

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«КУБАНСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»
Факультет истории, социологии и международных отношений

УТВЕРЖДАЮ:
Проректор по учебной работе,
качеству образования – первый
проректор

_____ Кагуров, И. В.
подпись
« _____ » _____ 2024 г.



РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)

Б1.В.ДВ.01.02 ИНТЕЛЛЕКТУАЛЬНЫЙ АНАЛИЗ ТЕКСТА В СОЦИОЛОГИЧЕСКИХ ИССЛЕДОВАНИЯХ

(код и наименование дисциплины в соответствии с учебным планом)

Направление подготовки/специальность _____ 39.03.01 Социология
(код и наименование направления подготовки/специальности)

Направленность (профиль) / специализация

_____ Социальная теория и прикладное социальное знание
(наименование направленности (профиля) / специализации)

Форма обучения _____ очная
(очная, очно-заочная, заочная)

Квалификация _____ бакалавр

Краснодар 2024

Рабочая программа дисциплины Б1.В.ДВ.01.02 Интеллектуальный анализ текста в социологических исследованиях составлена в соответствии с федеральным государственным образовательным стандартом высшего образования (ФГОС ВО) по направлению подготовки / специальности 39.03.01 Социология
код и наименование направления подготовки

Программу составил:

Т.А. Рунаев, доцент,
канд. социол. наук

И.О. Фамилия, должность, ученая степень, ученое звание



подпись

Рабочая программа дисциплины Б1.В.ДВ.01.02 Интеллектуальный анализ текста в социологических исследованиях утверждена на заседании кафедры (разработчика) социологии
протокол № 10 «окт» 03 2024 г.

Заведующий кафедрой (разработчика) Т.А. Хагуров

фамилия, инициалы



подпись

Рабочая программа обсуждена на заседании кафедры (выпускающей)
протокол № 10 «окт» 03 2024 г.

Заведующий кафедрой (выпускающей) Хагуров Т.А.

фамилия, инициалы



подпись

Утверждена на заседании учебно-методической комиссии факультета истории, социологии и международных отношений
протокол № 6 «15» март 2024 г.

Председатель УМК факультета Э.Г. Вартамян

фамилия, инициалы



подпись

Рецензенты:

Михайлов А.П. профессор кафедры социологии и философии ФГБОУ ВО «Адыгейский государственный университет», доктор социологических наук, профессор

Муха Виктория Николаевна

директор Центра социологических исследований, доцент кафедры социологии, правопедения и работы с персоналом ФГБОУ ВО «Кубанский государственный технологический университет», кандидат социологических наук

1 Цели и задачи изучения дисциплины (модуля).

1.1 Цель освоения дисциплины.

Формирование у студентов представлений об интеллектуальном анализе текста как цифровом инструменте работы с большими массивами текстовых данных, отражающих явления и процессы современного общества.

1.2 Задачи дисциплины.

Для достижения цели в ходе учебного процесса предполагается решить следующие задачи:

- познакомить студентов с возможностями современных цифровых технологий, позволяющих проводить анализ текстовой информации;
- выработать у обучающихся навыки сбора, систематизации и обработки массивов текстов, содержащих социально значимую информацию;
- развить способность проектирования индивидуального и (или) группового исследования с применением анализа текстовых массивов данных.

1.3 Место дисциплины (модуля) в структуре образовательной программы.

Дисциплина «Интеллектуальный анализ текста в социологических исследованиях» относится к дисциплинам Блока 1 «Часть, формируемая участниками образовательных отношений» (Дисциплины (модули) по выбору) учебного плана. Дисциплина находится в логической и содержательно-методической взаимосвязи с другими частями ООП и базируется на знаниях, полученных при изучении дисциплин общенаучного и профессионального цикла, таких как: «Теория измерений и анализ данных в социологии», «Современные компьютерные технологии в анализе социологической информации», «Методология и методы социологического исследования».

