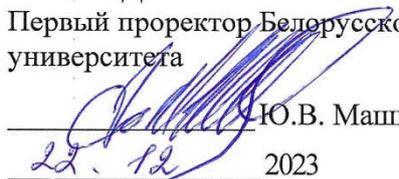


Межгосударственное образовательное учреждение высшего образования
«Белорусско-Российский университет»

УТВЕРЖДАЮ

Первый проректор Белорусско-Российского
университета


Ю.В. Машин

22.12 2023

Регистрационный № УД-090301/Б.Р.В.15/р

РАСПРЕДЕЛЕННЫЕ СИСТЕМЫ ОБРАБОТКИ ИНФОРМАЦИИ

(наименование дисциплины)

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Направление подготовки 09.03.01 Информатика и вычислительная техника

Направленность (профиль) Автоматизированные системы обработки информации и управления

Квалификация Бакалавр

	Форма обучения
	Очная
Курс	4
Семестр	7
Лекции, часы	30
Лабораторные занятия, часы	30
Зачет, семестр	7
Контактная работа по учебным занятиям, часы	60
Самостоятельная работа, часы	48
Всего часов / зачетных единиц	108/3

Кафедра-разработчик программы: Программное обеспечение информационных технологий

(название кафедры)

Составитель: К.В. Захарченков, к.т.н.

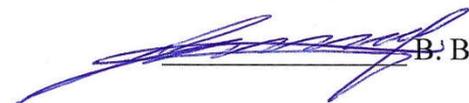
(И.О. Фамилия, ученая степень, ученое звание)

Могилев, 2023

Рабочая программа составлена в соответствии с федеральным государственным образовательным стандартом высшего профессионального образования по направлению подготовки 09.03.01 – “Информатика и вычислительная техника” (уровень бакалавриата), утвержденным приказом № 929 от 19.09.2017 и учебным планом, утвержденным Рег. № 090301-2.1 от 28.04.2023.

Рассмотрена и рекомендована к утверждению кафедрой «Программное обеспечение информационных технологий» «06» декабря 2023 г., протокол № 5.

Зав. кафедрой ПОИТ


В. В. Кутузов

Одобрена и рекомендована к утверждению Научно-методическим советом Белорусско-Российского университета

« 20 » декабря 2023 г., протокол № 3.

Зам. председателя
Научно-методического совета


С.А. Сухоцкий

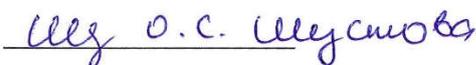
Рецензент:

И. В. Акиншева, заведующая кафедрой программного обеспечения информационных технологий МГУ им. А.А. Кулешова, канд. техн. наук, доцент

(И.О. Фамилия, должность, ученая степень, ученое звание рецензента)

Рабочая программа согласована:

Ведущий библиотекарь



Начальник учебно-методического
отдела


О.Е. Печковская

1 ПОЯСНИТЕЛЬНАЯ ЗАПИСКА

1.1 Цель учебной дисциплины

Основной целью преподавания дисциплины является получение теоретических знаний о методах сбора, обработки и анализа данных в распределенных системах, а также приобретение практических навыков автоматизированной обработки больших объемов данных в сложных программных системах.

Дисциплина «Распределенные системы обработки информации» обеспечивает приобретение знаний и умений в соответствии с государственным образовательным стандартом, содействует формированию мировоззрения и системного мышления.

1.2 Планируемые результаты изучения дисциплины

В результате освоения учебной дисциплины студент должен **знать:**

- процессы сбора, обработки и анализа больших объемов данных, основы машинного обучения, методы распределенного хранения и обработки больших объемов данных с использованием современных информационных технологий, методы глубокого анализа текста, методы визуализации данных.

уметь:

- решать задачи хранения, сбора, обработки и анализа больших объемов данных в распределенных системах.

владеть:

- навыками использования современных программных средств для хранения, сбора, обработки и анализа больших объемов данных в распределенных системах.

1.3 Место учебной дисциплины в системе подготовки студента

Дисциплина относится к блоку 1 «Часть Блока 1, формируемая участниками образовательных отношений»

Перечень учебных дисциплин, изучаемых ранее, усвоение которых необходимо для изучения данной дисциплины:

- программирование;
- базы данных;
- технологии разработки программных комплексов АСОИ.

