

Министерство цифрового развития, связи и массовых коммуникаций Российской Федерации
Федеральное государственное бюджетное образовательное учреждение высшего образования
«Сибирский государственный университет телекоммуникаций и информатики»
(СибГУТИ)
Уральский технический институт связи и информатики (филиал) в г. Екатеринбурге
(УрТИСИ СибГУТИ)

УТВЕРЖДАЮ
Директор УрТИСИ СибГУТИ
Минина Е.А.
« ____ » _____ 2023 г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Б1.В.01 Технологии обработки больших данных

Направление подготовки / специальность: **09.04.01 «Информатика и вычислительная техника»**

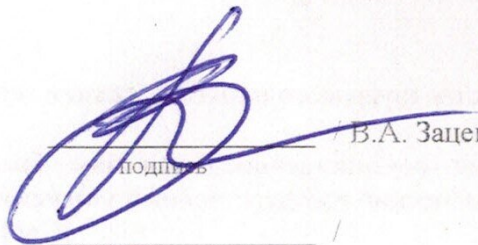
Направленность (профиль) /специализация: **Инженерия программного обеспечения и информационных систем**

Форма обучения: **очная, заочная**

Год набора: **2023**

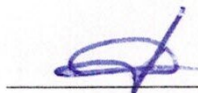
Екатеринбург, 2023

Разработчик (-и) рабочей программы:
к.п.н., доцент



_____ / В.А. Зацепин /
подпись

Утверждена на заседании кафедры информационных систем и технологий (ИСТ) протокол от
28.04.2023 г. № 10

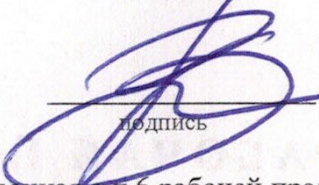
Заведующий кафедрой ИСТ


_____ / Д.И. Бурумбаев /
подпись

Согласовано:
Заведующий выпускающей кафедрой



_____ / Д.И. Бурумбаев /
подпись

Ответственный по ОПОП


_____ / В.А. Зацепин /
подпись

Основная и дополнительная литература, указанная в п.6 рабочей программы, имеется в наличии
в библиотеке института и ЭБС.

Заведующий библиотекой


_____ / С.Г. Торбенко /
подпись

1. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ

Дисциплина Б1.В.01 Технологии обработки больших данных относится к части, формируемой участниками образовательных отношений образовательной программы.

ПК-1 Способен к изучению, освоению и внедрению новых технологий работы с базами данных с обеспечением информационной безопасности	
Предшествующие дисциплины и практики	Б1.В.02 Документоориентированные базы данных Б1.В.ДВ.02 Дисциплины (модули) по выбору 2 (ДВ.2) Б1.В.ДВ.02.01 Перспективные технологии защиты информации Б1.В.ДВ.02.02 Анализ функционирования распределенных вычислительных систем
Дисциплины и практики, изучаемые одновременно с данной дисциплиной	Б3.01 Подготовка и сдача государственного экзамена Б3.02 Выполнение и защита выпускной квалификационной работы
Последующие дисциплины и практики	-

Дисциплина не может реализовываться с применением дистанционных образовательных технологий.

2. ТРЕБОВАНИЯ К РЕЗУЛЬТАТАМ ОСВОЕНИЯ ДИСЦИПЛИНЫ

В результате освоения дисциплины обучающийся должен демонстрировать результаты обучения, которые соотнесены с индикаторами достижения компетенций:

Код и наименование индикатора достижения компетенций	Результаты обучения по дисциплине
ПК-1 Способен к изучению, освоению и внедрению новых технологий работы с базами данных с обеспечением информационной безопасности	
ПК-1.1. Знает архитектуру систем баз данных, основные модели данных, этапы и методы проектирования Проектирование и администрирование баз данных	Знает основные технологии обработки баз данных; этапы жизненного цикла баз данных; виды и способы организации запросов к данным в реляционных моделях
ПК-1.2. Умеет применять языки программирования и инструментальные средства в профессиональной деятельности, обосновывать выбор необходимых инструментальных средства для создания и функционирования баз данных на предприятии	Умеет реализовывать этапы жизненного цикла баз данных; использовать технологии баз данных для обработки табличных данных; представлять отчеты по результатам обработки данных средствами СУБД
ПК-1.3. Владеет методами и средствами построения баз данных, демонстрировать способность и готовность к эксплуатации и администрированию баз данных с учетом требований по обеспечению информационной безопасности.	Владеет использованием основных моделей информационных технологий и способов их применения для решения задач в предметных областях; использованием классических приемов (алгоритмов) при решении корректных постановок классических задач.