1.4 Перечень планируемых результатов обучения по дисциплине (модулю), соотнесенных с планируемыми результатами освоения образовательной программы.

Изучение данной учебной дисциплины направлено на формирование у обучающихся профессиональных компетенций:

| Код и наименование индикатора | Результаты обучения по дисциплине |
|---|--|
| ОПК-3 Способен принимать участие в социологическом исследовании на всех этапах его проведения | |
| ОПК-3.1 Операционализирует задачи конкретного социологического исследования. | <i>знает</i> цифровой методологический инструментарий, предназначенный для проведения интеллектуального анализа текста в социологических исследованиях. |
| | <i>умеет</i> определять задачи, на решение которых направлены цифровые инструменты, предназначенные для интеллектуального анализа текста в социологических исследованиях. |
| | <i>владеет</i> навыками формирования целей и задач социологического исследования, основанного на интеллектуальном анализе текста. |
| ОПК-3.2 Предлагает пути проверки задач и гипотез исследования. | <i>знает</i> совокупность цифровых инструментов интеллектуального анализа текстов (приложений компьютера, программных алгоритмов), позволяющих решать задачи и проверять исследовательские гипотезы. |
| | <i>умеет</i> использовать цифровые инструменты интеллектуального анализа текстов (приложения компьютера, программные алгоритмы), направленные на решение задач и проверку исследовательских |

| | |
|--|--|
| | гипотез. |
| | <i>владеет</i> навыками применения цифровых инструментов интеллектуального анализа текстов (приложений компьютера, программных алгоритмов) для верификации эмпирических данных и проверки гипотез социологического исследования. |
| ОПК-3.3 Разрабатывает программные и методические документы социологического исследования. | <i>знает</i> основные этапы интеллектуального анализа текстов в виде препроцессинга, базовых методов, визуализации и постпроцессинга. |
| | <i>умеет</i> формировать программу социологического исследования в соответствии с этапами интеллектуального анализа текста. |
| | <i>владеет</i> навыками составления методологического инструментария с применением методов интеллектуального анализа текста. |
| ОПК-3.4 Решает организационные и методические вопросы сбора информации в соответствии с поставленными задачами и методической стратегией исследования; контролирует сбор социологических данных. | <i>знает</i> процедуру использования текстовой информации с применением цифровых инструментов (приложений компьютера, программных алгоритмов). |
| | <i>умеет</i> составлять программные коды для сбора данных, необходимых для применения интеллектуального анализа текста. |
| | <i>владеет</i> инструментами веб-скрайпинга и парсинга, направленного на изъятие и систематизацию текстовой информации из сети «Интернет». |
| ОПК-3.5 Оформляет научно-техническую документацию на всех этапах исследования. | <i>знает</i> особенности представления результатов интеллектуального анализа текста в соответствии с техническим заданием. |
| | <i>умеет</i> составлять аналитические справки и аналитические отчеты по результатам интеллектуального анализа текстов в социологических исследованиях. |
| | <i>владеет</i> навыками визуализации научно-исследовательских результатов в аналитической справке и (или) аналитическом отчете. |
| ОПК-4 Способен выявлять социально значимые проблемы и определять пути их решения на основе теоретических знаний и результатов социологических исследований | |
| ОПК-4.1 Демонстрирует возможности использования теоретических знаний и результатов социологических исследований для выявления социально значимых проблем. | <i>знает</i> тенденции в современных социологических исследованиях по интеллектуальному анализу текста |
| | <i>умеет</i> использовать виды текстового анализа (тематического моделирования, анализа тональности, латентно-семантического анализа) для определения социально значимых проблем. |
| | <i>владеет</i> способами презентации результатов обработанных текстовых данных с помощью цифровых инструментов. |
| ОПК-4.2 Выявляет социально значимые проблемы при использовании описательных, объяснительных и | <i>знает</i> основы тематического моделирования в рамках компьютерных приложений, предназначенных для осуществления интеллектуального анализа текста. |