Перечень учебных дисциплин (циклов дисциплин), которые будут опираться на данную дисциплину:

- технологии промышленного программирования (8 семестр);
- инструментальные средства промышленного программирования (8 семестр);
- технологии проектирования АСОИиУ.

Кроме того, знания, полученные при изучении дисциплины на лабораторных занятиях, будут использованы при прохождении преддипломной практики, а также при подготовке выпускной квалификационной работы.

1.4 Требования к освоению учебной дисциплины

Освоение данной учебной дисциплины должно обеспечивать формирование следующих компетенций:

Коды формируемых компетенций	Наименования формируемых компетенций
ПК-12	Способен обеспечивать информационную безопасность уровня баз данных
ПК-15	Способен организовать выполнение научно-исследовательских и опытно-конструкторских работ по закреплённой тематике

2 СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

Вклад дисциплины в формирование результатов обучения выпускника (компетенций) и достижение обобщенных результатов обучения происходит путём освоения содержания обучения и достижения частных результатов обучения, описанных в данном разделе.

2.1 Содержание учебной дисциплины

Номер тем	Наименование тем	Содержание	Коды формируемых компетенций
1	Введение.	Введение Область применения больших данных и их преимущества Структурированные и неструктурированные данные Данные на естественном языке Машинные данные Графовые и сетевые данные	ПК-12 ПК-15
2	Процесс анализа больших объемов данных	Сбор данных Подготовка данных Исследование данных Моделирование данных Визуальное представление результатов обработки данных	ПК-12 ПК-15
3	Программное обеспечение для автоматизации обработки больших объемов данных	Распределенные файловые системы Средства распределенного программирования приложений Программные средства интеграции данных Программные средства машинного обучения Программные средства сравнительного анализа данных Развертывание распределенной системы обработки данных	ПК-12 ПК-15
4	Машинное обучение	Понятие машинного обучения Типы машинного обучения Процесс моделирования при машинном обучении Обучение модели Проверка адекватности модели Прогнозирование новых наблюдений	ПК-15
5	Общие методы обработки больших объемов данных	Проектирование алгоритмов обработки больших объемов данных Проектирование архитектуры распределенной системы обработки больших объемов данных	ПК-15
6	Средства распределенного хранения и обработки больших данных	Проблемы создания распределенных баз данных Hadoop Spark	ПК-15
7	Глубокий анализ текста	Применение глубокого анализа текста Методы глубокого анализа текста Концепция набора слов Выделение основы и лемматизация Классификатор на базе дерева принятия решений	ПК-15

2.2 Учебно-методическая карта учебной дисциплины

№ недели	Лекции (наименование тем)	Часы	Лабораторные занятия	Часы	Самостоятельная работа, часы	Форма контроля знаний	Баллы (max)
Модуль 1							
1	Тема 1. Введение	2	Лр №1. Использование Hadoop для анализа больших объемов данных	2	4		
2	Тема 2. Процесс анализа больших объемов данных	2	Лр №1. Использование Hadoop для анализа больших объемов данных	2	3		
3	Тема 2. Процесс анализа больших объемов данных	2	Лр №1. Использование Hadoop для анализа больших объемов данных	2	3	ЗЛР	10
4	Тема 3. Программное обеспечение для автоматизации обработки больших объемов данных	2	Лр №2. Сбор данных для анализа	2	3		
5	Тема 4. Машинное обучение	2	Лр №2. Сбор данных для анализа	2	3		
6	Тема 4. Машинное обучение	2	Лр №2. Сбор данных для анализа	2	4	ЗЛР	15
7	Тема 4. Машинное обучение	2	Лр №3. Построение модели	2	3	КР	5
8	Тема 4. Машинное обучение	2	Лр №3. Построение модели	2	3	ПКУ	30
Модуль 2							
9	Тема 5. Общие методы обработки больших объемов данных	2	Лр №3. Построение модели	2	3	ЗЛР	8
10	Тема 5. Общие методы обработки больших объемов данных	2	Лр №4. Исследование модели	2	3		
11	Тема 6. Средства распределенного хранения и обработки больших данных	2	Лр №4. Исследование модели	2	3		
12	Тема 6. Средства распределенного хранения и обработки больших данных	2	Лр №4. Исследование модели	2	3	ЗЛР	8
13	Тема 6. Средства распределенного хранения и обработки больших данных	2	Лр №5. Анализ и визуальное представление результатов моделирования	2	3		
14	Тема 7. Глубокий анализ текста	2	Лр №5. Анализ и визуальное представление результатов моделирования	2	3	КР	5
15	Тема 7. Глубокий анализ текста	2	Лр №5. Анализ и визуальное представление результатов моделирования	2	4	ЗЛР ПКУ ПА (зачет)	9 30 40
	Итого	30		30	48		100

