



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)
ОСНОВЫ РАЗРАБОТКИ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ
«Б1.О.06.10 СБОР И ОБРАБОТКА ОТКРЫТЫХ ДАННЫХ»**

Уровень высшего образования: бакалавриат

Направление подготовки: 09.03.03 Прикладная информатика

Направленность(профиль) подготовки: Прикладная информатика

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 Груздев С. С.

Преподаватель кафедры Информатики и геоинформационных технологий Кондрашин Д. М.

Рабочая программа дисциплины (модуля) составлена в соответствии с требованиями ФГОС ВО по направлению подготовки 09.03.03 Прикладная информатика, утвержденного приказом Минобрнауки от 19.09.2017 № 922

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 - формирование компетенций по комплексному управлению жизненным циклом открытых наборов данных с использованием современных информационно-коммуникационных технологий

Задачи изучения дисциплины:

- Провести концептуальное проектирование набора данных, включая контроль его содержания;
- Оценить потенциальную востребованность набора данных посредством анализа метрик популярности и поисковых запросов;
- Освоить методологию создания набора данных с использованием мобильных геоинформационных систем;
- Освоить методологию автоматизированного сбора данных, включая разработку и реализацию алгоритмов;
- Выполнить загрузку данных и первичное ознакомление с их структурой в специализированном ПО;
- Идентифицировать и проанализировать потенциальные проблемы данных: пропуски, выбросы и ошибки;
- Осуществить трансформацию данных: кодирование признаков, создание новых признаков и сводных таблиц;
- Выбрать и применить формальные меры для оценки качества набора данных;
- Определить соответствие набора данных требованиям публикации, включая формирование метайнформации;
- Документировать результаты всех этапов сбора и обработки, включая определение условий пользования данными.

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

Компетенции, индикаторы и результаты обучения

УК-1 Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач

УК-1.1 Анализирует задачу, выделяя ее базовые составляющие

Знать:

УК-1.1/Зн2 Принципы концептуального проектирования и контроля содержания наборов данных

Уметь:

УК-1.1/Ум2 Выделять ключевые составляющие задачи на этапе определения содержания набора данных

Владеть:

УК-1.1/Нв2 Применять системный подход к структурированию задачи на основе требований к конечному набору данных

УК-1.2 Осуществляет поиск информации для решения поставленной задачи по различным типам запросов

Знать:

УК-1.2/Зн2 Методологии определения и оценки потенциальной востребованности информации

Уметь:

УК-1.2/Ум2 Использовать поисковые и аналитические методы для оценки спроса на данные

Владеть:

УК-1.2/Нв2 Принимать решения о целесообразности сбора набора данных на основе анализа его востребованности

УК-1.3 При обработке информации отличает факты от мнений, интерпретаций, оценок, формирует собственные мнения и суждения, аргументирует свои выводы, в том числе с применением философского понятийного аппарата

Знать:

УК-1.3/Зн2 Критерии различия фактологической информации и субъективных оценок в процессе анализа данных

Уметь:

УК-1.3/Ум2 Критически оценивать и аргументировать выводы, полученные после сбора и первичной обработки данных

Владеть:

УК-1.3/Нв2 Формировать обоснованные суждения о данных, интегрируя результаты сбора и анализа востребованности

УК-1.4 Осуществляет поиск, сбор и обработку информации для решения поставленной задачи

Знать:

УК-1.4/Зн2 Методы и этапы создания наборов данных с использованием различных программных средств

Уметь:

УК-1.4/Ум2 Организовывать и реализовывать процесс поиска и сбора информации с применением ИКТ

Владеть:

УК-1.4/Нв2 Обеспечивать целостность и прослеживаемость всего процесса от концептуального проектирования до получения первичного массива данных

ОПК-3 Способен решать стандартные задачи профессиональной деятельности на основе информационной и библиографической культуры с применением информационно-коммуникационных технологий и с учетом основных требований информационной безопасности

ОПК-3.1 Формулирует информационный запрос и эффективно осуществляет поиск необходимой для решения стандартных задач профессиональной деятельности информации

Знать:

ОПК-3.1/Зн3 Принципы формулирования запросов, направленных на определение популярности тематик и потенциального спроса на данные

Уметь:

ОПК-3.1/Ум3 Эффективно искать информацию, необходимую для обоснования необходимости создания набора данных

Владеть:

ОПК-3.1/Нв2 Использовать результаты поисковой активности для корректировки целей сбора данных

ОПК-3.2 Применяет принципы разработки и верификации структуры программного кода для создания или адаптации компонентов информационно-коммуникационных технологий, обеспечивая их соответствие архитектуре системы и требованиям заказчика.