| | |
|--|--|
| прогнозных моделей социальных явлений и процессов. | <i>умеет</i> применять историко-теоретический аппарат социологической науки для описания, объяснения и прогнозирования полученных результатов. |
| | <i>владеет</i> навыками выявления наиболее репрезентативных моделей интеллектуального анализа текстов, отражающих социальные феномены и явления. |
| ОПК-4.3 Формулирует задачи исследований для определения путей решения социально значимых проблем на основе теоретических знаний и результатов конкретных социологических исследований. | <i>знает</i> возможности применения интеллектуального анализа текста для решения прикладных задач (в маркетинговых исследованиях, в мониторинге СМИ и СМК и т.д.). |
| | <i>умеет</i> определять задачи, требуемые для применения интеллектуального анализа текста в рамках прикладных исследований. |
| | <i>владеет</i> цифровыми инструментами применения интеллектуального анализа текста для решения прикладных задач |

2. Структура и содержание дисциплины.

2.1 Распределение трудоёмкости дисциплины по видам работ.

Общая трудоёмкость дисциплины составляет 3 зач.ед. (108 часов), их распределение по видам работ представлено в таблице (для студентов ОФО)

| Вид учебной работы | Всего часов | Семестры (часы) | | | | |
|--|--------------------------------------|-----------------|-------------|---|---|---|
| | | 6 | | | | |
| Контактная работа, в том числе: | 68,2 | 68,2 | | | | |
| Аудиторные занятия (всего): | 64 | 64 | | | | |
| Занятия лекционного типа | 32 | 32 | - | - | - | |
| Лабораторные занятия | - | - | - | - | - | |
| Занятия семинарского типа (семинары, практические занятия) | 32 | 32 | - | - | - | |
| Иная контактная работа: | | | | | | |
| Контроль самостоятельной работы (КСР) | 4 | 4 | | | | |
| Промежуточная аттестация (ИКР) | 0,2 | 0,2 | | | | |
| Самостоятельная работа, в том числе: | 39,8 | 39,8 | | | | |
| Курсовая работа | | - | - | - | - | |
| Проработка учебного (теоретического) материала | 14 | 14 | - | - | - | |
| Выполнение индивидуальных заданий (подготовка проектов, презентаций) | 14 | 14 | - | - | - | |
| Реферат | 11,8 | 11,8 | - | - | - | |
| Подготовка к текущему контролю | - | - | - | - | - | |
| Контроль: | | | | | | |
| Подготовка к экзамену | | - | - | - | - | |
| Общая трудоемкость | час. | 108 | 108 | - | - | - |
| | в том числе контактная работа | 68,2 | 68,2 | | | |
| | зач.ед | 3 | 3 | | | |

2.2 Структура дисциплины:

Распределение видов учебной работы и их трудоемкости по разделам дисциплины.
Разделы дисциплины, изучаемые в 6 семестре (очная форма)

| № | Наименование разделов | Количество часов | | | | |
|----|---|------------------|-------------------|----|----|----------------------|
| | | Всего | Аудиторная работа | | | Внеаудиторная работа |
| | | | Л | ПЗ | ЛР | СРС |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 1. | Введение в интеллектуальный анализ текста | 10 | 4 | 4 | - | 2 |
| 2. | Основы работы с языком программирования R | 10 | 4 | 4 | - | 2 |
| 3. | Веб-скрайпинг | 12 | 4 | 4 | - | 4 |
| 4. | Преппроцессинг | 12 | 4 | 4 | - | 4 |
| 5. | Метод «мешок слов» | 12 | 4 | 4 | - | 4 |
| 6. | Анализ тональности | 12 | 4 | 4 | - | 4 |
| 7. | Тематическое моделирование | 12 | 4 | 4 | | 4 |
| 8. | Методы классификации текстов | 12 | 4 | 4 | | 4 |
| | Контроль самостоятельной работы (КСР) | 4 | - | - | - | - |
| | Промежуточная аттестация (ИКР) | 0,2 | - | - | - | - |
| | Реферат | 11,8 | - | - | - | - |
| | <i>Итого по дисциплине:</i> | 108 | 32 | 32 | - | 28 |

