контроля и ПА |
|---|-----------|--|-------------|---------------|--------------------|-----------|------------------------------------|
| <b>Раздел 1. Введение в Big Data</b>                      |           |  |             |               |                    |           |                                    |
| 1.  | Лек       | Введение в Big Data  | 2           | 7             |                    | УК-1,ПК-9 |                                    |
| 2.  | Лек       | Виды распределенных файловых систем. Распределенная файловая система Hadoop. Принцип доставки вычислений к данным. Метод Map/Reduce.                           | 2           | 7             |                    | УК-1,ПК-9 |                                    |
| 3.  | Ср        | Изучение теоретического курса  | 18          | 7             | 2                  | УК-1,ПК-9 |                                    |
| <b>Раздел 2. Базы данных NoSQL</b>                        |           |  |             |               |                    |           |                                    |
| 1.  | Лек       | Нереляционные базы данных (NoSQL). Проблемы масштабирования реляционных СУБД. Отличия нереляционных хранилищ от реляционных СУБД, области применения, примеры  | 2           | 7             |                    | УК-1,ПК-9 |                                    |
| 2.  | Лек       | Хранилища NoSQL: документоориентированные хранилища, СУБД MongoDB  | 2           | 7             |                    | УК-1,ПК-9 |                                    |
| 3.  | Лаб       | Введение в СУБД MongoDB  | 2           | 7             | 2                  | УК-1,ПК-9 |                                    |
| 4.  | Лаб       | Работа с БД в СУБД MongoDB   | 4           | 7             | 4                  | УК-1,ПК-9 |                                    |
| 5.  | Ср        | Изучение теоретического курса  | 26          | 7             | 4                  | УК-1,ПК-9 |                                    |
| <b>Раздел 3. Технологии высокопроизводительных систем</b> |           |  |             |               |                    |           |                                    |
| 1.  | Лек       | Технологии высокопроизводительных вычислений MPI, OpenMP. История развития. Новые подходы ориентированные на данные. Понятие высокопроизводительных вычислений | 2           | 7             |                    | УК-1,ПК-9 |                                    |
| 2.  | Лаб       | Основы OpenMP  | 6           | 7             | 6                  | УК-1,ПК-9 |                                    |
| 3.  | Ср        | Изучение теоретического курса  | 34          | 7             | 4                  | УК-1,ПК-9 |                                    |
| <b>Раздел 4. Модели и методы анализа больших данных</b>   |           |  |             |               |                    |           |                                    |
| 1.  | Лек       | Основные задачи анализа больших данных в информационных системах. Классификация современных методов анализа данных.  | 2           | 7             |                    | УК-1,ПК-9 |                                    |
| 2.  | Лек       | Инструменты работы с большими данными. Решение задач классификации, кластеризации, прогнозирования на основе больших данных.                                   | 4           | 7             |                    | УК-1,ПК-9 |                                    |
| 3.  | Лаб       | Инструменты работы с большими данными: платформа Python, библиотеки работы с табличными данными - Numpy и Pandas   | 6           | 7             | 6                  | УК-1,ПК-9 |                                    |

|  |       |   |    |   |   |           |  |
|--|-------|---|----|---|---|-----------|--|
| 4.   | Ср    | Изучение теоретического курса   | 36 | 7 | 4 | УК-1,ПК-9 |  |
| <b>Раздел 5. Технологии Big Data. Перспективы развития</b> |       |   |    |   |   |           |  |
| 1.   | Лек   | Применение (достоинства и недостатки) технологии Big Data. Точки роста технологий Big Data. Конвергенция Big Data и классических технологий высокопроизводительных вычислений | 2  | 7 |   | УК-1,ПК-9 |  |
| 2.   | Ср    | Изучение теоретического курса   | 30 | 7 | 4 | УК-1,ПК-9 |  |
| 3.   | Зачёт |   |    | 7 |   | УК-1,ПК-9 |  |

#### **4 Учебно-методическое обеспечение дисциплины**

##### **4.1 Печатные и электронные издания:**

1. Дюк В., Самойленко А. Data Mining. Учебный курс:. - Санкт-Петербург: Питер, 2001. - 366 с..
2. Мартишин С.А., Симонов В.Л. Базы данных: Работа с распределенными базами данных и файловыми системами на примере MongoDB и HDFS с использованием Node.js, Express.js, Apache Spark и Scala [Электронный ресурс]: Учебное пособие. - Москва: ООО "Научно-издательский центр ИНФРА-М", 2020. - 235 с. – Режим доступа: <https://znanium.com/catalog/document?id=346035> .
3. Варфоломеева А. О., Коряковский А.В. Информационные системы предприятия [Электронный ресурс]: Учебное пособие. - Москва: ООО "Научно-издательский центр ИНФРА-М", 2019. - 330 с. – Режим доступа: <https://znanium.com/catalog/document?id=335060> .
4. Макшанов А. В., Журавлев А. Е. Технологии интеллектуального анализа данных [Электронный ресурс]: учебное пособие. - Санкт-Петербург: Лань, 2018. - 212 с. – Режим доступа: <https://e.lanbook.com/book/109617> .

##### **4.3 Интернет-ресурсы, включая профессиональные базы данных и информационные справочные системы:**

1. Научная библиотека СФУ
2. Электронный каталог АБИС-ИРБИС
3. Система электронного обучения СФУ
4. Электронно-библиотечная система издательства «Лань»
5. Электронно-библиотечная система ZNANIUM.COM (ИНФРА-М)
6. Электронный образовательный курс «Технологии Big Data»

#### **5 Фонд оценочных средств**

Фонд оценочных средств является приложением к рабочей программе дисциплины (модуля), хранится на кафедре, обеспечивающей преподавание данной дисциплины (модуля).

#### **6 Материально-техническая база, необходимая для осуществления образовательного процесса по дисциплине (модулю)**

Лекционные занятия проводятся в лекционных аудиториях, оснащенных проекционным оборудованием, компьютером, рабочими местами для преподавателя и студентов, магнитно-маркерной или меловой доской.

Лабораторные работы и самостоятельная работа студентов выполняются в компьютерных классах, объединенных в локальную сеть с выходом в Интернет. Компьютерные классы оборудованы рабочими местами на 12 компьютеров.