

Министерство науки и высшего образования РФ  
Федеральное государственное автономное образовательное учреждение  
высшего образования  
«СИБИРСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»  
**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)**

Б1.В.13 Технологии Big Data

---

наименование дисциплины (модуля) в соответствии с учебным планом

Направление подготовки / специальность

09.03.03 Прикладная информатика

---

Направленность (профиль)

09.03.03.04 Прикладная информатика в государственном и  
муниципальном управлении

---

Форма обучения

очная

---

Год набора

2020

---

## РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)

Программу составили \_\_\_\_\_

к.ф.-м.н., Доцент, Буреева М. А.

\_\_\_\_\_  
должность, инициалы, фамилия

## 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

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

### 1.4 Особенности реализации дисциплины

Язык реализации дисциплины: Русский.

Дисциплина (модуль) реализуется с применением ЭО и ДОТ

URL-адрес и название электронного обучающего курса: <https://e.sfu-kras.ru/course/view.php?id=34135>.

## 2. Объем дисциплины (модуля)

| Вид учебной работы                         | Всего,<br>зачетных<br>единиц<br>(акад. час) | е |
|--|---|---|
|  |   | 1 |
| <b>Контактная работа с преподавателем:</b> | <b>1 (36)</b>                               |   |
| занятия лекционного типа                   | 0,5 (18)                                    |   |
| лабораторные работы                        | 0,5 (18)                                    |   |
| <b>Самостоятельная работа обучающихся:</b> | <b>2 (72)</b>                               |   |
| курсовое проектирование (КП)               | Нет   |   |
| курсовая работа (КР)                       | Нет   |   |

### 3 Содержание дисциплины (модуля)

#### 3.1 Разделы дисциплины и виды занятий (тематический план занятий)

| №<br>п/п   |  | Модули, темы (разделы) дисциплины   |  | Контактная работа, ак. час. |                    |                                     |                    |                                      |                    |                                  |                    |
|--|--|---|--|-----------------------------|--------------------|-------------------------------------|--------------------|--------------------------------------|--------------------|----------------------------------|--------------------|
|  |  |   |  | Занятия лекционного типа    |                    | Занятия семинарского типа           |                    |                                      |                    | Самостоятельная работа, ак. час. |                    |
|  |  |   |  |                             |                    | Семинары и/или Практические занятия |                    | Лабораторные работы и/или Практикумы |                    |                                  |                    |
|  |  |   |  | Всего                       | В том числе в ЭИОС | Всего                               | В том числе в ЭИОС | Всего                                | В том числе в ЭИОС | Всего                            | В том числе в ЭИОС |
| <b>1. Технологии высокопроизводительных систем</b> |  |   |  |                             |                    |                                     |                    |                                      |                    |                                  |                    |
|  |  | 1. Технологии высокопроизводительных вычислений MPI, OpenMP. История развития. Новые подходы ориентированные на данные. Понятие высокопроизводительных вычислений |  | 2                           |                    |                                     |                    |                                      |                    |                                  |                    |
|  |  | 2. Основы OpenMP  |  |                             |                    |                                     |                    | 6                                    |                    |                                  |                    |
|  |  | 3. Изучение теоретического курса  |  |                             |                    |                                     |                    |                                      |                    | 18 6                             |                    |
| <b>2. Модели и методы анализа больших данных</b>   |  |   |  |                             |                    |                                     |                    |                                      |                    |                                  |                    |
|  |  | 1. Основные задачи анализа больших данных в информационных системах. Классификация современных методов анализа данных.  |  | 2                           |                    |                                     |                    |                                      |                    |                                  |                    |
|  |  | 2. Инструменты работы с большими данными. Решение задач классификации, кластеризации, прогнозирования на основе больших данных.                                   |  | 4                           |                    |                                     |                    |                                      |                    |                                  |                    |

|   |    |  |  |  |  |    |  |    |    |
|---|----|--|--|--|--|----|--|----|----|
| 3. Инструменты работы с большими данными:<br>платформа Python, библиотеки работы с табличными<br>данными - Numpy и Pandas   |    |  |  |  |  | 6  |  |    |    |
| 4. Изучение теоретического курса  |    |  |  |  |  |    |  | 18 | 6  |
| <b>3. Технологии Big Data. Перспективы развития</b>   |    |  |  |  |  |    |  |    |    |
| 1. Применение (достоинства и недостатки) технологии<br>Big Data. Точки роста технологий Big Data.<br>Конвергенция Big Data и классических технологий<br>высокопроизводительных вычислений | 2  |  |  |  |  |    |  |    |    |
| 2. Изучение теоретического курса  |    |  |  |  |  |    |  | 16 |    |
| <b>4. Введение в Big Data</b>   |    |  |  |  |  |    |  |    |    |
| 1. Введение в Big Data  | 2  |  |  |  |  |    |  |    |    |
| 2. Виды распределенных файловых систем.<br>Распределенная файловая система Hadoop. Принцип<br>доставки вычислений к данным. Метод Map/Reduce.   | 2  |  |  |  |  |    |  |    |    |
| 3. Изучение теоретического курса  |    |  |  |  |  |    |  | 8  | 2  |
| <b>5. Базы данных NoSQL</b>   |    |  |  |  |  |    |  |    |    |
| 1. Нереляционные базы данных (NoSQL). Проблемы<br>масштабирования реляционных СУБД. Отличия<br>нереляционных хранилищ от реляционных СУБД,<br>области применения, примеры                 | 2  |  |  |  |  |    |  |    |    |
| 2. Хранилища NoSQL: документоориентированные<br>хранилища, СУБД MongoDB   | 2  |  |  |  |  |    |  |    |    |
| 3. Введение в СУБД MongoDB  |    |  |  |  |  | 2  |  |    |    |
| 4. Работа с БД в СУБД MongoDB   |    |  |  |  |  | 4  |  |    |    |
| 5. Изучение теоретического курса  |    |  |  |  |  |    |  | 12 | 4  |
| Всего   | 18 |  |  |  |  | 18 |  | 72 | 18 |

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

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

1. Дюк В., Самойленко А. Data Mining. Учебный курс(Санкт-Петербург: Питер).
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: Учебное пособие(Москва: ООО "Научно-издательский центр ИНФРА-М").
3. Варфоломеева А. О., Коряковский А.В. Информационные системы предприятия: Учебное пособие(Москва: ООО "Научно-издательский центр ИНФРА-М").
4. Макшанов А. В., Журавлев А. Е. Технологии интеллектуального анализа данных: учебное пособие(Санкт-Петербург: Лань).

### **4.2 Лицензионное и свободно распространяемое программное обеспечение, в том числе отечественного производства (программное обеспечение, на которое университет имеет лицензию, а также свободно распространяемое программное обеспечение):**

1. Программное обеспечение Microsoft Office 2007 и выше;
2. Актуальные версии веб-браузеров Google Chrom, Mozilla Firefox и др.

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

1. <https://elibrary.ru> – Электронно-библиотечная система elibrary
2. <http://www.znanium.com/> – Электронно-библиотечная система ZNANIUM.COM (ИНФРА-М)
3. <http://www.sfu-kras.ru/> – Сайт ФГАОУ ВО «Сибирский федеральный университет»
4. <http://khti.sfu-kras.ru/> – Сайт Хакасского технического института – филиала ФГАОУ ВО «Сибирский федеральный университет»

## **5 Фонд оценочных средств**

Оценочные средства находятся в приложении к рабочим программам дисциплин.

## **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.