

Министерство науки и высшего образования Российской Федерации
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ (НИ ТГУ)

Институт биологии, экологии, почвоведения, сельского и лесного хозяйства
(БИОЛОГИЧЕСКИЙ ИНСТИТУТ)



УТВЕРЖДАЮ:
Директор Биологического института

Д.С. Воробьев

« 22 » 06 2023 г.

Рабочая программа дисциплины

Информационные технологии в биологии

по направлению подготовки

06.03.01 Биология

Направленность (профиль) подготовки:
«Биология»

Форма обучения
Очная

Квалификация
Бакалавр

Год приема
2023

Код дисциплины в учебном плане: Б1.В.ДВ.04.01

СОГЛАСОВАНО:
Руководитель ОП
_____ Д.С. Воробьев

Председатель УМК
_____ А.Л. Борисенко

Томск – 2023

1. Цель и планируемые результаты освоения дисциплины (модуля)

Целью освоения дисциплины является формирование следующих компетенций:

– ОПК-6 – способность использовать в профессиональной деятельности основные методы математического анализа и моделирования, теоретических и экспериментальных исследований, приобретать новые математические и естественнонаучные знания, используя современные образовательные и информационные технологии.

Результатами освоения дисциплины являются следующие индикаторы достижения компетенций:

ИОПК-6.2. Применяет методы математического анализа и моделирования, теоретических и экспериментальных исследований

ИОПК-6.3. Приобретает новые математические и естественно-научные знания, используя современные образовательные и информационные технологии

2. Задачи освоения дисциплины

– Освоить аппарат поиска научно-биологической информации, использовать пакеты прикладных компьютерных программ, создавать базы экспериментальных биологических данных, работать с биологической информацией в глобальных компьютерных сетях для обработки, анализа и интерпретации данных, полученных в полевых и лабораторных исследованиях.

– Научиться владеть методами и умело их использовать при выполнении курсовых, выпускных работ и в дальнейшей своей профессиональной деятельности. Понимать проблему взаимосвязи эмпирического и теоретического знания в биологии.

3. Место дисциплины (модуля) в структуре образовательной программы

Дисциплина относится к части образовательной программы, формируемой участниками образовательных отношений, предлагается обучающимся на выбор.

4. Семестр(ы) освоения и форма(ы) промежуточной аттестации по дисциплине

Семестр 5, зачет.

5. Входные требования для освоения дисциплины

Для успешного освоения дисциплины требуются компетенции ИОПК-6.2, ИОПК-7.1; ИОПК-7.2 сформированные в ходе освоения образовательных программ предшествующего уровня образования.

Для успешного освоения дисциплины требуются результаты обучения по следующим дисциплинам: Математика, Информатика, Математические методы в биологии.

6. Язык реализации

Русский

7. Объем дисциплины (модуля)

Общая трудоемкость дисциплины составляет 2 з.е., 72 часов, из которых:

– лекции: 6 ч.;

– семинарские занятия: 0 ч.

– практические занятия: 28 ч.;

– лабораторные работы: 0 ч.

в том числе практическая подготовка: 0 ч.

Объем самостоятельной работы студента определен учебным планом.

8. Содержание дисциплины (модуля), структурированное по темам

Раздел 1. Введение в информационную биологию. Информация и информационные процессы. Информационные технологии.

Тема 1. Предмет, задачи и история развития информационной биологии. Разделы информационной биологии. Взаимодействие со смежными разделами науки и технологий.

Информация и информационные процессы. Теория информации. Единство информационных процессов в природе, обществе и технике. Основные свойства информации. Семиотика, семантика. Информатика, компьютерные науки, компьютеринг. Кибернетика, интеллектуальные системы, экспертные системы, «извлечение данных». Информационные технологии, аппаратные и программные средства информатизации.

Тема 2. Научно-информационная деятельность в биологии и экологии. Документы как материальные носители научной информации.

