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: 5BFEE1D286E9C95A0C0D2C611E434BD5
Владелец: Степанченко Алексей Леонидович
Действителен: с 23.09.2026 по 17.12.2027

«УТВЕРЖДАЮ»
Проректор
А. Л. Степанченко
18 апреля 2025 г.

ФОНД ОЦЕНОЧНЫХ СРЕДСТВ ПО ДИСЦИПЛИНЕ (МОДУЛЮ) «Б1.В.04 ЯЗЫКИ ПРОГРАММИРОВАНИЯ ДЛЯ РАБОТЫ С БОЛЬШИМИ ДАННЫМИ»

Уровень высшего образования: магистратура

Направление подготовки: 09.04.03 Прикладная информатика

Направленность (профиль) подготовки: Анализ больших пространственных данных

Квалификация (степень) выпускника: магистр

Форма обучения: очная

Год набора (приема на обучение): 2025

Срок получения образования: 2 года

Объем: в зачетных единицах: 3 з.е.
в академических часах: 108 ак.ч.

Москва, 2025

Разработчики:
Старший преподаватель кафедры Информатики и
геоинформационных технологий Стрельцов А. С.

Фонд оценочных средств по дисциплине (модулю) составлен в соответствии с требованиями ФГОС ВО по направлению подготовки 09.04.03 Прикладная информатика, утвержденного приказом Минобрнауки от 19.09.2017 № 916

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	14.10.2024, № 10-24

1. Планируемые результаты обучения, соотнесенные с планируемыми результатами освоения образовательной программы

ПК-1 Способен выполнять работы по сбору, обработке и анализу больших пространственных данных

ПК-1.1 Собирает большие пространственные данные

Знать:

ПК-1.1/Зн1 Методы и средства поиска и получения, систематизации и обработки аэрокосмической информации и снимков местности

Уметь:

ПК-1.1/Ум1 Планировать процессы и этапы обработки(получения) данных фотографической, оптико-электронной и других видов аэрокосмических съемок для решения определенного круга производственных и научных задач

Владеть:

ПК-1.1/Нв1 Методами планирования аэросъемочного процесса для получения изображений с заданными параметрами

ПК-1.2 Обрабатывает большие пространственные данные

Знать:

ПК-1.2/Зн1 Основы и физические принципы регистрации аэрокосмической информации о местности

ПК-1.2/Зн2 Технологии и алгоритмы сбора, обработки и анализа больших данных о Земле

Уметь:

ПК-1.2/Ум1 Осуществлять выбор аэрокосмических снимков, получаемых различными системами и комплексами дистанционного зондирования, в зависимости от решаемых практических или исследовательских задач

ПК-1.2/Ум2 Применять типовые технологии и алгоритмы сбора, обработки и анализа больших данных о Земле для решения практических задач

Владеть:

ПК-1.2/Нв1 Методами сбора и оценки качества пространственной информации, имеющей значения для подготовки и реализации проектов в соответствующих сферах профессиональной деятельности

ПК-1.2/Нв2 Навыками использования типовых технологий и алгоритмов сбора, обработки и анализа больших данных о Земле для решения практических задач

ПК-1.3 Анализирует большие пространственные данные

Знать:

ПК-1.3/Зн1 Основные виды и способы дистанционного зондирования и характеристики формируемых с их помощью снимков и изображений

ПК-1.3/Зн2 Технологии и алгоритмы анализа больших данных

ПК-1.3/Зн3 Программные среды и языки, применяемые для обработки больших объемов данных

Уметь:

ПК-1.3/Ум1 Ум1 проводить оценку фотографического качества аэрокосмических изображений и анализ соответствия его заданным в ТЗ на съемку параметрам

ПК-1.3/Ум2 Использовать научную и научно-техническую литературу в области обработки больших данных при программной реализации алгоритмов анализа больших данных

ПК-1.3/Ум3 Использовать применяемые в обработке больших данных программное обеспечение для решения аналитических и исследовательских задач профессиональной деятельности

Владеть:

ПК-1.3/Нв1 Навыками оценки качественных характеристик используемой при аэросъемке фотографической аппаратуры и соответствия их решаемой задаче

ПК-1.3/Нв2 Навыками использования типовых алгоритмов анализа больших данных для решения практических задач профессиональной деятельности

ПК-1.3/Нв3 Навыками использования специализированного программного обеспечения для анализа больших данных

