Министерство науки и высшего образования Российской Федерации

федеральное государственное бюджетное образовательное учреждение высшего образования
РОССИЙСКИЙ ГОСУДАРСТВЕННЫЙ ГИДРОМЕТЕОРОЛОГИЧЕСКИЙ УНИВЕРСИТЕТ

Кафедра прикладной информатики

Рабочая программа дисциплины

Б1.В.05.03 Хранение и обработка данных

Основная профессиональная образовательная программа
высшего образования по направлению подготовки

09.03.03 Прикладная информатика

Направленность (профиль):

«Прикладные информационные системы и технологии»

Уровень:
Бакалавриат

Форма обучения
Очная

Согласовано
Руководитель ОПОП

_____ **Колбина О.Н.**

Утверждаю
Проректор по учебной работе

_____**Н.О. Верещагина**

Рекомендована решением
Ученного совета института Информационных систем и геотехнологий
29.06.2023 г., протокол №7

Рассмотрена и утверждена на заседании кафедры
28.06.2023 г., протокол №6
И.о. заведующего кафедрой
прикладной информатики

_____ **Истомин Е.П.**

Автор-разработчик:
к.т.н. Колбина О.Н.

Санкт-Петербург 2023

1. Цель и задачи освоения дисциплины

Цель дисциплины — сформировать профессиональную компетентность и необходимый объем фундаментальных и прикладных знаний, умений и навыков в области эффективного управления большими объемами данных, их надежного хранения, быстрой и точной обработки, а также извлечения полезной информации для поддержки принятия решений.

Задачи:

1. Сформировать знание:
   - принципов работы с большими данными (Big Data) и распределенными системами хранения и обработки данных;
   - инструментов и библиотек для анализа данных (SQL).
2. Сформировать умение:
   - выполнять запросы к базам данных с использованием SQL и других языков запросов;
   - осуществлять ETL-процессы для загрузки, трансформации и выгрузки данных.
3. Сформировать владение:
   - умением оптимизировать запросы и структуры данных для повышения производительности;
   - способностью создавать и поддерживать надёжные и масштабируемые системы хранения и обработки данных.

2. Место дисциплины в структуре основной профессиональной образовательной программы

Дисциплина относится к части, формируемой участниками образовательных отношений, изучается в 7 семестре.

Изучению предшествуют следующие дисциплины:

«Траектория профессионального развития», «Моделирование предметной области исследования», «Основы профессиональной деятельности», «Цифровизация профессиональной деятельности», «Технологии программирования», «Основы социального взаимодействия», «Фундаментальные знания о личности и обществе», «Естественно-научные знания в профессиональной деятельности», «Основы проектной деятельности в профессиональной сфере», «Разработка визуальных систем», «Основы деловой коммуникации», «Создание IoT-систем», «Основы использования информационно-коммуникационных технологий для профессиональной деятельности».

Изучается параллельно с такими дисциплинами как:

«Разработка пользовательских интерфейсов», «Разработка документации пользовательского интерфейса», «Тестирование пользовательских интерфейсов», «Анализ и управление рисками», «Надежность технических систем», «Оптимизация программного кода по средством распределенных вычислений», «Исследование операций и методы оптимизации».

Дисциплина является базовой для дисциплин:

«Создание геоинформационной системы», «Основы внедрения информационных систем», «Разработка программы и методики испытаний», «Сопровождение внедренной информационной системы», «Принципы геоинформационного управления», производственная практика (проектно-технологическая), проектирование систем, преддипломной практики, выполнения и защиты выпускной квалификационной работы.

3. Перечень планируемых результатов обучения

Процесс изучения дисциплины направлен на формирование компетенций: ПК-5

Таблица 1. **Компетенции**

Код и наименование компетенции	Код и наименование индикатора достижения компетенции	Результаты обучения
ПК-5 Способен проектировать пользовательский интерфейс по готовому образцу или концепции интерфейса	ПК-5.3 Систематизировать и структурировать информацию путем создания баз данных	Знать: – принципов работы с большими данными (Big Data) и распределенными системами хранения и обработки данных; – инструментов и библиотек для анализа данных (SQL). Уметь: – выполнять запросы к базам данных с использованием SQL и других языков запросов; – осуществлять ETL-процессы для загрузки, трансформации и выгрузки данных. Владеть: – умением оптимизировать запросы и структуры данных для повышения производительности; – способностью создавать и поддерживать надёжные и масштабируемые системы хранения и обработки данных.