3. ОБЪЁМ ДИСЦИПЛИНЫ

Общая трудоемкость дисциплины составляет 4 зачетных единиц.

Дисциплина изучается:

по очной форме обучения – в 4 семестре.

по заочной форме обучения – на 2,3 курсе.

Форма промежуточной аттестации по дисциплине – экзамен

По дисциплине предусмотрена домашняя контрольная работа

3.1 Очная форма обучения (О)

Виды учебной работы	Всего часов	Семестр
		4
Аудиторная работа (всего)	60	60
Лекции (ЛК)	24	24
Лабораторные работы (ЛР)	-	-
Практические занятия (ПЗ)	34	34
В том числе в интерактивной форме	-	-
В том числе в форме практической подготовки	-	-
Предэкзаменационные консультации (ПК)	2	2
Самостоятельная работа (всего)	50	50
Работа над конспектами лекций	25	25
Подготовка к практическим занятиям	25	25
Подготовка к лабораторным работам	-	-
Выполнение курсового проекта	-	-
Выполнение курсовой работы	-	-
Выполнение РГР	-	-
Выполнение реферата	-	-
Контроль (всего)	34	34
Подготовка к сдаче экзамена	30	30
Сдача экзамена	4	4
Подготовка к сдаче зачета	-	-
Сдача зачета	-	-
Общая трудоемкость дисциплины	144	144

3.2 Заочная форма обучения (З)

Виды учебной работы	Всего часов	Курс	
		2	3
Аудиторная работа (всего)	20	4	16
Лекции (ЛК)	10	4	6
Лабораторные работы (ЛР)	-	-	-
Практические занятия (ПЗ)	8	-	8
В том числе в интерактивной форме	-	-	-
В том числе в форме практической подготовки	-	-	-
Предэкзаменационные консультации (ПК)	2	-	2
Самостоятельная работа (всего)	115	56	59
Работа над конспектами лекций	56	28	28
Подготовка к практическим занятиям	56	28	28
Подготовка к лабораторным работам	-	-	-
Выполнение курсового проекта	-	-	-
Выполнение курсовой работы	-	-	-
Выполнение РГР	-	-	-
Выполнение реферата	-	-	-
Выполнение домашней контрольной работы	3	-	3
Контроль (всего)	9	-	9
Подготовка к сдаче экзамена	6	-	6
Сдача экзамена	3	-	3
Подготовка к сдаче зачета	-	-	-
Сдача зачета	-	-	-
Общая трудоемкость дисциплины	144	60	84

4. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ ПО ВИДАМ УЧЕБНЫХ ЗАНЯТИЙ

4.1 Содержание лекционных занятий

№ раздела дисциплины	Наименование лекционных тем (разделов) дисциплины и их содержание	Объем в часах	
		О	З
1	Введение в обработку больших данных Определение больших данных и их роль в современном информационном мире. Основные вызовы и возможности обработки больших данных	3	1
2	Архитектура и экосистемы обработки больших данных Обзор распределенных архитектур и фреймворков для обработки больших данных, включая Hadoop и Spark. Роль хранилищ данных и кластеров в архитектуре	2	1
3	Хранение и управление большими данными Методы и технологии хранения и управления данными большого объема. Системы управления базами данных (СУБД) и их роль в обработке больших данных	2	1
4	Обработка данных в реальном времени Основы потоковой обработки данных и системы, такие как Apache Kafka. Анализ и применения потоковых данных	3	1
5	Методы обработки и анализа больших данных Основы обработки и анализа данных, включая SQL, NoSQL, и машинное обучение. Поиск и извлечение информации из больших данных	2	1
6	Безопасность и конфиденциальность данных в больших данных Вопросы безопасности и защиты данных в средах обработки больших данных. Методы шифрования и аутентификации	3	1
7	Интеграция и визуализация данных Методы интеграции различных источников данных в хранилища данных. Визуализация данных для лучшего понимания и принятия решений	2	1
8	Оптимизация производительности и масштабируемость Методы оптимизации запросов и обработки данных для повышения производительности. Масштабируемость систем обработки больших данных	2	1
9	Практические применения обработки больших данных Примеры применения технологий обработки больших данных в различных отраслях, включая финансы, медицину и маркетинг. Кейсы успешных проектов	3	1
10	Этические и правовые вопросы в обработке больших данных Рассмотрение этических и правовых аспектов сбора и использования больших данных. Соблюдение	2	1

