

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

Филологический факультет

УТВЕРЖДЕНО:
Декан
И. В. Тубалова

Рабочая программа дисциплины

Большие данные в филологических исследованиях

по направлению подготовки

45.04.01 Филология

Направленность (профиль) подготовки:
Академическая филология: современные исследовательские технологии

Форма обучения
Очная

Квалификация
Магистр

Год приема
2025

СОГЛАСОВАНО:
Руководитель ОП
Н.А. Мишанкина

Председатель УМК
Ю.А. Тихомирова

Томск – 2025

1. Цель и планируемые результаты освоения дисциплины

Целью освоения дисциплины является формирование следующих компетенций:

ОПК-2 Способен использовать в профессиональной деятельности, в том числе педагогической, знания современной научной парадигмы в области филологии и динамики ее развития, системы методологических принципов и методических приемов филологического исследования.

ОПК-3 Способен владеть широким спектром методов и приемов филологической работы с различными типами текстов..

ПК-1 Выполнение отдельных заданий в рамках решения исследовательских задач в сфере филологии под руководством более квалифицированного работника.

Результатами освоения дисциплины являются следующие индикаторы достижения компетенций:

ИОПК-2.1 Демонстрирует знание современной научной парадигмы в области филологии и динамики ее развития, системы методологических принципов и методических приемов филологического исследования.

ИОПК-3.1 Демонстрирует углубленные знания в избранной конкретной области филологии и владение характерным для нее спектром методов и приемов филологической работы с различными типами текстов.

ИПК-1.1 Владеет методами и способами решения научных задач по тематике проводимого исследования, под руководством более квалифицированного работника намечает путь решения исследовательской задачи, методологию и методику исследования.

2. Задачи освоения дисциплины

Знать: основные понятия и источники больших данных в филологии; типы и структуру метаданных; принципы визуализации информации.

Уметь: формулировать исследовательский вопрос, допускающий проверку с помощью анализа метаданных; собирать данные вручную и с помощью простых инструментов; очищать и структурировать данные; создавать базовые визуализации (графики, карты, сети).

Владеть: оффлайн- и онлайн-инструментами для визуализации (VoS, Datawrapper, RawGraphs); методологией интерпретации визуальных данных в филологическом контексте.

3. Место дисциплины в структуре образовательной программы

Дисциплина относится к Блоку 1 «Дисциплины (модули)».

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

Второй семестр, зачет с оценкой

5. Входные требования для освоения дисциплины

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

6. Язык реализации

Русский

7. Объем дисциплины

Общая трудоемкость дисциплины составляет 3 з.е., 108 часов, из которых:

-лекции: 6 ч.

-практические занятия: 12 ч.

в том числе практическая подготовка: 12 ч.

Объем самостоятельной работы студента определен учебным планом.

8. Содержание дисциплины, структурированное по темам

Тема 1: Введение. Что такое Big Data для филолога? От close reading к distant reading. Большие данные как источник новых филологических вопросов. Правильный исследовательский вопрос.

Тема 2: Сбор и "чистка" данных. Источники метаданных. Структура данных: что такое столбцы и строки. Важность единообразия. Создание датасета. Практика "чистки" данных: сортировка, поиск и замена.

Тема 3: Визуализация и анализ. Принципы визуальной аналитики. Типы графиков. Введение в анализ сетей. Визуализация и анализ данных.

Тема 4: Интерпретация и презентация данных. От цифры к смыслу. Как рассказывать историю, основанную на данных? Ограничения метода: что данные НЕ показывают. Риски и этика работы с данными.

9. Текущий контроль по дисциплине

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

Оценочные материалы текущего контроля размещены на сайте ТГУ в разделе «Информация об образовательной программе» - <https://www.tsu.ru/sveden/education/eduop>.

10. Порядок проведения и критерии оценивания промежуточной аттестации

Зачет с оценкой во втором семестре проводится устной форме по билетам. Экзаменационный билет состоит из трех частей. Продолжительность зачета с оценкой 1 час.

Оценочные материалы для проведения промежуточной аттестации размещены на сайте ТГУ в разделе «Информация об образовательной программе» - <https://www.tsu.ru/sveden/education/eduop>.

11. Учебно-методическое обеспечение

а) Электронный учебный курс по дисциплине в электронном университете «Moodle» - <https://moodle.tsu.ru/course/view.php?id=00000>

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

в) План семинарских / практических занятий по дисциплине.

г) Методические указания по проведению лабораторных работ.

д) Методические указания по организации самостоятельной работы студентов.

12. Перечень учебной литературы и ресурсов сети Интернет

а) основная литература:

Журнал "Digital Humanities"

б) дополнительная литература:

- Моррелл К. Визуализация информации. – СПб., 2014.

- Moretti F. Graphs, Maps, Trees: Abstract Models for a Literary History. – London, 2005.

- Бодрийяр Ж. Система вещей. – М., 1995. (для философского осмысления).

в) ресурсы сети Интернет:

- VoS
- Datawrapper
- RawGraphs
- Открытые источники данных: РГБ (lib.ru), Российская Книжная Палата, OpenLibrary.org и др.

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);
- публично доступные облачные технологии (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

Кашпур Валерия Викторовна, канд. филол. наук, доцент кафедры романо-германской и классической филологии