4. Структура и содержание дисциплины

4.1. Объем дисциплины

Общая трудоемкость дисциплины составляет 4 зачетные единицы, 144 академических часа.

Таблица 2. Объем дисциплины по видам учебных занятий в академических часах

Объём дисциплины	Очная форма обучения	
	Семестр	Итого
	7 семестр	
Зачётные единицы	4	4
Контактная работа обучающихся с преподавателем	72	72

(по видам аудиторных учебных занятий) – всего:		
в том числе:	-	-
– лекции	32	32
– занятия семинарского типа:	40	40
практические занятия	-	-
лабораторные занятия	40	40
Самостоятельная работа (далее – СРС) – всего:	70,84	70,84
в том числе:	-	-
– курсовая работа	-	-
– контрольная работа	-	-
Контроль:		
Текущий контроль успеваемости (ТКУ)	1	1
Промежуточная аттестация	0,16	0,16
ВСЕГО ЧАСОВ:	144	144
Вид промежуточной аттестации	Зачет	Зачет

4.2. Структура дисциплины

Таблица 3. Структура дисциплины

№	Тема дисциплины	Виды учебной работы, в т.ч. самостоятельная работа студентов, час.			Формы текущего контроля	Формируемые компетенции	Индикаторы достижения компетенций
		Лекции	Лабораторные занятия	Самостоятельная работа			
1	Основы реляционных баз данных и SQL	8	13	20,5	Устная защита результатов лабораторной работы №1, 2	ПК-5	ПК-5.3
2	Продвинутые SQL-техники и оптимизация запросов	6	7	9	Устная защита результатов лабораторной работы №3	ПК-5	ПК-5.3
3	Текущий контроль успеваемости (ТКУ)			1	Тест в Moodle	ПК-5	ПК-5.3
4	ETL-процессы и обработка больших объемов данных	10	13	30	Устная защита результатов лабораторной работы №4,5	ПК-5	ПК-5.3
5	Распределенные системы хранения и Big Data	8	7	10,34	Устная защита результатов лабораторной работы №6	ПК-5	ПК-5.3

№	Тема дисциплины	Виды учебной работы, в т.ч. самостоятельная работа студентов, час.			Формы текущего контроля	Формируемые компетенции	Индикаторы достижения компетенций
		Лекции	Лабораторные занятия	Самостоятельная работа			
-	ИТОГО	32	40	70,84	-	-	-

4.3. Содержание разделов дисциплины

Таблица 4. Содержание разделов дисциплины

№	Наименование темы дисциплины	Содержание	Компетенция
1	Тема 1. Основы реляционных баз данных и SQL	Реляционные модели данных (нормальные формы 1NF-3NF). Основы SQL: DDL (CREATE, ALTER, DROP), DML (INSERT, UPDATE, DELETE), DQL (SELECT с WHERE, ORDER BY, LIMIT). Агрегатные функции (COUNT, SUM, AVG, GROUP BY, HAVING). Типы JOIN (INNER, LEFT, RIGHT, FULL). Подзапросы и CTE (WITH). Индексация таблиц (B-tree, Hash).	ПК-5
2	Тема 2. Продвинутые SQL-техники и оптимизация запросов	Window functions (ROW_NUMBER, RANK, LAG/LEAD, PARTITION BY). Пивотинг/анпивотинг данных. Рекурсивные CTE для иерархических данных. EXPLAIN ANALYZE и планы запросов. Оптимизация медленных запросов (композитные индексы, покрывающие индексы, материализованные представления). Partitioning таблиц и vacuum/analyze.	ПК-5
3	Тема 3. ETL-процессы и обработка больших объемов данных	ETL архитектура (Extract, Transform, Load). SQL скрипты для data pipeline: мерж данных (MERGE/UPSERT), дедупликация (ROW_NUMBER+CTE), агрегация временных рядов. PostgreSQL COPY/FDW для bulk загрузки. Интеграция SQL с Python (psycopg2, SQLAlchemy). Обработка JSON/XML в PostgreSQL (jsonb, xpath).	ПК-5