	нормативных требований и регуляций в сфере обработки данных		
		ВСЕГО	24
			10

4.2 Содержание практических занятий

№ п/п	№ раздела дисциплины	Наименование практических занятий	Объем в часах	
			О	З
1	1	Введение в экосистему Hadoop	4	1
2	2	Обработка данных с использованием Apache MapReduce	3	1
3	3	Обработка потоков данных с Apache Kafka и Apache Flink	4	1
4	4	Использование NoSQL-баз данных для анализа больших данных	3	1
5	5	Машинное обучение на больших данных	3	1
6	6	Оптимизация запросов и производительности в больших данных	4	1
7	7	Интеграция и визуализация данных больших данных	3	1
8	8	Архитектура больших данных и масштабируемость	3	1
9	9	Безопасность и конфиденциальность данных в больших данных	4	-
10	10	Проекты и практические задания	3	-
			ВСЕГО	34
				8

5. ПЕРЕЧЕНЬ ИННОВАЦИОННЫХ ФОРМ УЧЕБНЫХ ЗАНЯТИЙ

№ п/п	Тема	Объем в часах		Вид учебных занятий	Используемые инновационные формы занятий
		О	З		
1	Безопасность и конфиденциальность данных в больших данных	-	-	лекция	дискуссия
2	Практические применения обработки больших данных	-	-	лекция	дискуссия
3	Машинное обучение на больших данных	-	-	Практическая работа	Моделирование рабочей ситуации
4	Оптимизация запросов и производительности в больших данных	-	-	Практическая работа	Моделирование рабочей ситуации
5	Интеграция и визуализация данных больших данных	-	-	Практическая работа	Моделирование рабочей ситуации
ВСЕГО		-	-		

6. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ПО ДИСЦИПЛИНЕ

6.1 Список основной литературы

6.1.1 Барбара Инь и Ванесса Ченг Технологии Big Data: наука о данных, машинное обучение и интеллектуальный анализ данных – Москва: Издательский дом «Питер», 2014 г. – 531 с.

6.1.2 Алексей Колосков Анализ данных с использованием Microsoft Excel – Москва: Издательский дом «Вильямс», 2015 г. – 462 с.

6.1.3 Матей Залевски Spark: высокопроизводительная обработка и анализ данных – Москва: Издательский дом «ДМК Пресс», 2016 г. – 139 с.

6.2 Список дополнительной литературы

6.2.1 Мэттью Л. Хатчисон, Бенджамин Каннинг, Оливер Джексон Введение в Hadoop – Москва: Издательский дом O'Reilly Media, 2021 г. – 590 с.

6.2.2 Роберт Кабанис R: анализ и визуализация данных – Санкт-Петербург: Издательский дом «Диалектика», 2020 г. – 200 с.

6.2.3 Wes McKinney, O'Reilly Python для анализа данных – Москва, 2021 г. – 439 с.

6.3 Информационное обеспечение (в т.ч. интернет-ресурсы).

1. Единая электронная образовательная среда института: URL:<http://aup.uisi.ru>

2. Научная электронная библиотека eLibrary. [Электронный ресурс] – Режим доступа: <http://www.elibrary.ru>.

3. Электронно-библиотечная система «IPRbooks» —(<http://www.iprbookshop.ru/>, доступ по паролю)

4. Полнотекстовая база данных УМП СибГУТИ — Режим доступа: (http://ellib.sibsutis.ru/cgi-bin/irbis64r_plus/cgiirbis_64_ft.exe?Z21ID=GUEST&C21COM=F&I21DBN=AUTHOR&P21DBN=ELLIB&Z21FLAGID=1, доступ по логину- паролю)

5. Полнотекстовая база данных ПГУТИ — Режим доступа: (http://ellib.sibsutis.ru/cgi-bin/irbis64r_plus/cgiirbis_64_ft.exe?Z21ID=GUEST&C21COM=F&I21DBN=AUTHOR&P21DBN=PGUTI&Z21FLAGID=1, доступ по паролю)

6. Архивы иностранных научных журналов на платформе НЭИКОН — Режим доступа: (<http://arch.neicon.ru/>, свободный доступ с ПК вуза – доступ по IP-адресу)

6.4 Нормативные правовые документы и иная правовая информация

Нормативные правовые акты и нормативные методические документы, иная правовая информация (при наличии).