Принятые обозначения:
Текущий контроль –

ЗЛР – защита лабораторной работы;
 КР – контрольная работа (тест);
 ПКУ – промежуточный контроль успеваемости.
 ПА - Промежуточная аттестация.

Итоговая оценка определяется как сумма текущего контроля и промежуточной аттестации и соответствует баллам:

Зачет

Оценка	Зачтено	Незачтено
Баллы	51–100	0–50

3 ОБРАЗОВАТЕЛЬНЫЕ ТЕХНОЛОГИИ

При изучении дисциплины используется модульно-рейтинговая система оценки знаний студентов. Применение форм и методов проведения занятий при изучении различных тем курса представлено в таблице.

№ п/п	Форма проведения занятия	Вид аудиторных занятий		Всего часов
		Лекции	Лабораторные занятия	
1	Мультимедиа	Лекции №1-7		30
2	С использованием ЭВМ		Лр№1-Лр№5	30
	ИТОГО			60

4 ОЦЕНОЧНЫЕ СРЕДСТВА

Используемые оценочные средства по учебной дисциплине представлены в таблице и хранятся на кафедре.

№ п/п	Вид оценочных средств	Количество комплектов
1	Вопросы к зачету	1
2	Задания для защиты лабораторных работ	5
3	Контрольная работа (тест)	2

5 МЕТОДИКА И КРИТЕРИИ ОЦЕНКИ КОМПЕТЕНЦИЙ СТУДЕНТОВ

5.1 Уровни сформированности компетенций

№ п/п	Уровни сформированности компетенции	Содержательное описание уровня	Результаты обучения
ПК-12: Способен обеспечивать информационную безопасность уровня баз данных			
ИПК-12.1: Применяет политику обеспечения информационной безопасности уровня баз данных			
1	Пороговый уровень	Знание политики обеспечения информационной безопасности уровня баз данных	Знание способов обеспечения информационной безопасности в распределенных базах данных

2	Продвинутый уровень	Применение политики обеспечения информационной безопасности уровня баз данных	Применение методов обеспечения информационной безопасности в распределенных базах данных
3	Высокий уровень	Синтез, оценка способов защиты информации в базах данных	Синтез способов защиты информации в распределенных базах данных
ПК-15: Способен организовать выполнение научно-исследовательских и опытно-конструкторских работ по закреплённой тематике			
ИПК-15.1: Способен применять современные методы и средства распределённой обработки информации			
1	Пороговый уровень	Знание современных методов сбора, анализа и обработки данных в распределённых системах	Знание и понимание процессов сбора, анализа и обработки данных в распределённых системах
2	Продвинутый уровень	Применение методов и средств обработки информации в распределённых системах при выполнении научно-исследовательских работ	Анализ данных при проведении научно-исследовательских работ с использованием современных методов и средств обработки информации в распределённых системах
3	Высокий уровень	Синтез и оценка способов обработки информации в распределённых системах при выполнении научно-исследовательских работ	Синтез способов анализа данных при проведении научно-исследовательских работ с использованием современных методов и средств обработки информации в распределённых системах