2. Показатели и критерии оценивания компетенций

Планируемые результаты освоения программы	Уровень овладения	Критерии оценивания (перечень сформированных ИДК/ЗУВ по уровням освоения)
ПК-1 Способен выполнять работы по сбору, обработке и анализу больших пространственных данных	Повышенный	ИДК ПК-1.1 Собирает большие пространственные данные Знать: ПК-1.1/Зн1 Методы и средства поиска и получения, систематизации и обработки аэрокосмической информации и снимков местности Уметь: ПК-1.1/Ум1 Планировать процессы и этапы обработки(получения) данных фотографической, оптико-электронной и других видов аэрокосмических съемок для решения определенного круга производственных и научных задач Владеть: ПК-1.1/Нв1 Методами планирования аэросъемочного процесса для получения изображений с заданными параметрами
ПК-1 Способен выполнять работы по сбору, обработке и анализу больших пространственных данных	Повышенный	ИДК ПК-1.2 Обрабатывает большие пространственные данные Знать: ПК-1.2/Зн1 Основы и физические принципы регистрации аэрокосмической информации о местности ПК-1.2/Зн2 Технологии и алгоритмы сбора, обработки и анализа больших данных о Земле Уметь: ПК-1.2/Ум1 Осуществлять выбор аэрокосмических снимков, получаемых различными системами и комплексами дистанционного зондирования, в зависимости от решаемых практических или исследовательских задач ПК-1.2/Ум2 Применять типовые технологии и алгоритмы сбора, обработки и анализа больших данных о Земле для решения практических задач

		Владеть: ПК-1.2/Нв1 Методами сбора и оценки качества пространственной информации, имеющей значения для подготовки и реализации проектов в соответствующих сферах профессиональной деятельности ПК-1.2/Нв2 Навыками использования типовых технологий и алгоритмов сбора, обработки и анализа больших данных о Земле для решения практических задач ИДК ПК-1.3 Анализирует большие пространственные данные Знать: ПК-1.3/Зн1 Основные виды и способы дистанционного зондирования и характеристики формируемых с их помощью снимков и изображений ПК-1.3/Зн2 Технологии и алгоритмы анализа больших данных ПК-1.3/Зн3 Программные среды и языки, применяемые для обработки больших объемов данных Уметь: ПК-1.3/Ум1 Ум1 проводить оценку фотографического качества аэрокосмических изображений и анализ соответствия его заданным в ТЗ на съемку параметрам ПК-1.3/Ум2 Использовать научную и научно-техническую литературу в области обработки больших данных при программной реализации алгоритмов анализа больших данных ПК-1.3/Ум3 Использовать применяемые в обработке больших данных программное обеспечение для решения аналитических и исследовательских задач профессиональной деятельности Владеть: ПК-1.3/Нв1 Навыками оценки качественных характеристик используемой при аэросъемке фотографической аппаратуры и соответствия их решаемой задаче ПК-1.3/Нв2 Навыками использования типовых алгоритмов анализа больших данных для решения практических задач профессиональной деятельности ПК-1.3/Нв3 Навыками использования специализированного программного обеспечения для анализа больших данных

	Базовый	ИДК ПК-1.1 Собирает большие пространственные данные Знать: ПК-1.1/Зн1 Методы и средства поиска и получения, систематизации и обработки аэрокосмической информации и снимков местности Уметь: ПК-1.1/Ум1 Планировать процессы и этапы обработки(получения) данных фотографической, оптико-электронной и других видов аэрокосмических съемок для решения определенного круга производственных и научных задач
	Базовый	ИДК ПК-1.2 Обрабатывает большие пространственные данные Знать: ПК-1.2/Зн2 Технологии и алгоритмы сбора, обработки и анализа больших данных о Земле ПК-1.2/Зн1 Основы и физические принципы регистрации аэрокосмической информации о местности Уметь: ПК-1.2/Ум2 Применять типовые технологии и алгоритмы сбора, обработки и анализа больших данных о Земле для решения практических задач ПК-1.2/Ум1 Осуществлять выбор аэрокосмических снимков, получаемых различными системами и комплексами дистанционного зондирования, в зависимости от решаемых практических или исследовательских задач ИДК ПК-1.3 Анализирует большие пространственные данные Знать: ПК-1.3/Зн3 Программные среды и языки, применяемые для обработки больших объемов данных ПК-1.3/Зн2 Технологии и алгоритмы анализа больших данных ПК-1.3/Зн1 Основные виды и способы дистанционного зондирования и характеристики формируемых с их помощью снимков и изображений Уметь: ПК-1.3/Ум3 Использовать применяемые в обработке больших данных программное обеспечение для решения аналитических и исследовательских задач профессиональной деятельности ПК-1.3/Ум2 Использовать научную и научно-техническую литературу в области обработки больших данных при программной реализации алгоритмов анализа больших данных

