



МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ
федеральное государственное бюджетное образовательное учреждение
высшего образования
МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ГЕОДЕЗИИ И КАРТОГРАФИИ
(МИИГАиК)

Факультет геоинформатики и информационной безопасности
Кафедра информатики и геоинформационных технологий



**ДОКУМЕНТ ПОДПИСАН
ЭЛЕКТРОННОЙ ПОДПИСЬЮ**

Сертификат: 5387FD2AF5B17FED816FDE9A1D4973B1
Владелец: Степанченко Алексей Леонидович
Действителен: с 21.07.2025 по 14.10.2026

«УТВЕРЖДАЮ»

Проректор
А. Л. Степанченко
16 февраля 2024 г.

**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)
«Б1.В.ДВ.04.02 БОЛЬШИЕ ПРОСТРАНСТВЕННЫЕ ДАННЫЕ»**

Уровень высшего образования: бакалавриат

Направление подготовки: 09.03.02 Информационные системы и технологии

Направленность(профиль) подготовки: Информационные системы и технологии

Квалификация (степень) выпускника: бакалавр

Форма обучения: очная

Год набора (приема на обучение): 2024

Срок получения образования: 4 года

Объем:
в зачетных единицах: 4 з.е.
в академических часах: 144 ак.ч.

Москва, 2024

Разработчики:

Преподаватель кафедры Информатики и геоинформационных технологий Кондрашин Д. М.

Рабочая программа дисциплины (модуля) составлена в соответствии с требованиями ФГОС ВО по направлению подготовки 09.03.02 Информационные системы и технологии, утвержденного приказом Минобрнауки от 19.09.2017 № 926

Согласование и утверждение

№	Подразделение	Ответственное лицо	ФИО	Виза	Дата, протокол (при наличии)
1	факультет Геоинформатики и информационной безопасности	Председатель ученого совета факультета	Матерухин А. В.	Рассмотрено	14.12.2023, № 7-23

Содержание

1. Цель и задачи освоения дисциплины (модуля)
2. Планируемые результаты обучения по дисциплине (модулю), соотнесенные с планируемыми результатами освоения образовательной программы
3. Место дисциплины в структуре ОП
4. Объем дисциплины и виды учебной работы
5. Содержание дисциплины
 - 5.1. Разделы, темы дисциплины и виды занятий
 - 5.2. Содержание разделов, тем дисциплины и формы текущего контроля
6. Оценочные материалы текущего контроля
7. Оценочные материалы промежуточной аттестации
8. Порядок проведения промежуточной аттестации
9. Материально-техническое и учебно-методическое обеспечение дисциплины
 - 9.1. Перечень основной и дополнительной учебной литературы
 - 9.2. Профессиональные базы данных и ресурсы «Интернет», к которым обеспечивается доступ обучающихся
 - 9.3. Программное обеспечение и информационно-справочные системы, используемые при осуществлении образовательного процесса по дисциплине
 - 9.4. Оборудование и технические средства обучения
10. Методические указания по освоению дисциплины (модуля)

1. Цель и задачи освоения дисциплины (модуля)

Цель освоения дисциплины - практическое применение современных программных средств и распределенных вычислительных парадигм для эффективного импорта, масштабируемой обработки, проведение пространственного анализа больших наборов векторных и растровых геоданных.

Задачи изучения дисциплины:

- Освоить концептуальные основы GeoBig Data и определить подходящий технологический стек для решения задач, связанных с масштабом;
- Владеть инструментами GDAL для быстрой предобработки и виртуализации больших наборов растровых данных ;
- Осуществлять эффективный импорт, хранение и управление большими векторными наборами данных в СУБД PostGIS с использованием утилит командной строки;
- Применять расширенные функции пространственного SQL в PostGIS для выполнения векторного анализа;
- Выполнять запросы к растровым данным, хранящимся в PostGIS Raster;
- Применять библиотеки Python для масштабируемого вычисления геопространственных индексов;
- Публиковать результаты проведенного анализа.