Знать:

ОПК-3.2/Зн3 Структуру и ключевые этапы автоматизированных и мобильных методов сбора данных с помощью ПО

Уметь:

ОПК-3.2/Ум3 Применять ИКТ-средства для автоматизации процессов сбора данных

Владеть:

ОПК-3.2/Нв3 Разрабатывать и адаптировать алгоритмы сбора данных, включая постобработку, с использованием соответствующего программного обеспечения

ОПК-3.3 Применяет информационно-коммуникационные технологии и программные средства для обработки, хранения, визуализации и представления данных в процессе решения стандартных задач профессиональной деятельности

Знать:

ОПК-3.3/Зн3 Основные этапы и процедуры предварительной обработки и анализа полученного набора данных

Уметь:

ОПК-3.3/Ум2 Использовать специализированное ПО для загрузки, очистки, кодирования и визуализации данных

Владеть:

ОПК-3.3/Нв2 Выполнять комплексную предварительную обработку данных, включая формирование рабочих гипотез и создание первичных визуализаций

ОПК-3.4 Реализует базовые меры и принципы обеспечения информационной безопасности при работе с данными, информационными системами и сетевыми ресурсами

Знать:

ОПК-3.4/Зн2 Требования к оформлению метаинформации и определению условий пользования набором данных

Уметь:

ОПК-3.4/Ум3 Применять меры для защиты данных на этапе сбора и обработки

Владеть:

ОПК-3.4/Нв2 Обеспечивать соответствие финального набора данных стандартам публикации, включая прозрачное описание условий использования и метаданных

ОПК-3.5 Анализирует, критически оценивает найденную информацию на достоверность и релевантность, а также корректно оформляет библиографические ссылки на используемые источники

Знать:

ОПК-3.5/Зн2 Критерии и методы оценки качества набора данных

Уметь:

ОПК-3.5/Ум3 Оценивать качество собранных данных

Владеть:

ОПК-3.5/Нв3 Обобщать результаты оценки качества и готовить финальный набор данных к публикации, подтверждая его соответствие заданным требованиям

3. Место дисциплины в структуре ОП

Дисциплина (модуль) Б1.О.06.10 «Сбор и обработка открытых данных» относится к обязательной 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	18	54	72	Зачет с оценкой
Всего	144	4	18	54	72	

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, соотношенные с результатами освоения программы
Раздел 1. Сбор открытых данных	75	9	28	38	УК-1.1 УК-1.2 УК-1.3 УК-1.4 ОПК-3.1 ОПК-3.2 ОПК-3.3 ОПК-3.4 ОПК-3.5
Тема 1.1. Концептуальное проектирование набора данных	14	3	4	7	
Тема 1.2. Определение потенциальной востребованности набора данных	16	3	4	9	
Тема 1.3. Создание набора данных с помощью программных средств	45	3	20	22	
Раздел 2. Обработка открытых данных	69	9	26	34	УК-1.1 УК-1.2 УК-1.3 УК-1.4 ОПК-3.1
Тема 2.1. Предварительная обработка набора данных	39	3	18	18	

Тема 2.2. Оценка качества набора данных	16	3	4	9	ОПК-3.1 ОПК-3.2 ОПК-3.3
Тема 2.3. Подготовка набора данных к публикации	14	3	4	7	ОПК-3.4 ОПК-3.5
Итого	144	18	54	72	

5.2. Содержание разделов, тем дисциплин и формы текущего контроля

Раздел 1. Сбор открытых данных

Тема 1.1. Концептуальное проектирование набора данных

Предварительный этап определения содержания набора данных; Определение содержания набора данных; Контроль содержания набора данных

Текущий контроль

Вид (форма) контроля
Практическая работа

Тема 1.2. Определение потенциальной востребованности набора данных

Определение популярности тематики; Анализ поисковых запросов; Востребованность набора данных по результатам опроса/анкетирования; Расчет общего коэффициента потенциальной востребованности набора данных