		ПК-1.3/Ум1 Ум1 проводить оценку фотографического качества аэрокосмических изображений и анализ соответствия его заданным в ТЗ на съемку параметрам
	Пороговый	ИДК ПК-1.1 Собирает большие пространственные данные Знать: ПК-1.1/Зн1 Методы и средства поиска и получения, систематизации и обработки аэрокосмической информации и снимков местности
	Пороговый	ИДК ПК-1.2 Обрабатывает большие пространственные данные Знать: ПК-1.2/Зн2 Технологии и алгоритмы сбора, обработки и анализа больших данных о Земле ПК-1.2/Зн1 Основы и физические принципы регистрации аэрокосмической информации о местности ИДК ПК-1.3 Анализирует большие пространственные данные Знать: ПК-1.3/Зн2 Технологии и алгоритмы анализа больших данных ПК-1.3/Зн3 Программные среды и языки, применяемые для обработки больших объемов данных ПК-1.3/Зн1 Основные виды и способы дистанционного зондирования и характеристики формируемых с их помощью снимков и изображений

3. Шкала оценивания

Оценка	Критерии оценивания (перечень сформированных ИДК/ЗУВ)
Зачтено	ПК-1.1 Собирает большие пространственные данные ПК-1.1/Зн1 Методы и средства поиска и получения, систематизации и обработки аэрокосмической информации и снимков местности ПК-1.2 Обрабатывает большие пространственные данные ПК-1.2/Зн1 Основы и физические принципы регистрации аэрокосмической информации о местности ПК-1.2/Зн2 Технологии и алгоритмы сбора, обработки и анализа больших данных о Земле ПК-1.3 Анализирует большие пространственные данные ПК-1.3/Зн1 Основные виды и способы дистанционного зондирования и характеристики формируемых с их помощью снимков и изображений ПК-1.3/Зн3 Программные среды и языки, применяемые для обработки больших объемов данных ПК-1.3/Зн2 Технологии и алгоритмы анализа больших данных
Не зачтено	Результаты обучения не сформированы на достаточном уровне

4. Контрольные мероприятия по дисциплине

Вид контроля	Форма контроля/Оценочное средство
Текущий контроль	Просмотр выполненных работ
Промежуточная аттестация	Зачет

№ п/п	Наименование раздела	Контролируемые ИДК	Вид контроля/ используемые оценочные материалы	
			Текущий	Промежут. аттестация
1	Основы Python для обработки больших данных	ПК-1.2	Просмотр выполненных работ	Зачет
2	Объектно-ориентированное программирование для масштабируемых решений	ПК-1.2	Просмотр выполненных работ	Зачет
3	NumPy и Pandas для работы с большими табличными данными	ПК-1.1 ПК-1.2 ПК-1.3	Просмотр выполненных работ	Зачет
4	Распределенные вычисления с Dask	ПК-1.2	Просмотр выполненных работ	Зачет
5	Обработка больших пространственных векторных данных	ПК-1.1 ПК-1.2 ПК-1.3	Просмотр выполненных работ	Зачет
6	Обработка больших растровых данных	ПК-1.1 ПК-1.2 ПК-1.3	Просмотр выполненных работ	Зачет
7	Облачные платформы для больших геоданных	ПК-1.1 ПК-1.2 ПК-1.3	Просмотр выполненных работ	Зачет
8	Оптимизация и профилирование	ПК-1.2	Просмотр выполненных работ	Зачет

5. 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, 1 семестр.

Зачет по дисциплине является формой промежуточной аттестации и проводится в форме тестирования.

Студентам будет предложен тест, содержащий 20 вопросов. На выполнение работы отводится

60 минут, в течение которых использование вспомогательных материалов и электронных устройств запрещено.