2. Планируемые результаты обучения по дисциплине (модулю), соотнесенные с планируемыми результатами освоения образовательной программы

Компетенции, индикаторы и результаты обучения

ПК-2 Способен разрабатывать требования и осуществлять проектирование программного обеспечения

ПК-2.1 Анализирует возможность реализации требований к программному обеспечению

Знать:

ПК-2.1/Зн5 Принципы работы с векторными и растровыми форматами в масштабе Big Data

Уметь:

ПК-2.1/Ум5 Провести сравнительный анализ двух альтернативных реализаций и обосновать выбор

Владеть:

ПК-2.1/Нв4 Навыками конфигурации и ограничениями распределенных сред при оценке масштабируемости решения

ПК-2.2 Проектирует программное обеспечения

Знать:

ПК-2.2/Зн5 Роли и взаимодействие ключевых компонентов стека (GDAL, PostGIS, GeoServer) в единой системе

Уметь:

ПК-2.2/Ум6 Проектировать структуру распределенного вычисления для выполнения сложного пространственного расчета

Владеть:

ПК-2.2/Нв4 Навыками проектирования схемы базы данных, включая правильное размещение пространственных столбцов и применение индексов

3. Место дисциплины в структуре ОП

Дисциплина (модуль) Б1.В.ДВ.04.02 «Большие пространственные данные» относится к формируемой участниками образовательных отношений части образовательной программы и изучается в семестре(ах): 7.

В процессе изучения дисциплины студент готовится к решению типов задач профессиональной деятельности, предусмотренных ФГОС ВО и образовательной программой.

4. Объем дисциплины (модуля) и виды учебной работы

Период обучения	Общая трудоемкость (часы)	Общая трудоемкость (ЗЕТ)	Лекционные занятия (часы)	Практические занятия (часы)	Самостоятельная работа (часы)	Промежуточная аттестация (часы)
7 семестр	144	4	34	34	76	Зачет с оценкой
Всего	144	4	34	34	76	

5. Содержание дисциплины (модуля)

5.1. Разделы, темы дисциплины и виды занятий (часы промежуточной аттестации не указываются)

Наименование раздела, темы	Всего	Лекционные занятия	Практические занятия	Самостоятельная работа	Планируемые результаты с обучения, соответствующие результатам освоения программы
Раздел 1. Основы работы с большими пространственными данными	36	10	8	18	ПК-2.1 ПК-2.2
Тема 1.1. Введение в большие пространственные данные	14	6	2	6	
Тема 1.2. Импорт и первичное исследование векторных данных	10		4	6	
Тема 1.3. Виртуальные растровые слои	12	4	2	6	
Раздел 2. Пространственный анализ на уровне систем управления базами данных	42	12	10	20	ПК-2.1 ПК-2.2
Тема 2.1. Расширенный пространственный SQL-анализ	18	6	4	8	
Тема 2.2. Оптимизация производительности в СУБД	10	2	2	6	

Тема 2.3. Основы работы с растровыми данными в СУБД	14	4	4	6	
Раздел 3. Распределенные вычисления для векторных данных	36	6	10	20	ПК-2.1 ПК-2.2
Тема 3.1. Настройка Python-среды для Big Data	8		2	6	
Тема 3.2. Работа с большими векторными наборами через Dask	16	4	4	8	
Тема 3.3. Распределенное пространственное соединение	12	2	4	6	
Раздел 4. Масштабируемая обработка растров с помощью Python	30	6	6	18	ПК-2.1 ПК-2.2
Тема 4.1. Настройка Python и Dask для геообработки	8		2	6	
Тема 4.2. Распределенный расчет индексов	10	2	2	6	
Тема 4.3. Публикация данных через WMS	12	4	2	6	
Итого	144	34	34	76	

5.2. Содержание разделов, тем дисциплин и формы текущего контроля

Раздел 1. Основы работы с большими пространственными данными

Тема 1.1. Введение в большие пространственные данные

Обзор концепций GeoBig Data; Обзор стека технологий (PostGIS, GeoPandas, облачные решения); Подключение PostGIS к QGIS.

Текущий контроль

Вид (форма) контроля
Тестирование

Тема 1.2. Импорт и первичное исследование векторных данных

Загрузка векторного набора данных; Импорт данных в PostGIS с использованием утилит; Выполнение базовых SQL-запросов; Настройка стилей в QGIS.

Текущий контроль

Вид (форма) контроля
Тестирование

Тема 1.3. Виртуальные растровые слои

Подготовка набора из множества небольших растровых файлов; Использование gdalbuildvrt для создания виртуального набора; Загрузка VRT в QGIS и анализ производительности.