Примечание: Л - лекции, ПЗ - практические занятия / семинары, ЛР - лабораторные занятия, СРС - самостоятельная работа студента

2.3 Содержание разделов дисциплины:

2.3.1 Занятия лекционного типа.

| № | Наименование раздела | Содержание раздела | Форма текущего контроля |
|----|---|---|-------------------------|
| 1 | 2 | 3 | 4 |
| 1. | Введение в интеллектуальный анализ текста | Введение: текст как объект познания социального мира. Контент-анализ: история, виды и возможности. Нарративный анализ. Дискурс-анализ: теория дискурса Лакло–Муфф, критический дискурс анализ (CDA) Н. Фэркло, дискурс-исторический подход (DHA) Р. Водак, дискурсивный анализ и социология знания (SKAD) Р. Келлера. Потребность в интеллектуальном анализе текстов в эпоху «Больших данных». Алгоритм интеллектуального анализа текста: постпроцессинг, основные операции, представление результатов, постпроцессинг. Спектр программного обеспечения для интеллектуального анализа текста. | Опрос |

| | | | |
|----|---|---|-------|
| 2. | Основы работы с языком программирования R | Алгоритм инсталлирования языка программирования R и приложения RStudio. Знакомство с интерфейсом приложения RStudio: панель скрипта, панель консоли, панель среды, вспомогательная панель. Базовое кодирование на языке программирования R: имена переменных, присвоение, векторы, матрицы, списки, фреймы данных. Пакеты в RStudio. Обзор пакетов интеллектуального анализа текстов. | Опрос |
| 3. | Веб-скрайпинг | Импорт текстов в RStudio. Веб-скрайпинг HTML. Создание циклов. Работа с изъятием текстов Telegram-каналов. Парсинг постов с метаданными из пабликов «ВКонтакте». Работа с API VK. Работа с API YouTube. Оптическое распознавание текстов (OCR). | Опрос |
| 4. | Препроцессинг | Составление и очистка корпуса документов. Ликвидация «шума» в текстах. Нормализация текста: стемминг и лемматизация. Особенности препроцессинга текстов интернет-коммуникации. | Опрос |
| 5. | Метод «мешок слов» | Введение в метод «мешка слов». Составление матрицы «термин – документ» (TDM). Составление матрицы «документ – термин» (DTM). Предельный уровень разреженности терминов. Виды взвешивания: term frequency и TF-IDF. Поиск корреляций терминов. | Опрос |
| 6. | Анализ тональности | Описание метода. Характеристика русскоязычных словарей тональностей (SentiMental, RuSentiLex, LinisCrowd, EmoLex). Способы построения авторских словарей тональностей. Алгоритм выявления тональности текста. Поиск среднего значения тональности в корпусе документов. Контент-анализ с помощью авторского словаря. | Опрос |
| 7. | Тематическое моделирование | Описание методов тематического моделирования. Анализ совпадений (co-occurrence): построение сетей терминов. Латентное размещение Дирихле (LDA): история метода и возможности использования в социологических исследованиях. Определение оптимального значения тем для LDA-модели. Коррелированное тематическое моделирование (CTM). Латентно-семантический анализ (LSA). | Опрос |
| 8. | Методы классификации текстов | Латентно-семантический анализ (LSA). История и особенности LSA. Расчет косинусного сходства словарного состава документов. Визуализация результатов в виде теплокарты. Кластерный анализ документов. Виды метрик кластерного анализа. Иерархическая кластеризация. Метод k-means. | Опрос |

2.3.2 Занятия семинарского типа.