Оценивание проводится по принципу «зачтено» или «не зачтено» Для успешной сдачи зачета и получения отметки «зачтено» студенту необходимо дать правильные ответы не менее чем на 60% вопросов, что составляет 12 правильных ответов из 20. В случае, если количество правильных ответов составит менее 12, зачет считается не сданным и выставляется отметка «не зачтено»

6. Оценочные материалы текущего контроля

Раздел 1. Основы Python для обработки больших данных
Контролируемые ИДК: ПК-1.2

Тема 1.2. Управляющие конструкции и обработка исключений
Форма контроля/оценочное средство: Просмотр выполненных работ

Тема 1.3. Модули, пакеты и генераторы
Форма контроля/оценочное средство: Просмотр выполненных работ

Раздел 2. Объектно-ориентированное программирование для масштабируемых решений
Контролируемые ИДК: ПК-1.2

Тема 2.2. ООП-рефакторинг и оптимизация
Форма контроля/оценочное средство: Просмотр выполненных работ

Раздел 3. NumPy и Pandas для работы с большими табличными данными
Контролируемые ИДК: ПК-1.1 ПК-1.2 ПК-1.3

Тема 3.2. Pandas для анализа больших датасетов
Форма контроля/оценочное средство: Просмотр выполненных работ

Раздел 4. Распределенные вычисления с Dask
Контролируемые ИДК: ПК-1.2

Тема 4.3. Dask Delayed для параллелизации функций
Форма контроля/оценочное средство: Просмотр выполненных работ

Раздел 5. Обработка больших пространственных векторных данных
Контролируемые ИДК: ПК-1.1 ПК-1.2 ПК-1.3

Тема 5.2. Dask-GeoPandas для распределенной обработки
Форма контроля/оценочное средство: Просмотр выполненных работ

Тема 5.3. API и потоковая обработка геоданных
Форма контроля/оценочное средство: Просмотр выполненных работ

Раздел 6. Обработка больших растровых данных
Контролируемые ИДК: ПК-1.1 ПК-1.2 ПК-1.3

Тема 6.3. Распределенная обработка растров

Форма контроля/оценочное средство: Просмотр выполненных работ

Раздел 7. Облачные платформы для больших геоданных
Контролируемые ИДК: ПК-1.1 ПК-1.2 ПК-1.3

Тема 7.3. Интеграция облачных и локальных вычислений
Форма контроля/оценочное средство: Просмотр выполненных работ

Раздел 8. Оптимизация и профилирование
Контролируемые ИДК: ПК-1.2

Тема 8.3. Профилирование и оптимизация памяти
Форма контроля/оценочное средство: Просмотр выполненных работ

7. Оценочные материалы промежуточной аттестации

1 семестр, Зачет
Контролируемые ИДК: ПК-1.1 ПК-1.2 ПК-1.3
Вопросы/Задания:

1. Какой модуль стандартной библиотеки Python предоставляет функции для создания эффективных итераторов?

2. Как называется библиотека, позволяющая масштабировать операции GeoPandas на наборы данных, превышающие объём оперативной памяти, за счёт разбиения на партиции и распределённой обработки?

3. Как называется тип данных Pandas, рекомендуемый для столбцов с небольшим количеством уникальных строковых значений с целью экономии памяти?

4. Дополните предложение термином:
Параметр _____ при чтении больших CSV-файлов средствами Pandas позволяет читать и обрабатывать файл по частям (чанкам), а не загружать его целиком в память.

5. Дополните предложение термином:
Формат хранения данных _____ является колоночным и хорошо подходит для аналитических запросов к большим датасетам, в отличие от строкового формата CSV.

6. Чем Dask DataFrame принципиально отличается от Pandas DataFrame? Раскройте содержание партиционирования и поясните его роль в этом отличии.

7. Что такое GIL (Global Interpreter Lock)? Опишите его влияние на многопоточное выполнение CPU-интенсивных задач в Python и назовите способы его обхода, упомянутые в программе дисциплины.

8. Раскройте содержание понятия ленивых вычислений (lazy evaluation) применительно к библиотеке Dask: дайте определение, поясните роль графа задач и объясните, чем ленивое выполнение отличается от немедленного (eager) выполнения, характерного для обычного Pandas.

9. Раскройте содержание оптимизации пространственных соединений (spatial join) больших векторных данных в GeoPandas: объясните, зачем нужен пространственный индекс, и опишите, как его отсутствие сказывается на вычислительной сложности операции соединения.

10. Раскройте содержание архитектуры Google Earth Engine (GEE): объясните, в чём заключается принципиальное отличие серверной обработки данных GEE от традиционного подхода локальной обработки данных, и укажите, какое практическое следствие это имеет для работы с данными планетарного масштаба.

11. Расположите этапы чтения большого CSV-файла по частям (chunks) средствами Pandas и агрегации результата в правильной последовательности.