Текущий контроль

Вид (форма) контроля
Тестирование

Раздел 2. Пространственный анализ на уровне систем управления базами данных

Тема 2.1. Расширенный пространственный SQL-анализ

Выполнение операций буферизации и пересечения; Анализ близости и создание матриц расстояний; Сохранение результатов анализа в новые таблицы.

Текущий контроль

Вид (форма) контроля
Тестирование

Тема 2.2. Оптимизация производительности в СУБД

Создание GiST индексов для пространственных столбцов; Сравнение скорости запросов с индексом и без него.

Текущий контроль

Вид (форма) контроля
Тестирование

Тема 2.3. Основы работы с растровыми данными в СУБД

Настройка расширения PostGIS Raster; Импорт набора растровых данных (raster2pgsql); Выполнение запросов к растровым данным.

Текущий контроль

Вид (форма) контроля
Тестирование

Раздел 3. Распределенные вычисления для векторных данных

Тема 3.1. Настройка Python-среды для Big Data

Настройка Python, GeoPandas, Dask; Обзор архитектуры Dask (Scheduler, Workers).

Текущий контроль

Вид (форма) контроля
Тестирование

Тема 3.2. Работа с большими векторными наборами через Dask

Чтение GeoJSON/CSV-файла в Dask GeoDataFrame; Выполнение негеометрических операций (фльтрация, агрегация) в распределенном режиме; Преобразование части данных обратно в GeoPandas для финальной обработки.

Текущий контроль

Вид (форма) контроля
Тестирование

Тема 3.3. Распределенное пространственное соединение

Разделение набора данных на несколько Dask DataFrames; Выполнение пространственного соединения с использованием Dask; Сохранение результата в PostGIS или CSV.

Текущий контроль
Вид (форма) контроля
Тестирование

Раздел 4. Масштабируемая обработка растров с помощью Python

Тема 4.1. Настройка Python и Dask для геообработки

Настройка Rasterio, Dask; Изучение концепции chunking в контексте GeoTIFF.

Текущий контроль
Вид (форма) контроля
Тестирование

Тема 4.2. Распределенный расчет индексов

Чтение многополосного растра с поддержкой Dask через Rasterio; Написание сценария для параллельного вычисления NDVI; Сравнение времени выполнения с нераспределенным методом.

Текущий контроль
Вид (форма) контроля
Тестирование

Тема 4.3. Публикация данных через WMS

Настройка GeoServer; Создание рабочей директории и подключение к базе PostGIS; Публикация обработанного растрового слоя как WMS-сервиса.

Текущий контроль
Вид (форма) контроля
Тестирование

6. Порядок проведения промежуточной аттестации

Порядок проведения промежуточной аттестации регламентируется локальными нормативными актами Университета:

- Положением о формах, периодичности, порядке текущего контроля успеваемости и промежуточной аттестации обучающихся по основным образовательным программам высшего образования – программам бакалавриата, программам специалитета, программам магистратуры в федеральном государственном бюджетном образовательном учреждении высшего образования «Московский государственный университет геодезии и картографии», п.6 «Организация и проведение промежуточной аттестации».
- Положением о порядке осуществления образовательного процесса, проведения текущего контроля успеваемости и промежуточной аттестации обучающихся с применением электронного обучения и дистанционных образовательных технологий в федеральном государственном бюджетном образовательном учреждении высшего образования «Московский государственный университет геодезии и картографии», п.8 «Проведение промежуточной аттестации с применением ЭО и ДОТ».

Промежуточная аттестация: Зачет с оценкой, 7 семестр.

7. Оценочные материалы текущего контроля

Раздел 1. Основы работы с большими пространственными данными

Контролируемые ИДК: ПК-2.1 ПК-2.2

Тема 1.1. Введение в большие пространственные данные

Форма контроля/оценочное средство: Тестирование

Вопросы/Задания:

1. Какой инструмент в QGIS используется для подключения к базе данных PostGIS и просмотра таблиц?

- A. Vector Layers Panel
- Б. DB Manager
- В. Processing Toolbox
- Г. Layer Properties

2. Что является ключевой характеристикой (V) GeoBig Data?