№	Наименование темы дисциплины	Содержание	Компетенция
4	Тема 4. Распределенные системы хранения и Big Data	Принципы CAP-теоремы и BASE vs ACID. Распределенные SQL (CockroachDB, YugabyteDB, TiDB). NoSQL в сравнении с RDBMS (ключ-значение, document, column-family). Apache Spark SQL для распределенной обработки. Hadoop HDFS и Hive для petabyte данных. Kafka и SQL streaming (ksqlDB). Шардинг и репликация в PostgreSQL (Citus).	ПК-5

4.4. Содержание занятий семинарского типа

Таблица 5. Содержание лабораторных занятий

№ темы дисциплины	Тематика лабораторных работ	Всего часов	Формируемые компетенции
1	Лабораторная работа №1. Основы SQL: DDL, DML, простые запросы SELECT	6	ПК-5
2	Лабораторная работа №2. Комплексные SQL запросы с JOIN, агрегатами и оконными функциями	7	ПК-5
2	Лабораторная работа №3. Оптимизация SQL запросов: индексы и EXPLAIN ANALYZE	7	ПК-5
3	Лабораторная работа №4. ETL-процесс: трансформация и загрузка данных через SQL скрипты	7	ПК-5
3	Лабораторная работа №5. Обработка JSON/NoSQL данных в PostgreSQL	6	ПК-5
4	Лабораторная работа №6. Шардинг и репликация в распределенных БД (PostgreSQL Citus)	7	ПК-5
-	ВСЕГО	40	

5. Перечень учебно-методического обеспечения самостоятельной работы обучающихся по дисциплине

1. Электронный учебный курс «Хранение и обработка данных» в системе Moodle [Электронный ресурс]. Режим доступа: https://moodle.rshu.ru/course/view.php?id=2420

6. Оценочные средства для текущего контроля успеваемости и промежуточной аттестации по итогам освоения дисциплины

Учет успеваемости обучающегося по дисциплине осуществляется по 100-балльной шкале.

Таблица 6. Учёт успеваемости обучающегося по дисциплине

Учет успеваемости	Количество баллов
Максимальное количество баллов по дисциплине за один семестр:	100
− максимальное количество баллов за выполнение всех видов текущего контроля	70
− максимальное количество баллов за прохождение промежуточной аттестации	30

6.1. Текущий контроль

Задания, методика выполнения и критерии оценивания текущего контроля по разделам дисциплины представлены в Фонде оценочных средств по данной дисциплине.

6.2. Промежуточная аттестация

Перечень вопросов и критерии оценивания ответов на вопросы в билете по темам дисциплины представлены в Фонде оценочных средств по данной дисциплине.

Форма промежуточной аттестации по дисциплине: зачет**.**

Форма проведения зачета: устный ответ на один вопрос в билете и индивидуальное практическое задание.

6.3. Балльно-рейтинговая система оценивания

Таблица 7. Распределение баллов по видам учебной работы

Вид учебной работы, за которую ставятся баллы	Баллы
Текущий контроль успеваемости	0-70
Промежуточная аттестация	0-30
ИТОГО	0-100

Таблица 7.1. Распределение баллов по текущему контролю

№	Вид работ	Min	Max
1. Обязательная часть			
1.1 Текущий контроль успеваемости по проверке форсированности остаточных знаний			
Текущий контроль успеваемости. Тест:		4	10
1.2 Выполнение практических работ:			
1.2.1	Лабораторная работа №1. Основы SQL: DDL, DML, простые запросы SELECT	4	6
1.2.2	Лабораторная работа №2. Комплексные SQL запросы с JOIN, агрегатами и оконными функциями	4	8
1.2.3	Лабораторная работа №3. Оптимизация SQL запросов: индексы и EXPLAIN ANALYZE	4	10
1.2.4	Лабораторная работа №4. ETL-процесс: трансформация и загрузка данных через SQL скрипты	4	10
1.2.5	Лабораторная работа №5. Обработка JSON/NoSQL данных в PostgreSQL	4	8
1.2.6	Лабораторная работа №6. Шардинг и репликация в распределенных БД (PostgreSQL Citus)	4	8
Итого баллов по обязательной части		20	50
2. Вариативная часть			