| № | Наименование раздела | Содержание раздела | Форма текущего контроля |
|----|---|---|--------------------------------|
| 1 | 2 | 3 | 4 |
| 1. | Введение в интеллектуальный анализ текста | Введение: текст как объект познания социального мира. Контент-анализ: история, виды и возможности. Нарративный анализ. Дискурс-анализ: теория дискурса Лакло–Муфф, критический дискурс анализ (CDA) Н. Фэркло, дискурс-исторический подход (DHA) Р. Водак, дискурсивный анализ и социология знания (SKAD) Р. Келлера. Потребность в интеллектуальном анализе текстов в эпоху «Больших данных». Алгоритм интеллектуального анализа текста: постпроцессинг, основные операции, представление результатов, постпроцессинг. Спектр программного обеспечения для интеллектуального анализа текста. | Опрос, тест |
| 2. | Основы работы с языком программирования R | Алгоритм инсталлирования языка программирования R и приложения RStudio. Знакомство с интерфейсом приложения RStudio: панель скрипта, панель консоли, панель среды, вспомогательная панель. Базовое кодирование на языке программирования R: имена переменных, присвоение, векторы, матрицы, списки, фреймы данных. Пакеты в RStudio. Обзор пакетов интеллектуального анализа текстов. | Опрос, тест, проектное задание |
| 3. | Веб-скрайпинг | Импорт текстов в RStudio. Веб-скрайпинг HTML. Создание циклов. Работа с изъятием текстов Telegram-каналов. Парсинг постов с метаданными из пабликов «ВКонтакте». Работа с API VK. Работа с API YouTube. Оптическое распознавание текстов (OCR). | Опрос, тест, проектное задание |
| 4. | Препроцессинг | Составление и очистка корпуса документов. Ликвидация «шума» в текстах. Нормализация текста: стемминг и лемматизация. Особенности препроцессинга текстов интернет-коммуникации. | Опрос, тест, проектное задание |
| 5. | Метод «мешок слов» | Введение в метод «мешка слов». Составление матрицы «термин – документ» (TDM). Составление матрицы «документ – термин» (DTM). Предельный уровень разреженности терминов. Виды взвешивания: term frequency и TF-IDF. Поиск корреляций терминов. | Опрос, тест, проектное задание |
| 6. | Анализ тональности | Описание метода. Характеристика русскоязычных словарей тональностей (SentiMental, RuSentiLex, LinisCrowd, EmoLex). Способы построения авторских словарей тональностей. Алгоритм выявления тональности текста. Поиск среднего значения тональности в корпусе документов. Контент-анализ с помощью авторского словаря. | Опрос, тест, проектное задание |

| | | | |
|----|------------------------------|--|--------------------------------|
| 7. | Тематическое моделирование | Описание методов тематического моделирования. Анализ совпадений (co-occurrence): построение сетей терминов. Латентное размещение Дирихле (LDA): история метода и возможности использования в социологических исследованиях. Определение оптимального значения тем для LDA-модели. Коррелированное тематическое моделирование (CTM). Латентно-семантический анализ (LSA). | Опрос, тест, проектное задание |
| 8. | Методы классификации текстов | Латентно-семантический анализ (LSA). История и особенности LSA. Расчет косинусного сходства словарного состава документов. Визуализация результатов в виде теплокарты. Кластерный анализ документов. Виды метрик кластерного анализа. Иерархическая кластеризация. Метод k-means. | Опрос, тест, проектное задание |

2.3.3 Лабораторные занятия.

Лабораторные занятия - не предусмотрены

2.3.4 Примерная тематика курсовых работ (проектов)

Курсовые работы (проекты) - не предусмотрены

2.4 Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине (модулю)

Учебно-методические материалы для самостоятельной работы обучающихся из числа инвалидов и лиц с ограниченными возможностями здоровья (ОВЗ) предоставляются в формах, адаптированных к ограничениям их здоровья и восприятия информации:

Для лиц с нарушениями зрения:

- в печатной форме увеличенным шрифтом,
- в форме электронного документа.