- A. Объем (Volume)
- Б. Разнообразие (Variety)
- В. Точность (Precision)
- Г. Скорость (Velocity)

Тема 1.2. Импорт и первичное исследование векторных данных

Форма контроля/оценочное средство: Тестирование

Вопросы/Задания:

1. Какая утилита используется для загрузки шейп-файла в PostGIS из командной строки?

- A. ogr2ogr
- Б. shp2pgsql
- В. raster2pgsql
- Г. gdalinfo

2. Какие команды SQL используются для первичного исследования векторных данных в PostGIS?

- A. SELECT * FROM table LIMIT 10;
- Б. SELECT ST_GeometryType(geom) FROM table;
- В. CREATE INDEX ...
- Г. SELECT count(*) FROM table;

Тема 1.3. Виртуальные растровые слои

Форма контроля/оценочное средство: Тестирование

Вопросы/Задания:

1. Какую команду необходимо вызвать для создания VRT из набора перекрывающихся GeoTIFF файлов?

- A. gdal_translate
- Б. gdalwarp
- В. gdalbuildvrt
- Г. rasterio.open()

2. Какие преимущества дает VRT по сравнению с физическим слиянием (merge) большого количества растровых тайлов?

- A. Нет физического копирования данных (виртуальный доступ).
- Б. Улучшение производительности в некоторых ГИС-клиентах.
- В. Постоянное сжатие всех исходных файлов.
- Г. Возможность выполнения SQL-запросов к слою.

Раздел 2. Пространственный анализ на уровне систем управления базами данных

Тема 2.1. Расширенный пространственный SQL-анализ

Форма контроля/оценочное средство: Тестирование

Вопросы/Задания:

1. Какая функция PostGIS используется для создания буфера радиусом 100 метров вокруг всех полигонов в таблице zones?

- A. ST_Intersection(geom, 100)
- Б. ST_Buffer(geom, 100)
- В. ST_Distance(geom, 100)
- Г. ST_Union(geom, 100)

2. Какие операции в SQL-анализе позволяют определить взаимоотношения между двумя векторными слоями?

- A. ST_Area()
- Б. ST_Intersects()
- В. ST_Contains()
- Г. ST_Overlaps()

Тема 2.2. Оптимизация производительности в СУБД

Форма контроля/оценочное средство: Тестирование

Вопросы/Задания:

1. Какой тип индекса является предпочтительным для ускорения пространственных запросов в PostGIS?

- A. B-tree
- Б. GIN
- В. GiST
- Г. Hash

2. Какие действия необходимо выполнить для диагностики того, что пространственный запрос выполняется медленно, несмотря на наличие индекса?

- A. Выполнить EXPLAIN ANALYZE для запроса.
- Б. Удалить геометрию из таблицы.
- В. Выполнить VACUUM ANALYZE на таблице.
- Г. Заменить SRID.

Тема 2.3. Основы работы с растровыми данными в СУБД

Форма контроля/оценочное средство: Тестирование

Вопросы/Задания:

1. Какая утилита используется для загрузки растровых данных (GeoTIFF) в PostGIS Raster?

- A. shp2pgsql
- Б. gdalbuildvrt
- В. raster2pgsql
- Г. ogr2ogr

2. Какие шаги необходимы для извлечения среднего значения растрового слоя в пределах заданного полигона с помощью PostGIS?

- A. Активировать расширение PostGIS_Raster.
- Б. Использовать функцию, комбинирующую растровый доступ и ST_Intersects.
- В. Использовать функцию ST_Value или ST_DumpAsPolygons.
- Г. Выполнить команду gdal_translate.

Раздел 3. Распределенные вычисления для векторных данных

Контролируемые ИДК: ПК-2.1 ПК-2.2

Тема 3.1. Настройка Python-среды для Big Data

Форма контроля/оценочное средство: Тестирование

Вопросы/Задания:

1. В архитектуре Dask, какая компонента отвечает за отслеживание выполнения задач и управление рабочими узлами?

- A. Worker
- Б. Client
- В. Scheduler
- Г. Manager

2. Какие библиотеки необходимо установить для работы с большими векторными данными в распределенном режиме, как это показано в курсе?