2.1	Реферат по дисциплине	1	4
2.2	Участие в НИРС	15	25
2.3	Участие в региональной олимпиаде по профилю	10	20
2.3.1	участие	10	10
2.3.2	призер	20	20
2.4	Публикация в индексируемом журнале (совместно с преподавателем)	10	10
2.5	Акселерационная программа/ проект Росмолодежи	20	40
2.5.1	участие	20	20
2.5.2	грант	40	40
2.6	Конспект параграфа литературы в рамках одной лабораторной работы дисциплины	2	10
2.7	Создание одного запроса SQL в рамках одной темы дисциплины	3	10
2.8	Разработка ETL-процесса в рамках одной лабораторной работы дисциплины	4	10
	Промежуточная аттестация по дисциплине	0	30
Итого баллов по вариативной части		40	60
Итого баллов по дисциплине			100

Таблица 7.2. Конвертация баллов в итоговую оценку

Оценка	Баллы
Зачтено	40-100
Незачтено	0-39

7. Методические рекомендации для обучающихся по освоению дисциплины

Методические рекомендации ко всем видам аудиторных занятий, а также методические рекомендации по организации самостоятельной работы, в том числе по подготовке к текущему контролю и промежуточной аттестации представлены в Методических рекомендации для обучающихся по освоению дисциплины.

8. Учебно-методическое и информационное обеспечение дисциплины

8.1. Перечень основной и дополнительной учебной литературы

Основная литература:

1. Советов, Б. Я. Информационные технологии : учебник для вузов / Б. Я. Советов, В. В. Цехановский. — 8-е изд., перераб. и доп. — Москва : Издательство Юрайт, 2023. — 414 с. — (Высшее образование). — ISBN 978-5-534-20054-6. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: https://urait.ru/bcode/559897
2. Воронов, М. В. Автоматическое управление. Управление организационными системами. Цифровые платформы : учебник для вузов / М. В. Воронов, В. И. Пименов, И. А. Небаев. — Москва : Издательство Юрайт, 2023. — 475 с. — (Высшее образование). — ISBN 978-5-534-19845-4. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: https://urait.ru/bcode/569236

Дополнительная литература:

1. Трофимов, В. В. Глобальные и локальные сети : учебник для вузов / В. В. Трофимов, М. И. Барабанова, В. И. Кияев. — 4-е изд., перераб. и доп. — Москва : Издательство Юрайт, 2023. — 151 с. — (Высшее образование). — ISBN 978-5-534-20428-5. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: https://urait.ru/bcode/568695
2. А. Ю. Сидоренко Технологии искусственного интеллекта в геоинформационных системах / А. Ю. Сидоренко, С. Ю. Степанов, Я. А. Петров [и др.] // Информационные технологии и системы: управление, экономика, транспорт, право. – 2019. – № 3(35). – С. 115-122. – EDN PXETFE.
3. Астапчук, В. А. Корпоративные информационные системы: требования при проектировании: учебное пособие для вузов / В. А. Астапчук, П. В. Терещенко. — 2-е изд., испр. и доп. — Москва: Издательство Юрайт, 2021. — 113 с. — (Высшее образование). — ISBN 978-5-534-08546-4. — Текст: электронный // Образовательная платформа Юрайт [сайт]. — URL: https://urait.ru/bcode/472111
4. Берджесс, Э. Искусственный интеллект - для вашего бизнеса: практическое руководство / Э. Берджесс. - Москва: Интеллектуальная Литература, 2021. - 232 с. - ISBN 9-785-907274-81-5. - Текст: электронный. - URL: https://znanium.com/catalog/product/1842395
5. Мартын, И. А. Определение внутренних волн по данным спутникового зондирования / И. А. Мартын // Информационные технологии и системы: управление, экономика, транспорт, право. – 2020. – № 2(38). – С. 23-30. – EDN OJWCND.