Теория научной информации, её предметная область (информационная деятельность; документы как материальные носители научной информации; средства обеспечения информационных процессов), задачи (изучение (научно-)информационных процессов; разработка методов анализа семантической информации; разработка методов фиксации научной информации в документальном виде; разработка информационно-поисковых систем, использование электронно-вычислительных машин для обеспечения информационных процессов) и технологические этапы (сбор информации, аналитико-синтетическая обработка, хранение, информационный поиск, распространение научно-технической информации).

Типология научно-технической информации. Информационная структура научного документа: новые знания, релевантная и нерелевантная информация, помехи (искажения), информационный шум.

Первичная научно-техническая информация: публикации, отчетные и другие документы, не имеющие статуса публикаций. Периодические, продолжающиеся и разовые издания, их регистрация. Международные системы регистрации изданий и научных статей. Обмен печатными работами между специалистами, нормативно-правовые ограничения, этика обмена.

Вторичная научно-техническая информация: библиография первой и второй степени. Каталоги и их разновидности (алфавитные, систематические, предметные и др.). Текущая и регистрационная библиография. Информационные издания: сигнальная информация, реферативные журналы и их аналоги на электронных носителях. Реферативные и аналитические обзоры. Кодификация научных публикаций по тематической направленности: система универсальной десятичной кодификации (УДК).

Тема 3. Информационные подходы к анализу биологических и экологических процессов и систем.

Информационные подходы (семантический, биокибернетический, семиотический) к анализу биологических и экологических процессов и систем. Системный подход, редукционизм и холизм. Биологические системы, их особенности. Биологическая информация; наследственная информация, ее защита и восстановление. Нейробиологическая информация, биологические системы управления, гомеостаз. Биоинформатика, геномика, их предмет и перспектива развития. Биоинформационные системы, on-line информационные базы данных. Информационные аспекты изучения структуры и динамики биологического разнообразия. Вычислительная экология, основные направления исследований в этой области. Экологическое моделирование, варианты моделей экологических процессов и систем. Профессии будущего.

Раздел 2. Технологии документального оформления результатов биологических и экологических исследований.

Тема 1. Компьютерные технологии работы с текстовой информацией.

Использование персонального компьютера в научно-исследовательской деятельности для оформления учебной и научной документации, отчетных материалов.

Особенности текстового оформления отдельных видов научных работ, в том числе квалификационных. Текстовые редакторы, основные форматы текстовых файлов, их особенности и применение. Транслитерация текста. Синтаксический контроль, проверка грамматики и орфографии, редактирование и рецензирование. Средства автопоиска и автозамены. Внедрение дополнительных объектов оформления (таблицы, графические изображения и пр.). Редактирование больших документов. Конвертация текстовых файлов из формата в формат. Создание гипертекстовых документов. Работа с файлами формата PDF, создание заметок и комментариев. Интеллектуальные информационные технологии. Технологии распознавания образов: программные продукты, используемые для распознавания текста и особенности их практического использования. Системы электронного перевода текстовой информации и основные возможности их использования в профессиональной деятельности биологов и экологов.

Тема 2. Цитирование научных публикаций.

Менеджмент цитирования. Правила оформления списков информационных источников, требования стандартов и методических указаний ВАК, правила для авторов журнальных научных публикаций. Менеджеры библиографических данных и системы управления библиографической информацией, их возможности и практическое использование: поиск источников, составление «картотеки», цитирование, формирование библиографических списков.

Тема 3. Специализированные программы для работы с библиографией.

Библиографические менеджеры. Reference Management Software: бесплатные (например, Zotero, Mendeley, Qiqa и др.) и проприетарные (например, EndNote, RefWorks и др.), программы и онлайн-сервисы (например, EndNote Web, Citeulike). Сравнительные преимущества и недостатки. Проблема выбора.

Раздел 3. Информационные технологии анализа данных.

Тема 1. Статистические данные.

Дескриптивные и графические методы анализа данных. Статистическое оценивание. Статистическая проверка гипотез. Исследование зависимостей. Методы многомерного статистического анализа. Программное обеспечение анализа данных на персональных компьютерах.

Создание и разработка электронной таблицы.