- A. GeoPandas
- Б. PostGIS
- В. Dask
- Г. Rasterio

Тема 3.2. Работа с большими векторными наборами через Dask

Форма контроля/оценочное средство: Тестирование

Вопросы/Задания:

1. Какую команду вы используете для чтения очень большого GeoJSON файла, чтобы он был автоматически разделен на чанки Dask?

- A. `pd.read_csv()`
- Б. `gpd.read_file()`
- В. `ddf = dd.read_json(...)`
- Г. `gpd.read_file(..., engine='dask')`

2. Какие операции с Dask GeoDataFrame (DDF) могут быть выполнены в распределенном режиме, не требуя немедленного сбора данных в Pandas?

- A. Фильтрация по атрибутивным полям.
- Б. Группировка и агрегация (`groupby().sum()`).
- В. Запись результата обратно в PostGIS.
- Г. Визуализация с помощью стандартного Matplotlib.

Тема 3.3. Распределенное пространственное соединение

Форма контроля/оценочное средство: Тестирование

Вопросы/Задания:

1. В контексте Dask, как называется операция, которая ищет объекты, находящиеся в радиусе X друг от друга, если данные разделены на несколько Dask DataFrames?

- A. `merge`
- Б. `join`
- В. `Spatial Join`
- Г. `set_index`

2. Какие шаги необходимы для безопасного выполнения пространственного соединения между двумя очень большими наборами данных с помощью Dask?

- A. Убедиться, что оба набора данных являются Dask DataFrames.
- Б. Обеспечить правильное "картирование" (mapping) геометрий для минимизации передачи данных между рабочими узлами.
- В. Использовать только буферизацию SQL.
- Г. Использовать функцию, поддерживающую Dask (например, из GeoPandas/Dask интеграции).

Раздел 4. Масштабируемая обработка растров с помощью Python

Тема 4.1. Настройка Python и Dask для геообработки

Форма контроля/оценочное средство: Тестирование

Вопросы/Задания:

1. При чтении большого GeoTIFF с помощью Rasterio и Dask, какая опция в функции open() критически важна для активации распределенной обработки?

- А. driver='GTiff'
- Б. compress='LZW'
- В. chunks
- Г. crs='EPSG:4326'

2. Что включает в себя концепция chunking для растровых данных, обрабатываемых Dask/Rasterio?

- А. Файл физически разбивается на мелкие файлы.
- Б. Каждый чанк может быть прочитан независимо.
- В. Чанки являются многомерными массивами NumPy.
- Г. Чанки всегда имеют форму 1*N.

Тема 4.2. Распределенный расчет индексов

Форма контроля/оценочное средство: Тестирование

Вопросы/Задания:

1. При расчете индекса NDVI на распределенном растре, какой Dask-совместимый инструмент используется для сохранения результата в новый GeoTIFF с сохранением chunking-структуры?

- А. gdal_translate
- Б. rasterio.open()
- В. Dask Array .to_raster()
- Г. shutil.copy()

2. Какие преимущества дает использование Dask Arrays (по сравнению с обычными NumPy) при расчете сложных математических индексов на больших растрах?

- А. Поддержка данных, превышающих RAM.
- Б. Ленивое вычисление (lazy evaluation).
- В. Автоматическое распараллеливание.
- Г. Сжатие данных в памяти.

Тема 4.3. Публикация данных через WMS

Форма контроля/оценочное средство: Тестирование

Вопросы/Задания:

1. Какой стандарт OGC используется для публикации обработанных растровых данных (например, NDVI) в виде динамического картографического сервиса?

- А. WMS (Web Map Service)
- Б. WFS (Web Feature Service)
- В. WCS (Web Coverage Service)
- Г. SOS (Sensor Observation Service)

2. Какие основные компоненты необходимы для настройки и запуска WMS-сервиса, который публикует данные из PostGIS?

- А. GeoServer
- Б. Подключение к PostGIS
- В. Стили (SLD)
- Г. WMS-прокси

8. Оценочные материалы промежуточной аттестации

7 семестр, Зачет с оценкой

Контролируемые ИДК: ПК-2.1 ПК-2.2

Вопросы/Задания:

1. Какой инструмент командной строки (CLI) используется для импорта большого шейп-файла (SHP) в базу данных PostGIS?