8.2. Перечень ресурсов информационно-телекоммуникационной сети "Интернет"

1. Образовательная платформа Нетология [Электронный ресурс]. Режим доступа: https://netology.ru/
2. Образовательная платформа Яндекс Практикум [Электронный ресурс]. Режим доступа: https://practicum.yandex.ru/
3. Образовательная платформа GeekBrains [Электронный ресурс]. Режим доступа: https://gb.ru/
4. Образовательная платформа Skillbox [Электронный ресурс]. Режим доступа: https://skillbox.ru/
5. Образовательная платформа SkillFactory [Электронный ресурс]. Режим доступа: https://skillfactory.ru/
6. Образовательная платформа Открытое образование [Электронный ресурс]. Режим доступа: https://openedu.ru/
7. Образовательная платформа Лекториум [Электронный ресурс]. Режим доступа: https://www.lektorium.tv/

8.3. Перечень программного обеспечения

1. Операционная система: Astra linux [Электронный ресурс]. Режим доступа: https://astralinux.ru/
2. Операционная система: Alt linux [Электронный ресурс]. Режим доступа: https://www.basealt.ru/alt-education/
3. Браузер: Яндекс браузер [Электронный ресурс]. Режим доступа: https://browser.yandex.ru/
4. Файловый архиватор: 7-zip [Электронный ресурс]. Режим доступа: https://www.7-zip.org/

5. Файловый менеджер: Far-manager [Электронный ресурс]. Режим доступа: https://farmanager.com/
6. Офисный пакет: OpenOffice [Электронный ресурс]. Режим доступа: https://www.openoffice.org/ru/
7. Разработка 2D и 3D визуализации данных [Электронный ресурс]. Режим доступа: https://www.goldensoftware.com/
8. PostgreSQL 17 (Community Edition) [Электронный ресурс]. Режим доступа: https://www.postgresql.org/download/

8.4. Перечень информационных справочных систем

1. Веб-геоинформационная платформа [Электронный ресурс]. Режим доступа: https://kosmosnimki.ru/
2. Веб-портал в области ГИС и ДЗЗ [Электронный ресурс]. Режим доступа: https://gis-lab.info/
3. Веб-портал в области свободного программного обеспечения [Электронный ресурс]. Режим доступа: http://www.opennet.ru/
4. Веб-портал в области современных технологий [Электронный ресурс]. Режим доступа: https://www.computerra.ru/
5. Информационный портал «ГИС-ассоциация: Межрегиональная общественная организация содействия развитию рынка геоинформационных технологий и услуг» [Электронный ресурс]. Режим доступа: http://www.gisa.ru/
6. Информационный портал «Научная Россия» [Электронный ресурс]. Режим доступа: https://scientificrussia.ru/
7. Сетевое издание «CNews» («СиНьюс») [Электронный ресурс]. Режим доступа: https://www.cnews.ru/
8. Сетевое издание «IT-World: Мир цифровых и информационных технологий» [Электронный ресурс]. Режим доступа: https://www.it-world.ru/

8.5. Перечень профессиональных баз данных

1. База данных исследований Центра стратегических разработок [Электронный ресурс]. Режим доступа: https://www.csr.ru/ru/research/
2. База данных международных индексов научного цитирования Scopus [Электронный ресурс]. Режим доступа: http://www.scopus.com/
3. База данных международных индексов научного цитирования Web of Science [Электронный ресурс]. Режим доступа: http://webofscience.com/
4. База данных НП «Международное Исследовательское Агентство «Евразийский Монитор» [Электронный ресурс]. Режим доступа: https://eurasiamonitor.org/issliedovaniia
5. База книг и публикаций электронной библиотеки «Наука и Техника» [Электронный ресурс]. Режим доступа: https://n-t.ru/
6. Базы данных официальной статистики Федеральной службы государственной статистики [Электронный ресурс]. Режим доступа: https://rosstat.gov.ru/statistic
7. Геопортал данных ДЗЗ Роскосмоса [Электронный ресурс]. Режим доступа: https://gptl.ru/
8. Электронная библиотечная система «Znanium» [Электронный ресурс]. Режим доступа: https://znanium.ru/
9. Электронная библиотечная система «Юрайт» [Электронный ресурс]. Режим доступа: https://urait.ru/
10. Электронная научная библиотека «Elibrary» [Электронный ресурс]. Режим доступа: https://elibrary.ru/