7 МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ И ТРЕБУЕМОЕ ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ

Наименование аудиторий, кабинетов, лабораторий	Вид занятий	Оборудование, программное обеспечение
Учебная аудитория для проведения занятий лекционного типа	лекционные занятия	<p>Оснащение: 55 посадочных мест, офисная мебель, доска магнитно-маркерная, персональный компьютер PowerColor, монитор АОС, проектор Viewsonic, экран настенный.</p> <p>Программное обеспечение: операционная система Windows 10, Microsoft Office, PDF24, Google Chrome, Microsoft Visual Studio, Kaspersky Endpoint Security, Microsoft SQL Server 2019, Pascal ABC.NET, Python 3.10.7, Pip for Python, PyCharm Community Edition 2022.2.1, Foxit PDF Reader.</p>
Учебная аудитория для проведения практических занятий.	практические занятия	<p>Оснащение: 23 посадочных мест, офисная мебель, доска магнитно-маркерная, персональные компьютеры Athlon, мониторы АОС.</p> <p>Программное обеспечение: операционная система Windows 10, Microsoft Office, PDF24, Google Chrome, Microsoft Visual Studio, Microsoft Visual Studio Code, IntelliJ IDEA Community Edition 2022.1.3, Maple 12, Kaspersky Endpoint Security, Mathcad 2001 Professional, Microsoft SQL Server 2019, Pascal ABC.NET, Python 3.10.8, Pip for Python, Sublime Text 3, PyCharm Community Edition 2022.1.3, VLC Media Player, Foxit PDF Reader.</p>
Учебная аудитория для проведения групповых, индивидуальных консультаций, текущего контроля и промежуточной аттестации	Групповые и индивидуальные консультации текущий контроль, промежуточная аттестация	<p>Оснащение: 23 посадочных мест, офисная мебель, доска магнитно-маркерная, персональные компьютеры Athlon, мониторы АОС.</p> <p>Программное обеспечение: операционная система Windows 10, Microsoft Office, PDF24, Google Chrome, Microsoft Visual Studio, Microsoft Visual Studio Code, IntelliJ IDEA Community Edition 2022.1.3, Maple 12, Kaspersky Endpoint Security, Mathcad 2001 Professional, Microsoft SQL Server 2019, Pascal ABC.NET, Python 3.10.8, Pip for Python, Sublime Text 3, PyCharm</p>

		Community Edition 2022.1.3, VLC Media Player, Foxit PDF Reader.
Помещение для самостоятельной работы	самостоятельная работа	<p>Оснащение: 23 посадочных мест, офисная мебель, доска магнито-маркерная, проектор Sanyo, экран настенный, персональные компьютеры Intel Core 2 Duo, мониторы Samsung.</p> <p>Программное обеспечение: операционная система Windows 7, Microsoft Office, PDF24, Google Chrome, Microsoft Visual Studio Code, Kaspersky Endpoint Security, Python 3.8.10, Pip for Python, VLC Media Player, Foxit PDF Reader.</p>

8. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ ПО ОСВОЕНИЮ ДИСЦИПЛИН

8.1 Подготовка к лекционным, практическим занятиям

8.1.1 Подготовка к лекциям

На лекциях необходимо вести конспектирование учебного материала, обращать внимание на категории, формулировки, раскрывающие содержание научных явлений и процессов, научные выводы и практические рекомендации.

Конспектирование лекций – сложный вид аудиторной работы, предполагающий интенсивную умственную деятельность студента. Целесообразно сначала понять основную мысль, излагаемую лектором, а затем записать ее. Желательно оставлять поля, на которых при самостоятельной работе с конспектом можно сделать дополнительные записи и отметить непонятные вопросы.

Конспект лекции лучше подразделять на пункты в соответствии с вопросами плана лекции, предложенными преподавателем. Следует обращать внимание на акценты, выводы, которые делает лектор, отмечая наиболее важные моменты в лекционном материале.

Во время лекции можно задавать преподавателю уточняющие вопросы с целью освоения теоретических положений, разрешения спорных вопросов.

8.1.3 Подготовка к практическим занятиям

Подготовку к практическим занятиям следует начинать с ознакомления плана практического занятия, который отражает содержание предложенной темы. Изучение вопросов плана основывается на проработке текущего материала лекции, а затем изучении основной и дополнительной литературы. Новые понятия по изучаемой теме необходимо выучить и внести в глоссарий, который целесообразно вести с самого начала изучения курса.

Результат такой работы должен проявиться в способности студента свободно ответить на теоретические вопросы практикума,

8.2 Самостоятельная работа студентов

Успешное освоение компетенций, формируемых данной учебной дисциплиной, предполагает оптимальное использование времени самостоятельной работы.

