

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

Биологический  
институт

Д.С. Воробьев

24 » марта 20 22 г.

Рабочая программа дисциплины

**Информационные технологии в гидробиологии**

по направлению подготовки

**06.04.01 Биология**

Направленность (профиль) подготовки:

**«Фундаментальная и прикладная биология»**

Форма обучения

**Очная**

Квалификация

**Магистр**

Год приема

**2022**

Код дисциплины в учебном плане: Б1.В.ДВ.02.07

СОГЛАСОВАНО:

Руководитель ОП

Д.С. Воробьев

Председатель УМК

А.Л. Борисенко

Томск – 2022

## **1. Цель и планируемые результаты освоения дисциплины (модуля)**

Целью освоения дисциплины является формирование следующих компетенций:

– ОПК-6 – способность творчески применять и модифицировать современные компьютерные технологии, работать с профессиональными базами данных, профессионально оформлять и представлять результаты новых разработок;

– ПК-1 – способность обрабатывать и использовать научную и научно-техническую информацию при решении исследовательских задач в соответствии с профилем (направленностью) магистерской программы.

Результатами освоения дисциплины являются следующие индикаторы достижения компетенций:

ИОПК-6.1 – Описывает разнообразие, пути и перспективы применения компьютерных технологий в современной биологии;

ИОПК-6.2 – Использует компьютерные технологии и профессиональные базы данных при планировании профессиональной деятельности, обосновывает их выбор;

ИПК-1.2 – Осуществляет поиск, анализ и обобщение научной и научно-технической информации при решении конкретных исследовательских задач.

## **2. Задачи освоения дисциплины**

– Освоить современные программные технологии для хранения, поиска, накопления, структуризации, анализа и визуализации научных данных.

– Научиться применять понятийный аппарат многофакторного анализа данных для решения практических задач профессиональной деятельности.

## **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 1, зачет.

## **5. Входные требования для освоения дисциплины**

Для успешного освоения дисциплины требуются компетенции, сформированные в ходе освоения образовательных программ предшествующего уровня образования.

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

## **6. Язык реализации**

Русский

## **7. Объем дисциплины (модуля)**

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

– лекции: 8 ч.;

– семинарские занятия: 18 ч.

– практические занятия: 0 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. Обзор современных информационных технологий

Тема 1. Базы on-line

Работа с базами данных (FishBase, Catalog of Fishes, GBIF, NCBI и др.)

Тема 2. Язык программирования R как инструмент работы с on-line базами.

Знакомство с R - пакетами для эффективной работы с этими базами (rfishbase, rfisheries, taxon).

Тема 3. Средства для эффективной работы с электронными документами.

Система компьютерной вёрстки TeX / LaTeX. Синтаксис языка разметки. Функции и их основные категории.

Библиографические менеджеры.

Раздел 2. Основные статистики

Тема 1. Базовые статистики

Описательная статистика (Descriptive statistics). Корреляционная матрица. Критерий Стьюдента сравнения средних. Группировка и однофакторная ANOVA.

Тема 2. Основные законы распределения.

Нормальное распределение. Параметры и оценки. Выборочное распределение. Центральная предельная теорема. Понятие и статистические характеристики ряда распределения. Генерация случайных чисел. Метод вероятностной бумаги (Normal Probability Plot).

Тема 3. Статистика выводов.

Проверка гипотез. t-распределение. Применение t-теста для двухвыборочных данных. Применение непараметрического теста для парных данных. Двухвыборочный t-тест. Проверка равенства дисперсии. Применение t-теста для двухвыборочных данных. Применение непараметрического теста для двухвыборочных данных. ANOVA.

Раздел 3. Современные методы визуализации данных

Тема 1. Пакеты прикладных программ.

Графические возможности наиболее популярных пакетов прикладных программ MS Excel, Statistica, SPSS, Past, SSPS и др.

Тема 2. Основы программирования в среде R.

Возможности базового пакета. Продвинутое средства визуализации – пакет ggplot2, lattice. Создание интерактивной графики средствами пакета shiny и plotly.

Раздел 4. Обзор методов многомерной статистики

Тема 1. Линейное многомерное моделирование взаимосвязей.

Линейная регрессионная модель. Описание модуля Multiple Regression.

Тема 2. Классификационный анализ.

Дискриминантный анализ. Кластерный анализ. Деревья классификации и их свойства. Вычислительные методы.

Тема 3. Методы редукции данных.

Факторный анализ (Factor Analysis). Метод главных компонент и классификация (PrincipalComponents & Classification). Многомерный анализ данных.

## **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 первом семестре** проводится в форме аналитической работы, итог аналитический отчет, с использованием современной инфографики. Продолжительность зачета 1,5 часа.

Работа заключается в проведении анализа данных по примерной структуре.

Анализ переменных, формулирование гипотез, проверка данных, выбор адекватных методов анализа, выбор методов визуализации, создание отчета в виде схемы научной публикации.

«зачтено» ставится, если обучающийся даёт полный и правильный ответ, отвечает связно, последовательно, самостоятельно делает выводы либо если в ответе допущены 1–2 неточности, которые учащийся легко исправляет сам или с небольшой помощью преподавателя.

«не зачтено» ставится, если обучающийся обнаруживает незнание большей части материала, не может самостоятельно сделать выводы, речь прерывиста и непоследовательна, допускаются грубые ошибки, которые не исправляются даже с помощью преподавателя.

## **11. Учебно-методическое обеспечение**

а) Электронный учебный курс по дисциплине в электронном университете «Moodle» - <https://moodle.tsu.ru/course/view.php?id=18901>

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

## **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.

Ассанж Дж. Шифропанки. Свобода и будущее Интернета [Электронный ресурс] / Дж. Ассанж, Э. Мюллер-Магун, Дж. Аппельбаум, Ж. Циммерман. – М. : Азбука, 2014. – 224 с.

Беляков Н.С. TEX для всех: оформление учебных и научных работ в системе LATEX / Н.С. Беляков, В.Е. Палаш, П.А. Садовский – Москва : ЛИБРОКОМ, 2012 – 203 с. - URL: <http://sun.tsu.ru/limit/2016/000458943/000458943.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:

– Wickham, H. (2021). Mastering shiny. "O'Reilly Media, Inc.". – URL: <https://mastering-shiny.org/preface.html>

Froese, R. and D. Pauly. Editors. 2022. FishBase. World Wide Web electronic publication. [www.fishbase.org](http://www.fishbase.org), version (02/2022).

Fricke R., Eschmeyer W. N., Van der Laan R. Catalog of fishes: genera, species, references //California Academy of Sciences, San Francisco, CA, USA <http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp>. – 2018.

GBIF. GBIF occurrence download. – 2018.

## **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 OneNote, MS Office

Publisher, MS Outlook, MS Office Web Apps (Word Excel MS PowerPoint Outlook), R, RStudio;

– 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, доцент