11. Электронная научная библиотека «КиберЛенинка» [Электронный ресурс]. Режим доступа: https://cyberleninka.ru/
12. Национальное управления океанических и атмосферных исследований NOAA [Электронный ресурс]. Режим доступа: https://www.iaea.org/
13. ЕСИМО – межведомственная федеральная информационная система. Единая государственная система информации об обстановке в мировом океане [Электронный ресурс]. Режим доступа: https://esimo.ru/
14. Федеральная служба государственной статистики (Профессиональная база данных) [Электронный ресурс]. Режим доступа: http://www.gks.ru/
15. Официальная статистика РФ ЕМИСС [Электронный ресурс]. Режим доступа: https://www.fedstat.ru/

9. Материально-техническое обеспечение дисциплины

Материально-техническое обеспечение программы соответствует действующим санитарно-техническим и противопожарным правилам и нормам и обеспечивает проведение всех видов аудиторных занятий и самостоятельной работы студентов.

Учебная аудитории для проведения занятий лекционного типа — укомплектована специализированной (учебной) мебелью, набором демонстрационного оборудования и учебно-наглядными пособиями, обеспечивающими тематические иллюстрации, соответствующие рабочим учебным программам дисциплин .

Учебная аудитория для групповых и индивидуальных консультаций — укомплектована специализированной (учебной) мебелью, техническими средствами обучения, служащими для представления учебной информации.

Учебная аудитория для текущего контроля и промежуточной аттестации — укомплектована специализированной (учебной) мебелью, техническими средствами обучения, служащими для представления учебной информации.

Учебно-научный лабораторный центр «ИНФОГЕО» — укомплектована специализированной (учебной) мебелью, компьютерами, служащими для работы с информацией.

Помещение для самостоятельной работы — укомплектовано специализированной (учебной) мебелью, оснащено компьютерной техникой с возможностью подключения к сети "Интернет" и обеспечено доступом в электронную информационно-образовательную среду организации.

Помещение для хранения и профилактического обслуживания учебного оборудования.

10. Особенности освоения дисциплины для инвалидов и лиц с ограниченными возможностями здоровья

Обучение обучающихся с ограниченными возможностями здоровья при необходимости осуществляется на основе адаптированной рабочей программы с использованием специальных методов обучения и дидактических материалов, составленных с учетом особенностей психофизического развития, индивидуальных возможностей и состояния здоровья таких обучающихся (обучающегося).

При определении формы проведения занятий с обучающимся-инвалидом учитываются рекомендации, содержащиеся в индивидуальной программе реабилитации инвалида, относительно рекомендованных условий и видов труда.

При необходимости для обучающихся из числа инвалидов и лиц с ограниченными возможностями здоровья создаются специальные рабочие места с учетом нарушенных функций и ограничений жизнедеятельности.

11. Возможность применения электронного обучения и дистанционных образовательных технологий

Дисциплина может реализовываться с применением электронного обучения и дистанционных образовательных технологий.

Рассмотрено и рекомендовано к использованию в учебном процессе на 24/25 учебный год без изменений*

Протокол заседания кафедры прикладной информатики от 27.06.2024 №6

Рассмотрено и рекомендовано к использованию в учебном процессе на 25/26 учебный год без изменений*

Протокол заседания кафедры прикладной информатики от 20.06.2025 №6

Рассмотрено и рекомендовано к использованию в учебном процессе на 26/27 учебный год без изменений*

Протокол заседания кафедры прикладной информатики от 11.06.2026 №12

*Заполняется при ежегодном пересмотре программы, если в неё не внесены изменения