- A. ogr2ogr
- Б. shp2pgsql
- B. raster2pgsql
- Г. gdal_translate

2. Какое расширение PostgreSQL необходимо активировать для работы с растровыми данными внутри базы данных?

- A. PostGIS_Vector
- Б. PostGIS_Topology
- B. PostGIS_Raster
- Г. PostGIS_Tiler

3. Какая структура данных в библиотеке Dask позволяет эффективно работать с геопространственными данными, разделяя их на части, которые не помещаются в оперативную память?

- A. Dask DataFrame
- Б. Dask Array
- B. Dask GeoDataFrame
- Г. Dask Cluster

4. Какой тип индекса в PostGIS используется для ускорения пространственных запросов, таких как пересечение или поиск ближайших соседей?

- A. B-tree
- Б. Hash
- B. GiST
- Г. GIN

5. Какая функция PostGIS используется для создания буфера вокруг векторного объекта?

- A. ST_Intersection
- Б. ST_Distance
- B. ST_Buffer
- Г. ST_Union

6. Для чего в первую очередь используется утилита gdalbuildvrt?

- A. Сжатие растровых данных.
- Б. Объединение нескольких растров в один логический виртуальный слой (VRT).
- B. Вычисление NDVI.
- Г. Создание индексов в PostGIS.

7. Какая библиотека Python необходима для чтения многополосных растров с поддержкой Dask и концепции chunking?

- A. GeoPandas
- Б. Rasterio
- B. NumPy
- Г. Matplotlib

8. Какая основная задача у Scheduler в архитектуре Dask?

- A. Хранение данных.
- Б. Управление и распределение задач между рабочими узлами (Workers).
- B. Выполнение финальной визуализации.

Г. Создание пространственных индексов.

9. Какая функция PostGIS используется для определения, пересекается ли один объект с другим?

- А. ST_Contains
- Б. ST_Touches
- В. ST_Intersects
- Г. ST_Distance

10. Какую роль играет DB Manager в QGIS при работе с PostGIS?

- А. Он выполняет сложные пространственные SQL-запросы.
- Б. Он служит графическим интерфейсом для подключения и просмотра таблиц базы данных.
- В. Он управляет правами доступа к СУБД.
- Г. Он конвертирует GeoTIFF в VRT.

11. Для чего в контексте распределенной обработки растров используется концепция chunking?

- А. Для сжатия выходного файла.
- Б. Для разделения большого раstra на небольшие части, которые могут обрабатываться независимо.
- В. Для изменения системы координат.
- Г. Для создания векторных слоев.

12. Какой инструмент CLI используется для публикации пространственных данных через WMS-сервис?

- А. raster2pgsql
- Б. gdal_translate
- В. GeoServer (как сервис)
- Г. shp2pgsql

13. Какие операции можно выполнить на уровне расширенного пространственного SQL-анализа в PostGIS?

- А. Вычисление площади объекта.
- Б. Создание буфера.
- В. Создание матрицы расстояний.
- Г. Фильтрация по атрибутивным полям.
- Д. Изменение SRID.

14. Какие преимущества дает создание GiST индекса для пространственной таблицы в PostGIS?

- А. Ускорение поиска объектов в пределах заданной области.
- Б. Ускорение стандартных негеометрических SQL-запросов.
- В. Ускорение операций, использующих пространственные предикаты (ST_Intersects).
- Г. Уменьшение физического размера базы данных.

15. Какие типы негеометрических операций можно выполнять на Dask GeoDataFrame в распределенном режиме?

- А. Фильтрация по атрибутам.
- Б. Пересчет NDVI.
- В. Агрегация (например, groupby().mean()).
- Г. Пространственное соединение.

16. Каковы основные преимущества использования Dask GeoDataFrame по сравнению с обычным GeoPandas DataFrame при работе с очень большими файлами?

- А. Возможность работать с данными, превышающими объем оперативной памяти.
- Б. Автоматическое управление параллельными вычислениями.
- В. Встроенная поддержка PostGIS.
- Г. Поддержка пространственного соединения в распределенном режиме.

17. Какие форматы данных являются распространенными при работе с большими наборами данных в контексте Big Data?

- А. CSV (с координатами X/Y).

- Б. GeoJSON.
- В. Shapefile (SHP).
- Г. Parquet.

18. Что необходимо учитывать при сравнении скорости запросов в PostGIS?