Целесообразно дорабатывать свой конспект лекции, делая в нем соответствующие записи из литературы, рекомендованной преподавателем и предусмотренной учебной программой.

Все задания к практическим занятиям, а также задания, вынесенные на самостоятельную работу, рекомендуется выполнять непосредственно после соответствующей темы лекционного курса, что способствует лучшему усвоению материала, позволяет своевременно выявить и устранить «пробелы» в знаниях, систематизировать ранее пройденный материал, на его основе приступить к получению новых знаний и овладению навыками.

Самостоятельная работа во внеаудиторное время состоит из:

- повторение лекционного материала;
- подготовки к практическим занятиям работам;
- изучения учебно-методической и научной литературы;
- изучения нормативно-правовых актов;
- решения задач, выданных на практических занятиях;
- подготовки к контрольным работам, тестированию и т. д.;
- выполнения домашней контрольной работы;
- проведение самоконтроля путем ответов на вопросы текущего контроля знаний, решения представленных в учебно-методических материалах дисциплины задач, тестов.

8.3 Подготовка к промежуточной аттестации

При подготовке к промежуточной аттестации необходимо:

- внимательно изучить перечень вопросов и определить, в каких источниках находятся сведения, необходимые для ответа на них;
- внимательно прочитать рекомендуемую литературу;
- составить краткие конспекты ответов (планы ответов).

Освоение дисциплины предусматривает посещение лекционных занятий, выполнение и защиту практических работ, самостоятельной работы.

Текущий контроль достижения результатов обучения по дисциплине включает следующие процедуры:

- решение индивидуальных задач на практических занятиях;
- контроль самостоятельной работы, осуществляемый на каждом практическом занятии;

Промежуточный контроль достижения результатов обучения по дисциплине проводится в следующих формах:

- экзамен

Для проведения текущего контроля и промежуточной аттестации используются оценочные средства, описание которых представлено в Приложении 1 и на сайте (<http://www.aup.uisi.ru>).

9. ОСОБЕННОСТИ РЕАЛИЗАЦИИ ДИСЦИПЛИНЫ ДЛЯ ЛИЦ С ОГРАНИЧЕННЫМИ ВОЗМОЖНОСТЯМИ ЗДОРОВЬЯ

Для реализации дисциплины используются материально-технические условия, программное обеспечение и доступная среда, созданные в институте. Учебные материалы предоставляются обучающимся в доступной форме (в т.ч. в ЭИОС) с применением программного обеспечения:

Балаболка — программа, которая предназначена для воспроизведения вслух текстовых файлов самых разнообразных форматов, среди них: DOC, DOCX, DjVu, FB2, PDF и многие другие. Программа Балаболка умеет воспроизводить текст, набираемый на клавиатуре, осуществляет проверку орфографии;

Экранная лупа – программа экранного увеличения.

Для контактной и самостоятельной работы используются мультимедийные комплексы, электронные учебники и учебные пособия, адаптированные к ограничениям здоровья обучающихся имеющиеся в электронно-библиотечных системах «IPR SMART//IPRbooks», «Образовательная платформа Юрайт».

Промежуточная аттестация и текущий контроль по дисциплине осуществляется в соответствии с фондом оценочных средств в формах, адаптированных к ограничениям здоровья и восприятия информации обучающихся.

Задания предоставляется в доступной форме:

для лиц с нарушениями зрения: в устной форме или в форме электронного документа с использованием специализированного программного обеспечения;

для лиц с нарушениями слуха: в печатной форме или в форме электронного документа;

для лиц с нарушениями опорно-двигательного аппарата: в устной форме или в печатной форме, или в форме электронного документа.

Ответы на вопросы и выполненные задания обучающиеся предоставляют в доступной форме:

для лиц с нарушениями зрения: в устной форме или в письменной форме с помощью ассистента, в форме электронного документа с использованием специализированного программного обеспечения;

для лиц с нарушениями слуха: в электронном виде или в письменной форме;

для лиц с нарушениями опорно-двигательного аппарата: в устной форме или письменной форме, или в форме электронного документа (возможно с помощью ассистента).

При проведении текущего контроля и промежуточной аттестации обучающимся из числа инвалидов и лиц с ограниченными возможностями здоровья предоставляется дополнительное время для подготовки и ответа (по их заявлению).

Для инвалидов и лиц с ограниченными возможностями здоровья учебные занятия по дисциплине проводятся в ДОТ и/или в специально оборудованной аудитории (по их заявлению).