

**Федеральное государственное бюджетное образовательное учреждение
высшего образования
Московский государственный университет имени М.В. Ломоносова
Географический факультет**

УТВЕРЖДАЮ
Декан географического факультета,
член-корр. РАН Добролюбов С.А.

**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ
Автоматизированная обработка данных в страноведении**

Уровень высшего образования:
магистратура

Направление подготовки:
05.04.02 «География»

Направленность (профиль) ОПОП:
«Социально-экономическая география зарубежных стран»

Форма обучения:
очная

Рабочая программа рассмотрена и одобрена
Учебно-методической комиссией географического факультета
(протокол № 12, дата 8 декабря 2021 г.)

Москва 2021

Рабочая программа дисциплины разработана в соответствии с самостоятельно установленным МГУ образовательным стандартом (ОС МГУ) для реализуемых основных профессиональных образовательных программ высшего образования по направлению подготовки «География» (*программы бакалавриата, магистратуры, реализуемые последовательно по схеме интегрированной подготовки*).

ОС МГУ утверждены решением Ученого совета МГУ имени М.В. Ломоносова (приказ по МГУ № 1383 от 30 декабря 2020 года).

Год (годы) приема на обучение: 2021

1. Место дисциплины в структуре ОПОП — относится к вариативной части ОПОП, является дисциплиной по выбору.
2. Входные требования для освоения дисциплины, предварительные условия: базируется на знании основ геоинформатики, количественных методов в социально-экономической географии, методов анализа пространственных данных в социально-экономической географии, эконометрики, социологии.
3. Планируемые результаты обучения по дисциплине, соотнесенные с требуемыми компетенциями выпускников

Компетенции выпускников (коды)	Индикаторы (показатели) достижения компетенций	Планируемые результаты обучения по дисциплине, сопряженные с компетенциями
<p>МПК-5 (<i>формируется частично</i>) Способен анализировать причины и проявления географической неоднородности социально-экономического развития стран мира.</p>	<p>МПК-5.1. Умеет анализировать причины и проявления географической неоднородности социально-экономического развития стран мира.</p>	<p>Знать: основные структуры данных, применяемые в программировании; основы синтаксиса языка разметки HTML; подходы к автоматизированному поиску и систематизации информации в сети Интернет; основные понятия сетевого анализа; методы моделирования транспортных сетей; алгоритмы решения задач основные понятия теории фракталов; задачи социально-экономической географии, в которых применяется теория фракталов на графах.</p> <p>Уметь: создавать простые программные решения для автоматизированного сбора пространственной информации в сети Интернет, решать задачи по моделированию транспортных сетей и динамики городской застройки с использованием специального программного обеспечения.</p> <p>Владеть: основами программирования на языке Python, в том числе с использованием специализированных библиотек; навыками работы в специализированном программном обеспечении для решения задач на графах; навыками работы в специализированном программном обеспечении для фрактального анализа.</p>

4. Объем дисциплины 2 з.е., в том числе 36 академических часов на контактную работу обучающихся с преподавателем, 36 академических часов на самостоятельную работу обучающихся.

5. Формат обучения не предполагает электронного обучения и использования дистанционных образовательных технологий (за исключением форс-мажорных обстоятельств – пандемии и т.п.).

6. Содержание дисциплины, структурированное по темам (разделам) с указанием отведенного на них количества академических или астрономических часов и виды учебных занятий

Наименование и краткое содержание разделов и тем дисциплины, Форма промежуточной аттестации по дисциплине	Всего (часы)	В том числе						
		Контактная работа (работа во взаимодействии с преподавателем) <i>Виды контактной работы, часы*</i>					Самостоятельная работа обучающегося <i>Виды самостоятельной работы, часы</i>	
		Занятия лекционного типа	Занятия семинарского типа	Групповые консультации	Индивидуальные консультации	Всего	Домашняя работа	Всего
Тема 1. Основы программирования на языке Python	10	2	4	-	-	6	4	4
Тема 2. Работа с данными в Python	16	2	6	-	-	8	8	8
Тема 3. Веб-скрепинг в социально-экономической географии	18	4	6	-	-	10	8	8
Тема 4. Сетевой анализ в социально-экономической географии.	14	2	4	-	-	6	8	8
Тема 5. Использование теории фракталов в социально-экономической географии	10	4	2	-	-	6	4	4
Промежуточная аттестация — зачет	4	<i>Устный зачет</i>					4	
Итого	72	36					36	

Содержание лекций, семинаров

Тема 1. Основы программирования на языке Python

Язык Python: история развития, сферы применения. Типы и структуры данных в Python. Создание и присвоение значений переменных. Арифметические операции. Операции над строками. Условия. Циклы. Функции и подпрограммы. Импорт библиотек. Комментарии в коде. Принятые правила написания и оформления кода, PEP 8.