- А. Размер обрабатываемого пространственного индекса.
- Б. Количество рабочих узлов Dask.
- В. Использование и наличие GiST индекса.
- Г. Тип используемого процессора.

19. Сформулируйте основную проблему (вызов), которую решают концепции GeoBig Data.

20. Объясните, что такое "chunking" в контексте Rasterio/Dask и почему это критично для обработки больших растров.

9. Материально-техническое и учебно-методическое обеспечение дисциплины

9.1. Перечень основной и дополнительной учебной литературы

Основная литература

1. Макшанов А. В. Большие данные. Big Data: учебник для вузов: учебник для вузов / А. В. Макшанов, А. Е. Журавлев, Л. Н. Тындыкарь. - 4-е изд., стер. - Санкт-Петербург: Лань, 2024. - 188 с. - 978-5-507-47346-5. - Текст: электронный. // Издательство Лань: [сайт]. - URL: <https://e.lanbook.com/img/cover/book/362318.jpg> (дата обращения: 16.03.2025). - Режим доступа: по подписке

2. Мамедли Р. Э. Большие данные и NoSQL базы данных: учебное пособие для вузов: учебное пособие для вузов / Р. Э. Мамедли, Т. Б. Казиахмедов. - Санкт-Петербург: Лань, 2024. - 92 с. - 978-5-507-49873-4. - Текст: электронный. // Издательство Лань: [сайт]. - URL: <https://e.lanbook.com/img/cover/book/434051.jpg> (дата обращения: 16.03.2025). - Режим доступа: по подписке

Дополнительная литература

1. Сараджишвили С. Э. Цифровая обработка многомерных сигналов и Большие Данные: учебное пособие: учебное пособие / С. Э. Сараджишвили, И. А. Воронков. - Цифровая обработка многомерных сигналов и Большие Данные - Санкт-Петербург: Санкт-Петербургский политехнический университет Петра Великого, 2023. - 183 с. - 978-5-7422-7890-0. - Текст: электронный. // IPR SMART: [сайт]. - URL: <https://www.iprbookshop.ru/143008.html> (дата обращения: 16.03.2025). - Режим доступа: по подписке

9.2. Профессиональные базы данных и ресурсы «Интернет», к которым обеспечивается доступ обучающихся

Профессиональные базы данных

Не используются.

Ресурсы «Интернет»

1. <https://www.dask.org/> - Сайт, посвященный библиотеке Dask для параллельных вычислений на Python.

9.3. Программное обеспечение и информационно-справочные системы, используемые при осуществлении образовательного процесса по дисциплине

*Перечень программного обеспечения
(обновление производится по мере появления новых версий программы)*
Не используется.

*Перечень информационно-справочных систем
(обновление выполняется еженедельно)*
Не используется.

9.4. Оборудование и технические средства обучения

Компьютерный класс
Автоматизированное рабочее место (АРМ) - 12 шт.
Интерактивная доска - 1 шт.
Проектор - 1 шт.

10. Методические указания по освоению дисциплины (модуля)

С целью успешного освоения данного предмета студенту необходимо выполнять домашние работы по указанным преподавателям темам и представленным на занятиях материалам. Для более успешного выполнения работы студенту предлагается использование учебно-методических материалов, соответствующих указанной теме.

Самостоятельная работа студентов описывается и регулируется:

- Методическими рекомендациями по дисциплине;
- Методическими рекомендациями по организации самостоятельной работы студентов МИИГАиК;

Описание возможностей изучения дисциплины лицами с ОВЗ и инвалидами

При необходимости обучения студентов-инвалидов и лиц с ограниченными возможностями здоровья аудиторные занятия могут быть заменены или дополнены изучением полнотекстовых лекций, презентаций, видео- и аудиоматериалов в электронной информационно-образовательной среде (ЭИОС) университета. Индивидуальные задания подбираются в адаптированных к ограничениям здоровья формах (письменно или устно, в форме презентаций). Выбор методов обучения зависит от их доступности для инвалидов и лиц с ограниченными возможностями здоровья.

Рабочая программа дисциплины для инвалидов и лиц с ограниченными возможностями здоровья разрабатывается индивидуально с учетом особенностей психофизического развития, индивидуальных возможностей и состояния здоровья таких обучающихся.