Тема 2. Назначение электронных таблиц. Основные принципы работы с электронной таблицей MS Excel. Эффективная работа с формулами и функциями. Математические, статистические. Текстовые функции. Логические функции. Функции даты и времени. Основы работы с пакетом анализа «Анализ данных».

Тема 3. Дескриптивные и графические методы анализа данных.

Применение языка программирования для статистического анализа R. Статистическое оценивание. Статистическая проверка гипотез. Исследование зависимостей. Методы многомерного статистического анализа.

Раздел 4. R - язык программирования для статистической обработки данных и работы с графикой

Тема 1. Основы работы. Базовый синтаксис.

Зачем использовать R? Получение и установка R. Работа в R. Рабочее пространство. Пакеты. Установка, загрузка, получение информации о пакете. Пакетная обработка. Работа с большими массивами данных. Вектора, Массивы, Матрицы, Дата фреймы, таблицы.

Тема 2. Функционал для расчета статистик.

Функции, для расчета описательных статистик для не сгруппированных данных.

Функции, для расчета описательных статистик для сгруппированных данных. Написание собственных функций. Функции проверки статистических гипотез. Одновыборочные гипотезы, двухвыборочные гипотезы. Гипотезы множественного сравнения. Дисперсионный анализ. Линейная регрессия и корреляция. Множественная

регрессия и корреляция. Нелинейная регрессии. Логистическая регрессия. Непараметрическая статистика.

Тема 3. Визуализация данных в научных исследованиях и образовании, компьютерные презентации.

Визуализация данных в научных исследованиях и образовании, компьютерные презентации. Программные средства создания диаграмм и графиков, встроенные функции создания графиков и диаграмм стандартных программных пакетов, научная графика и сплайны. Правила оформления графиков и диаграмм, иллюстрирующих научные работы.

Пакеты графического представления информации (ggplot2, lattice).

Создание и демонстрация компьютерных презентаций. Редакторы компьютерных презентаций, принципы и технология создания презентаций. Особенности презентаций, сопровождающих научные доклады и отражающих материалы квалификационных работ.

Тема 4. Создание отчетов.

R Markdown документы. Использование продуктивный notebook interface, чтобы сплести текст и код повествования для получения элегантного форматированного вывода. Использование нескольких языков, включая R, Python и SQL.

Создание интерактивных веб-приложений для визуализации данных с Shiny.

9. Текущий контроль по дисциплине

Текущий контроль по дисциплине проводится путем контроля посещаемости, проведения контрольных работ, тестов по лекционному материалу и фиксируется в форме контрольной точки не менее одного раза в семестр.

10. Порядок проведения и критерии оценивания промежуточной аттестации

Зачет в пятом семестре проводится в ходе выполнения двух практических работ. Продолжительность зачета 1,5 часа.

1. Создание макета Курсовой работы. Реализация основных требований, предъявляемых к курсовым работам.

2. Практическая работа по применению языка программирования R при анализе данных (Статистический анализ и визуализация результатов).

Результаты зачета определяются оценками «зачтено», «незачтено».

«зачтено» ставится, если обучающийся выполняет две работы предлагаемые к зачету работы. При создании макета курсовой работы выполняет все требования, указанные в методических рекомендациях (Оформление работ: методические указания по оформлению курсовых и выпускных квалификационных работ, 2021). При решении задачи по анализу данных выполнены все задания.

«не зачтено» ставится, если обучающийся не выполнил содержащиеся в методических руководствах требования или выполнил задания не в полном объеме.

11. Учебно-методическое обеспечение

а) Электронный учебный курс по дисциплине в электронном университете «Moodle» - <https://moodle.tsu.ru/course/view.php?id=16937>

б) Оценочные материалы текущего контроля и промежуточной аттестации по дисциплине.

в) Методические указания по проведению лабораторных работ.

г) Методические указания по организации самостоятельной работы студентов.

12. Перечень учебной литературы и ресурсов сети Интернет

а) основная литература:

– Кабаков Р.И. R в действии. Анализ и визуализация данных на языке R / Р.И. Кабаков. – М. : ДМК Пресс, 2014. – 580 с.