5.2 Методика оценки знаний, умений и навыков студентов

Результаты обучения	Оценочные средства
ПК-12: Способен обеспечивать информационную безопасность уровня баз данных	
Знание способов обеспечения информационной безопасности в распределённых базах данных	Задание для защиты лабораторных работ №№ 1-2. Вопросы к зачету. Контрольная работа (тест)
Применение методов обеспечения информационной безопасности в распределённых базах данных	Задание для защиты лабораторных работ №№ 1-2. Вопросы к зачету. Контрольная работа (тест)
Синтез способов защиты информации в распределённых базах данных	Задание для защиты лабораторных работ №№ 1-2. Вопросы к зачету. Контрольная работа (тест)
ПК-15: Способен организовать выполнение научно-исследовательских и опытно-конструкторских работ по закреплённой тематике	
Знание и понимание процессов сбора, анализа и обработки данных в распределённых системах	Задание для защиты лабораторных работ №№ 1-5. Вопросы к зачету. Контрольная работа (тест)
Анализ данных при проведении научно-исследовательских работ с использованием современных методов и средств обработки информации в распределённых системах	Задание для защиты лабораторных работ №№ 1-5. Вопросы к зачету. Контрольная работа (тест)
Синтез способов анализа данных при проведении научно-исследовательских работ с использованием современных методов и средств обработки информации в распределённых системах	Задание для защиты лабораторных работ №№ 1-5. Вопросы к зачету. Контрольная работа (тест)

5.3 Критерии оценки лабораторных работ

Выполненная и защищенная лабораторная работа оценивается в 8, 9 и 10, 15 баллов максимум. При этом от 3 до 5 баллов начисляется за выполнение работы и от 2 до 10 баллов за оформление отчета и защиту работы в зависимости от качества оформления и уровня знаний студента по тематике работы. Если по окончании модуля лабораторная работа выполнена, но не защищена, то баллы по ней не начисляются и она попадает в разряд задолженности.

5.4 Критерии оценки контрольной работы (теста)

За семестр выполняются две контрольные работы (теста). Контрольная работа представляет собой случайную выборку из 20 вопросов. Каждый правильный ответ на вопрос оценивается в 0,25 балла. Итоговая оценка получается простым суммированием.

5.5 Критерии оценки зачета

Контрольное задание включает 2 теоретических вопроса. Теоретические вопросы выбираются из разных дидактических единиц. Каждый вопрос оценивается положительной оценкой в диапазоне от 10 до 20 баллов. Ответы на вопросы оцениваются по следующим критериям.

Теоретические вопросы:

- ◆ **19-20 баллов** – студент глубоко понимает пройденный материал, отвечает четко и всесторонне, умеет оценивать факты, использует научную терминологию, самостоятельно рассуждает, отличается способностью обосновать выводы и разъяснять их в логической последовательности, дает развернутый ответ на поставленный вопрос и четко отвечает на дополнительные вопросы.
- ◆ **17-18 баллов** – студент глубоко понимает пройденный материал, отвечает четко и всесторонне, умеет оценивать факты, самостоятельно рассуждает, отличается способностью обосновать выводы и разъяснять их в логической последовательности, но допускает отдельные неточности, в том числе и на дополнительные вопросы.
- ◆ **15-16 баллов** – студент хорошо понимает пройденный материал, отвечает правильно, умеет оценивать факты, самостоятельно рассуждает, обосновывает выводы и разъясняет их, но допускает ошибки общего характера.
- ◆ **13-14 баллов** – студент понимает пройденный материал, но не может теоретически обосновать некоторые выводы, допускает ошибки общего характера.
- ◆ **11-12 баллов** – студент отвечает в основном правильно на поставленный вопрос, но чувствуется механическое заучивание материала, отсутствует логическая последовательность при изложении ответа, не может ответить на дополнительные вопросы.
- ◆ **10 баллов** – в ответе студента имеются существенные недостатки, материал охвачен «половинчато», в рассуждениях допускаются ошибки
- ◆ **Ниже 10 баллов** – студент имеет общее представление о вопросе, ответ студента правилен лишь частично, при разъяснении материала допускаются серьезные ошибки, отсутствует техническая терминология, не может исправить ошибки с помощью наводящих вопросов;

6 МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ ПО ОРГАНИЗАЦИИ И ВЫПОЛНЕНИЮ САМОСТОЯТЕЛЬНОЙ РАБОТЫ СТУДЕНТОВ ПО УЧЕБНОЙ ДИСЦИПЛИНЕ

Самостоятельная работа студентов (СРС) направлена на закрепление и углубление освоения учебного материала, развитие практических умений. СРС включает следующие виды самостоятельной работы студентов:

- проработка тем (вопросов), вынесенных на самостоятельное изучение;

- конспектирование учебной литературы;
- подготовка докладов.