Текущий контроль

Вид (форма) контроля
Практическая работа

Тема 1.3. Создание набора данных с помощью программных средств

Использование мобильных геоинформационных систем: Подготовительный этап, Полевой этап, Камеральный этап, Обработка результатов полевого исследования; Автоматизированный сбор данных (Веб-скрейпинг): Подготовительный этап (этап «разведки»), Выбор программных средств автоматизации, Разработка и реализация алгоритма, Постобработка собранных данных

Текущий контроль

Вид (форма) контроля
Практическая работа

Раздел 2. Обработка открытых данных

Тема 2.1. Предварительная обработка набора данных

Загрузка данных в специализированное ПО; Ознакомление с полученными данными; Определение перечня возможных потенциальных проблем; Формирование рабочих гипотез относительно данных; Изучение специфики этапов обработки данных; Изучение свойств столбцов; Обнаружение пропущенных значений; Анализ выбросов и ошибочных данных; Кодирование категориальных признаков; Создание новых признаков; Создание сводных таблиц; Визуализация данных; Документирование результатов

Текущий контроль

Вид (форма) контроля
Практическая работа

Тема 2.2. Оценка качества набора данных

Указание единицы качества данных; Выбор меры качества; Определение метода оценки качества данных; Определение оценки качества данных на выходе; Обобщение результатов оценки качества

Текущий контроль
Вид (форма) контроля
Практическая работа

Тема 2.3. Подготовка набора данных к публикации

Определение соответствия набора данных требованиям; Определение метаинформации набора данных; Определение условий пользования набором данных

Текущий контроль
Вид (форма) контроля
Практическая работа

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. Сбор открытых данных

Контролируемые ИДК: ОПК-3.1 ОПК-3.2 ОПК-3.3 ОПК-3.4 ОПК-3.5 УК-1.1 УК-1.2 УК-1.3 УК-1.4

Тема 1.1. Концептуальное проектирование набора данных

Форма контроля/оценочное средство: Практическая работа

Вопросы/Задания:

1. Концептуальное проектирование набора данных
1. Ознакомиться с существующими наборами открытых данных на официальных сайтах публикаторов открытых данных
2. Среди предложенных тематик выбрать ту, которая связана с вашей деятельностью или в

области которой вы имеете знания

3. Проанализировать в каких сферах (категориях) больше всего публикуют наборы открытых данных
4. Описать какие атрибутивные характеристики должен иметь набор данных
5. Определить какой тип данных должен иметь каждый из атрибутов набора данных
6. Дать название набору данных. Название должно точно отражать его предметную область
7. Ознакомиться с нормативно-правовой документацией, регулирующих публикацию данных
8. Определить не попадает ли содержание набора данных под какой-либо из критериев

Тема 1.2. Определение потенциальной востребованности набора данных

Форма контроля/оценочное средство: Практическая работа

Вопросы/Задания:

1. Определение потенциальной востребованности набора данных
1. Определение популярности тематики
2. Анализ поисковых запросов
3. Востребованность набора данных по результатам опроса/анкетирования
4. Расчет общего коэффициента потенциальной востребованности набора данных

Тема 1.3. Создание набора данных с помощью программных средств

Форма контроля/оценочное средство: Практическая работа

Вопросы/Задания:

1. Создание набора данных с помощью программных средств
- При создании набора данных можно использовать один из предложенных способов: использование средств записи данных, использование мобильных средств для сбора пространственных данных и автоматизированный сбор данных.

Раздел 2. Обработка открытых данных

Контролируемые ИДК: ОПК-3.1 ОПК-3.2 ОПК-3.3 ОПК-3.4 ОПК-3.5 УК-1.1 УК-1.2 УК-1.3 УК-1.4

Тема 2.1. Предварительная обработка набора данных

Форма контроля/оценочное средство: Практическая работа

Вопросы/Задания:

1. Предварительная обработка набора данных
1. Загрузка данных в специализированное ПО.
2. Ознакомление с полученными данными.
3. Определение перечня возможных потенциальных проблем.
4. Формирование рабочих гипотез относительно данных.
5. Изучение специфики этапов обработки данных
6. Изучение свойств столбцов
7. Обнаружение пропущенных значений
8. Анализ выбросов и ошибочных данных
9. Кодирование категориальных признаков
10. Создание новых признаков
11. Создание сводных таблиц
12. Визуализация данных
13. Документирование результатов

Тема 2.2. Оценка качества набора данных

Форма контроля/оценочное средство: Практическая работа

Вопросы/Задания:

1. Оценка качества набора данных

Первым этапом необходимо указать единицы качества данных.
Во втором этапе необходимо выбрать меры качества.
Третий этап – определение метода оценки качества данных.
Четвертый этап – определение оценки качества данных на выходе.