– Шипунов А.Б. Наглядная статистика. Используем R! / А.Б. Шипунов [и др.] – Москва: ДМК Пресс, 2014. - 298 с.

– Оформление работ : методические указания по оформлению курсовых и выпускных квалификационных работ / Томский государственный университет, Научная библиотека ; составители: Е. Ю. Кичигина, С. М. Григорьевская. – Томск, 2021. – 64 с. https://www.lib.tsu.ru/win/produkcija/metodichka/NB_Metodichka_2021_god.pdf

б) дополнительная литература:

– Zuur A., Ieno E. N., Meesters E. A Beginner's Guide to R. – Springer Science & Business Media, 2009. // 10.1007/978-0-387-93837-0

– Kordon A. Applied Statistics Using SPSS, STATISTICA, MATLAB and R. – Berlin, Heidelberg : Springer-Verlag, [Electronic resource]. – 2007 – URL: <http://dx.doi.org/10.1007/978-3-540-71972-4>. Дата обращения: 29.09.2016.

в) ресурсы сети Интернет:

– Kabacoff Robert I. Quick-R accessing the power of R [Электронный ресурс], 2014. – URL: <http://www.statmethods.net/>

– Çetinkaya-Rundel M. et al. Master Statistics with R / M. Çetinkaya-Rundel, D. Banks, C. Rundel, M.A. Clyde [Электронный ресурс] / Duke university. – Электрон. дан. – URL: <https://www.coursera.org/specializations/statistics>

– Бондарев Т., Федотов П. Программирование на Python [Электронный ресурс] / Институт биоинформатики. – Электрон. дан. – Спб. : Bioinformatics Institute, 2016. – URL: <https://stepik.org/course/Программирование-на-Python-67>

– SpringerLink [Electronic resource] / Springer International Publishing AG, Part of Springer Science+Business Media. – Electronic data. – Cham, Switzerland, [s. n.]. – URL: <http://link.springer.com/>

– Введение в биоинформатику (Introduction to Bioinformatics) [Электронный ресурс] / Санкт-Петербургский государственный университет. – Электрон. дан. – Спб., 2014. – URL: <https://www.coursera.org/learn/bioinformatika/home/welcome>

– Мاستицкий С. R: Анализ и визуализация данных [Электронный ресурс] – Электрон. дан. – 2011-2016 – URL: <http://r-analytics.blogspot.ru>

13. Перечень информационных технологий

а) лицензионное и свободно распространяемое программное обеспечение:

– Microsoft Office Standart 2013 Russian: пакет программ. Включает приложения: MS Office Word, MS Office Excel, MS Office PowerPoint, MS Office On-eNote, MS Office Publisher, MS Outlook, MS Office Web Apps (Word Excel MS PowerPoint Outlook);

– R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.

– публично доступные облачные технологии (Google Docs, Яндекс диск и т.п.).

б) информационные справочные системы:

– Электронный каталог Научной библиотеки ТГУ – <http://chamo.lib.tsu.ru/search/query?locale=ru&theme=system>

– Электронная библиотека (репозиторий) ТГУ – <http://vital.lib.tsu.ru/vital/access/manager/Index>

– ЭБС Лань – <http://e.lanbook.com/>

– ЭБС Консультант студента – <http://www.studentlibrary.ru/>

– Образовательная платформа Юрайт – <https://urait.ru/>

– ЭБС ZNANIUM.com – <https://znanium.com/>

– ЭБС IPRbooks – <http://www.iprbookshop.ru/>

14. Материально-техническое обеспечение

Аудитории для проведения занятий лекционного типа.

Аудитории для проведения занятий семинарского типа, индивидуальных и групповых консультаций, текущего контроля и промежуточной аттестации.

Помещения для самостоятельной работы, оснащенные компьютерной техникой и доступом к сети Интернет, в электронную информационно-образовательную среду и к информационным справочным системам.

15. Информация о разработчиках

Бабкина Ирина Борисовна, канд. биол. наук, кафедра ихтиологии и гидробиологии Биологический институт, доцент