Дистрибутив Anaconda: назначение, возможности, сферы применения. Управление средами (environments), установка библиотек.

Среда Jupyter Notebook: установка, использование. Ячейки кода, Markdown-ячейки. Файл блокнота (*.ipynb). Способы обмена кодом.

Тема 2. Работа с данными в Python

Работа с файлами и папками в Python. Библиотека os. Чтение и запись файлов стандартными средствами. Специализированные библиотеки для чтения табличных данных: csv, pandas.

Работа с табличными данными средствами библиотеки pandas. Датафрейм, инструменты работы с датафреймом. Панельные данные. Представление панельных данных в «широком» и «длинном» формате.

Работа с пространственными данными в Python. Библиотека geopandas. Библиотеки osgeo: GDAL, OGR, OSR.

Тема 3. Веб-скрепинг в социально-экономической географии

Язык разметки HTML. Веб-страница как иерархический XML-документ. Основные структурные элементы веб-страницы.

Библиотека Selenium и модуль Webdriver: назначение, возможности, сферы применения. Selenium Webdriver. Имитация действий пользователя средствами Webdriver. Поиск содержимого по тегам в структуре HTML-документа.

Обращение к API сетевых ресурсов. Библиотека requests. Метод GET, метод POST.

Примеры задач, решаемых с помощью веб-скрепинга. Парсинг поисковой выдачи. Сбор рассчитанного времени маршрутов в картографических веб-сервисах. Агрегирование информации о частоте и стоимости перелетов.

Тема 4. Сетевой анализ в социально-экономической географии

Понятие о сетевом анализе. Варианты представления транспортных сетей. Геометрическая сетевая модель. Транспортные сети. Нахождение наилучшего маршрута. Нахождение ближайших обслуживающих объектов. Построение областей обслуживания. Составление матрицы стоимости «источник-назначение». Решение маршрутно-транспортной задачи. Анализ местоположения-распределения. Виды задач местоположение-распределение. Сравнительная характеристика возможностей анализа геометрических и транспортных сетей. Оценка транспортной доступности территории. Классические алгоритмы поиска кратчайших путей. Алгоритм Дейкстры. Эвристические алгоритмы: лучший-первый, A*, M*. Задача коммивояжера: эйлеров граф и гамильтонов граф. Эвристики для задачи коммивояжера. Понятие о кратчайшем и простейшем пути. Индекс простоты. Профиль простоты. Индексы Канского. Число Кёнига. Индекс Шимбела. Стоимость сети. Понятие центральности. Центральность по промежуточности. Центральность по близости. Центральность по гармонике. Центральность по степени.

Тема 5. Использование теории фракталов в социально-экономической географии

Проблема моделирования роста городской застройки: top-down и bottom-up подходы. Новая парадигма — теория сложных динамических самоорганизующихся систем в урбанистике: понятия равновесия, эмерджентности, скейлинга и иерархии. Фрактальность

морфологии как следствие самоорганизации: ограничения евклидовой геометрии, моно- и мультифракталы, фрактальная размерность. Методы определения фрактальной размерности: box-counting, метод дилатации и корреляционный анализ. Интерпретация фрактальной размерности. Обзор задач с применением фрактального анализа

7. Фонд оценочных средств для оценивания результатов обучения по дисциплине:

Домашние задания:

1. Написание демонстрационной программы на Python, с использованием базовых возможностей языка.
2. Написание программы чтения пространственных данных, преобразования системы координат, вычисления статистических характеристик на основе позиционных и семантических данных, визуализация средствами библиотеки Matplotlib и экспорта результата.
3. Создание скрипта, в автоматическом режиме получающего текстовую информацию, путём парсинга содержимого веб-страницы.
4. Создание скрипта для автоматического обращения к картографическим веб-сервисам и получения пространственных данных в структурированной форме.
5. Решение маршрутно-транспортной задачи на примере города. Расчёт центральностей для элементов сети. Моделирование происшествий и анализ изменения центральности по промежуточности.
6. Задание по фракталам.

Примерный перечень вопросов для зачета

1. Типы и структуры данных в Python
2. Переменные в Python. Операции с числами и строками.
3. Условия и циклы в Python
4. Функции в Python. Правила определения и вызова функций
5. Правила оформления кода на языке Python в соответствии с PEP8
6. Работа с файлами и папками средствами языка Python. Библиотеки os, shutil
7. Установка и управление библиотеками языка Python. Дистрибутивы Python. Anaconda. Среды (environments) Anaconda и основные команды управления ими.
8. Работа с табличными данными в Python. Панель. Широкий и длинный формат представления табличных данных. Библиотеки для работы с табличными данными (pandas)
9. Работа с пространственными данными в Python. Библиотеки shapely, pygeos, geopandas, GDAL/OGR
10. Язык разметки HTML. Веб-страница как структурированный текст на языке HTML. Основные теги HTML
11. Selenium WebDriver. Имитация действий пользователя в браузере
12. Веб-скреппинг. Поиск по тегам в HTML документе.
13. Понятие API. Методы GET, POST
14. Понятие фрактала, виды фракталов. Фрактальная размерность.
15. Задачи, решаемые с применением фрактального анализа.