Для СРС рекомендуется использовать источники, приведенные в п. 7.

7 УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ УЧЕБНОЙ ДИСЦИПЛИНЫ

7.1 Основная литература

№ п/п	Библиографическое описание	Гриф	Количество экземпляров/URL
1	Мартишин, С. А. 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 / С.А. Мартишин, В.Л. Симонов, М.В. Храпченко. — М. : ИНФРА-М, 2021. — 235 с.	Рекомендовано в качестве учебного пособия для студентов, бакалавров и магистрантов направлений подготовки «Информатика и вычислительная техника», «Информационные системы и технологии», «Программная инженерия», «Прикладная математика и информатика», «Прикладная информатика», «Управление в технических системах», «Бизнес-информатика», «Педагогическое образование» (профиль «Информатика»)	https://znanium.com/catalog/product/1214862

7.2 Дополнительная литература

№ п/п	Библиографическое описание	Гриф	Количество экземпляров/URL
1	Варфоломеева, А. О. Информационные системы предприятия : учебное пособие / А.О. Варфоломеева, А.В. Коряковский, В.П. Романов. — 2-е изд., перераб. и доп. — М. : ИНФРА-М, 2022. — 330 с. — (Высшее образование: Бакалавриат).	Рекомендовано УМО по образованию в области прикладной информатики в качестве учебного пособия для студентов ВУЗов, обучающихся по направлению 09.03.03 «Прикладная информатика» и экономическим специальностям	https://znanium.com/catalog/product/1844303
2	Григорьев, А. А. Передача, хранение и обработка больших объемов научных данных : учебное пособие / А.А. Григорьев, Е.А. Исаев, П.А. Тарасов. — М. : ИНФРА-М, 2024. — 207 с. — (Высшее образование: Бакалавриат).	Рекомендовано Межрегиональным учебно-методическим советом профессионального образования в качестве учебного пособия для студентов ВУЗов, обучающихся по направлениям подготовки 01.03.02 «Прикладная	https://znanium.com/catalog/product/1073525

		математика и информатика», 09.03.01 «Информатика и вычислительная техника», 38.03.05 Бизнес-информатика» квалификация (степень «бакалавр»)	
3	Келлехер, Д. Наука о данных: базовый курс / Джон Келлехер, Брендан Тирни ; пер. с англ.. - М. : Альпина Паблшер, 2020. - 222 с.	–	https://znanium.com/catalog/product/1221800

7.3 Перечень ресурсов сети Интернет по изучаемой дисциплине

1. *Intuit.ru*
2. *Edx.com*

7.4 Перечень наглядных и других пособий, методических рекомендаций по проведению учебных занятий, а также методических материалов к используемым в образовательном процессе техническим средствам

7.4.1 Методические рекомендации

Захарченков К. В., Мрочек Т.В. Распределенные системы обработки информации. Методические рекомендации к лабораторным работам для студентов направления подготовки 09 03 01 «Информатика и вычислительная техника» дневной формы обучения (электронный вариант).

7.4.3 Информационные технологии

- Тема 1. Введение.
- Тема 2. Процесс анализа больших объемов данных.
- Тема 3. Программное обеспечение для автоматизации обработки больших объемов данных.
- Тема 4. Машинное обучение.
- Тема 5. Общие методы обработки больших объемов данных.
- Тема 6. Средства распределенного хранения и обработки больших данных.
- Тема 7. Глубокий анализ текста.

7.4.3 Перечень программного обеспечения, используемого в образовательном процессе

1. *MS Visual Studio 2022 Community Edition (свободно распространяемое)*
2. *Eclipse (свободно распространяемое)*

8 МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ УЧЕБНОЙ ДИСЦИПЛИНЫ

Материально-техническое обеспечение дисциплины содержится в паспорте лаборатории а. 517/2, рег. № паспорта лаборатории № ПУЛ - 4 517/2-23; в паспорте лаборатории а. 518/2, рег. № паспорта лаборатории № ПУЛ - 4 518/2-23.