1. Накопление промежуточного результата обработки чанка в общем аккумуляторе.
2. Получение очередного чанка данных в виде отдельного DataFrame.
3. Повторение получения и обработки чанков до достижения конца файла, за которым следует получение итогового результата.
4. Открытие CSV-файла с указанием параметра chunksize.
5. Выполнение необходимой обработки (фильтрации, агрегации) над текущим чанком.

12. Расположите этапы построения пространственного индекса и его использования для пространственного соединения в GeoPandas в правильной последовательности.

1. Формирование итогового результата соединения на основе объектов, прошедших точную геометрическую проверку.
2. Точная геометрическая проверка пересечения только для найденных кандидатов, а не для всех объектов индексированного слоя.
3. Построение пространственного индекса (например, R-дерева) для одного из соединяемых слоёв.
4. Для каждого объекта другого слоя – быстрый поиск по индексу кандидатов, потенциально пересекающихся с этим объектом.

13. Расположите этапы обработки временной серии спутниковых снимков в Google Earth Engine в правильной последовательности.

1. Экспорт итогового результата в облачное хранилище.
2. Определение интересующей коллекции изображений (ImageCollection).
3. Агрегация результатов по всей отфильтрованной коллекции (например, вычисление среднего значения индекса за период).
4. Применение операции reduceRegion для получения значения по заданному региону.
5. Применение функции обработки к каждому изображению коллекции (например, вычисление индекса).
6. Фильтрация коллекции по временному диапазону и территории интереса.

14. Расположите этапы асинхронной обработки запросов к API геосервиса с использованием aiohttp в правильной последовательности.

1. Формирование списка асинхронных задач-запросов к API.
2. Сбор результатов всех завершившихся запросов после их одновременного выполнения.
3. Одновременный запуск всех задач с ожиданием их параллельного выполнения.

15. Расположите этапы профилирования и оптимизации памяти Python-скрипта, обрабатывающего большие данные, в правильной последовательности.

1. Анализ отчёта профилировщика и определение конкретных участков кода или структур данных, требующих оптимизации.
2. Повторный запуск профилировщика для проверки эффекта внесённых изменений.
3. Запуск скрипта с профилировщиком памяти (например, memory_profiler) для выявления мест наибольшего потребления памяти.
4. Внесение изменений в код (например, использование более компактных типов данных или обработка данных по частям).

16. Установите соответствие между библиотекой экосистемы Python для работы с большими геопространственными данными и типом данных, с которым она работает.

GeoPandas - Распределённые и параллельные вычисления над данными любого из перечисленных типов
RasterIO - Многомерные размеченные массивы (например, временные серии растровых снимков)
Xarray/rioxarray - Большие векторные (точечные, линейные, полигональные) геопространственные данные
Dask - Большие растровые файлы, чтение по окнам

17. Установите соответствие между форматом хранения данных и его характерной особенностью.

Parquet - Формат, разработанный для быстрой записи и чтения DataFrame между R и Python
Feather - Колоночный формат, хорошо подходящий для аналитических запросов к большим датасетам
Zarr - Простой текстовый строковый формат без колоночной оптимизации для аналитических запросов
CSV - Иерархический формат для хранения больших объёмов научных данных
HDF5 - Формат для хранения чанкованных, сжатых N-мерных массивов, облачно-оптимизированная альтернатива других форматов

18. Установите соответствие между инструментом профилирования Python-кода и тем, что он измеряет.

Line_profiler - Потребление оперативной памяти кодом
memory_profiler - Число вызовов и суммарное время выполнения каждой функции программы
cProfile - Время выполнения кода построчно

19. Установите соответствие между технологией (модулем) параллелизма Python и типом задач, для которого она подходит.

Multiprocessing - Задачи, ограниченные ожиданием ввода-вывода, с явным асинхронным управлением переключением
threading - CPU-интенсивные задачи, ускоряемые компиляцией критичных участков кода в машинный код
asyncio - Задачи, ограниченные ожиданием ввода-вывода, в пределах одного процесса
Numba/Cython - CPU-интенсивные задачи, требующие обхода GIL через независимые процессы

20. Установите соответствие между понятием Google Earth Engine и его определением.

ImageCollection - Предварительно загруженный и оптимизированный для обработки на серверах Google набор общедоступных геопространственных данных
reduceRegion - Стек или временная серия изображений, объединённых в единую коллекцию
Каталог данных GEE - Операция агрегации значений пикселей в пределах указанного региона