16. Фрактальность морфологии городской застройки. Способы оценки фрактальной размерности при изучении морфологии городской застройки
17. Основные понятия теории графов. Виды графов, элементы графов.
18. Понятие центральности в географических сетях.
19. Эвристические алгоритмы поиска кратчайших путей.
20. Индекс простоты и профиль простоты — возможности интерпретации.
21. Решение задачи коммивояжера.
22. Типовые задачи сетевого анализа: анализ местоположения-распределения.
23. Типовые задачи сетевого анализа: матрица «источник — назначение».
24. Центральность по промежуточности и её применение.
25. Индексы Канского и их интерпретация.
26. Транспортные и геометрические сети: сходства и различия.

Промежуточная аттестация по итогам освоения дисциплины — зачёт в устной форме

Оценка РО и соответствующие виды оценочных средств	Незачет	Зачет
Знания (виды оценочных средств: устный опрос, домашние задания)	Фрагментарные знания или отсутствие знаний	Сформированные систематические знания или общие, но не структурированные знания
Умения (виды оценочных средств: домашние задания)	В целом успешное, но не систематическое умение или отсутствие умений	Успешное и систематическое умение или в целом успешное, но содержащее отдельные пробелы умение (допускает неточности непринципиального характера)
Навыки (владения, опыт деятельности) (виды оценочных средств: домашние задания)	Наличие отдельных навыков или отсутствие навыков	Сформированные навыки (владения), применяемые при решении задач или, в целом, сформированные навыки (владения), но используемые не в активной форме

8. Ресурсное обеспечение:

- Перечень основной и дополнительной учебной литературы

а) основная литература:

Доусон М. Програмируем на Python, СПб.: Питер, 2014. — 416 с.: ил. — 3-е изд.

Лурье И. К. Геоинформационное картографирование. Методы геоинформатики и цифровой обработки космических снимков. — Книжный дом Университет Москва, 2016. — 424 с.

б) дополнительная литература:

Геоинформатика. Толковый словарь основных терминов. Под ред. А.М.Берлянта, А.В.Кошкарёва. – М.: ГИС-Ассоциация, 1999. – 204 с.

Лурье И.К., Косиков А.Г. Теория и практика цифровой обработки изображений. Учебное пособие. - М.: Научный мир, 2003.

Халугин Е.И., Жалковский Е.А., Жданов Н.Д. Цифровые карты. М.: Недра, 1992. - 416 с.

Geoprocessing with Python Chris Garrard, May 2016, Manning, 360 p.

Learning Python, Fifth Edition by Mark Lutz, Published by O'Reilly Media, Inc., 1005

Gravenstein Highway North, Sebastopol, CA 95472, 2013, pp. 1594

Barthelemy M. Morphogenesis of spatial networks, 2018. – 342 p. DOI: <https://doi.org/10.1007/978-3-319-20565-6>.

- Перечень лицензионного программного обеспечения:

Не требуется

- Перечень профессиональных баз данных и информационных справочных систем:

- Менеджер пакетов и библиотек (Anaconda или Miniconda)

- Программа для написания скриптов (Jupyter Notebook)

- Перечень ресурсов информационно-телекоммуникационной сети «Интернет»:

- <https://www.e-education.psu.edu/geog485/node/91> — Веб страница курса GIS Programming and Software Development Penn State University(GEOG 485) – <http://gis-lab.info> сайт сообщества специалистов в области ГИС и ДЗЗ

- <https://docs.python.org/3.5/> — официальная документация языка программирования Python

- <https://stackoverflow.com/> — система вопросов и ответов о программировании

- Описание материально-технической базы

- Учебная аудитория (компьютерный класс) с мультимедийным проектором

9. Язык преподавания: русский.

10. Преподаватель (преподаватели): Ответственный за курс — Елманова Дарья Сергеевна, с.н.с., к.г.н.; преподаватели: Елманова Дарья Сергеевна, с.н.с., к.г.н.; Энтин Андрей Леонидович, с.н.с., к.г.н.; Карпачевский Андрей Михайлович, н.с., к.г.н.

11. Разработчики программы: Елманова Дарья Сергеевна, кафедра социально-экономической географии зарубежных стран, с.н.с., к.г.н.; Энтин Андрей Леонидович, кафедра картографии и геоинформатики, с.н.с., к.г.н.; Карпачевский Андрей Михайлович, кафедра картографии и геоинформатики, н.с., к.г.н.