Тема 2.3. Подготовка набора данных к публикации

Форма контроля/оценочное средство: Практическая работа

Вопросы/Задания:

1. Подготовка набора данных к публикации
- Определение соответствия набора данных требованиям.
Определение метаинформации набора данных.
Определение условий пользования набором данных.

8. Оценочные материалы промежуточной аттестации

7 семестр, Зачет с оценкой

Контролируемые ИДК: УК-1.1 УК-1.2 УК-1.3 УК-1.4 ОПК-3.1 ОПК-3.2 ОПК-3.3 ОПК-3.4 ОПК-3.5

Вопросы/Задания:

1. Какой этап предварительной обработки данных включает в себя изучение свойств столбцов?

- А) Загрузка данных
- б) Ознакомление с данными
- в) Определение потенциальных проблем
- г) Формирование рабочих гипотез

2. Какой метод используется для оценки качества данных?

- А) Однозначная оценка пригодности/непригодности
- б) Метод наименьших квадратов
- в) Метод главных компонент
- г) Метод контрольных сумм

3. Какие этапы включает в себя процесс создания набора данных с помощью мобильных геоинформационных систем?

- А) Подготовительный этап
- б) Полевой этап
- в) Камеральный этап
- г) Обработка результатов
- д) Верификация данных

4. Какие действия выполняются на этапе предварительной обработки набора данных?

- А) Загрузка данных в специализированное ПО
- б) Кодирование категориальных признаков
- в) Создание новых признаков
- г) Публикация данных
- д) Визуализация данных

5. Опишите процесс определения потенциальной востребованности набора данных

6. Перечислите и кратко опишите этапы автоматизированного сбора данных

7. Какие основные требования предъявляются к набору данных перед публикацией?

8. Опишите процесс оценки качества набора данных

9. Какие методы используются для выявления пропущенных значений в наборе данных?

10. Какие методы можно использовать для кодирования категориальных признаков?

9. Материально-техническое и учебно-методическое обеспечение дисциплины

9.1. Перечень основной и дополнительной учебной литературы

Основная литература

1. Титов А. Н. Python. Обработка данных: учебно-методическое пособие: учебно-методическое пособие / А. Н. Титов, Р. Ф. Тагиева. - Python. Обработка данных - Казань: Издательство КНИТУ, 2022. - 104 с. - 978-5-7882-3171-6. - Текст: электронный. // IPR SMART: [сайт]. - URL: <https://www.iprbookshop.ru/129220.html> (дата обращения: 16.03.2025). - Режим доступа: по подписке

2. Применение геоинформационной системы QGIS 3 для решения практических задач: лабораторный практикум: лабораторный практикум / составители: П. С. Куприенко, Е. П. Вялова, Н. В. Ильина. - Применение геоинформационной системы QGIS 3 для решения практических задач - Воронеж: Воронежский государственный технический университет, ЭБС АСВ, 2024. - 109 с. - 978-5-7731-1197-9. - Текст: электронный. // IPR SMART: [сайт]. - URL: <https://www.iprbookshop.ru/146999.html> (дата обращения: 16.03.2025). - Режим доступа: по подписке

Дополнительная литература

1. Лысаков К. Ф. Практическое программирование на Python: учебное пособие: учебное пособие / К. Ф. Лысаков. - Практическое программирование на Python - Новосибирск: Новосибирский государственный университет, 2023. - 75 с. - 2227-8397. - Текст: электронный. // IPR SMART: [сайт]. - URL: <https://www.iprbookshop.ru/134584.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://data.gov.spb.ru/> - Портал открытых данных Санкт-Петербурга
2. <https://data.gov.ru/> - Портал открытых данных РФ
3. <https://data.gov35.ru/> - Портал открытых данных Вологодской области
4. <https://data.mos.ru/> - Портал открытых данных Москвы
5. <https://data.yarregion.ru/> - Портал открытых данных Ярославской области
6. <https://fedstat.ru/> - Единая межведомственная статистическая система ЕМИСС
7. <https://qgis.org/> - Свободно распространяемое ПО ГИС с открытым кодом QGIS
8. <https://nextgis.ru/nextgis-mobile/> - Мобильная ГИС для работы на местности
9. <https://qfield.org/> - Мобильная ГИС для работы на местности

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.

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.