Для лиц с нарушениями слуха:

- в печатной форме,
- в форме электронного документа.

Для лиц с нарушениями опорно-двигательного аппарата:

- в печатной форме,
- в форме электронного документа.

Данный перечень может быть конкретизирован в зависимости от контингента обучающихся.

3. Образовательные технологии.

При реализации различных видов учебной работы по дисциплине «Интеллектуальный анализ текста в социологических исследованиях», используются следующие образовательные технологии: активные и интерактивные формы проведения занятий - интерактивные и проблемные лекции; опросы; самостоятельная работа - работа с публикациями в предметной области дисциплины; подготовка рефератов, выполнение практических занятий.

Для лиц с ограниченными возможностями здоровья предусмотрена организация консультаций с использованием электронной почты.

4. Оценочные средства для текущего контроля успеваемости и промежуточной аттестации.

Оценочные средства предназначены для контроля и оценки образовательных достижений

обучающихся, освоивших программу учебной дисциплины «Интеллектуальный анализ текста в социологических исследованиях».

Оценочные средства включает контрольные материалы для проведения текущего контроля в форме тестовых заданий, доклада-презентации по проблемным вопросам, разноуровневых заданий, ролевой игры, ситуационных задач и промежуточной аттестации в форме вопросов и заданий к зачету

Структура оценочных средств для текущей и промежуточной аттестации

| № п/п | Код и наименование индикатора | Результат обучения | Наименование оценочного средства | |
|-------|---|--|---|---|
| | | | Текущий контроль | Промежуточная аттестация |
| 1 | ОПК-3.1 Операционализирует задачи конкретного социологического исследования. | <i>знает</i> цифровой методологический инструментарий, предназначенный для проведения интеллектуального анализа текста в социологических исследованиях. | контрольный опрос (КО); тестирование (Т) | Вопрос на зачете, Итоговое тестирование |
| | | <i>умеет</i> определять задачи, на решение которых направлены цифровые инструменты, предназначенные для интеллектуального анализа текста в социологических исследованиях. | разработка проекта (РП) | Вопрос на зачете, Итоговое проектное задание |
| | | <i>владеет</i> навыками формирования целей и задач социологического исследования, основанного на интеллектуальном анализе текста. | разработка проекта (РП) | Вопрос на зачете, Итоговое проектное задание |
| 2 | ОПК-3.2 Предлагает пути проверки задач и гипотез исследования. | <i>знает</i> совокупность цифровых инструментов интеллектуального анализа текстов (приложений компьютера, программных алгоритмов), позволяющих решать задачи и проверять исследовательские гипотезы. | контрольный опрос (КО); тестирование (Т) | Вопрос на зачете |
| | | <i>умеет</i> использовать цифровые инструменты интеллектуального анализа текстов (приложения компьютера, программные алгоритмы), | разработка проекта (РП) | Вопрос на зачете, Итоговое проектное задание |

| | | | | |
|---|--|--|--|--|
| | | направленные на решение задач и проверку исследовательских гипотез. | | |
| | | <i>владеет</i> навыками применения цифровых инструментов интеллектуального анализа текстов (приложений компьютера, программных алгоритмов) для верификации эмпирических данных и проверки гипотез социологического исследования. | разработка проекта (РП) | Вопрос на зачете, Итоговое проектное задание |
| 3 | ОПК-3.3 Разрабатывает программные и методические документы социологического исследования. | <i>знает</i> основные этапы интеллектуального анализа текстов в виде препроцессинга, базовых методов, визуализации и постпроцессинга. | контрольный опрос (КО); тестирование (Т) | Вопрос на зачете, Итоговое тестирование |
| | | <i>умеет</i> формировать программу социологического исследования в соответствии с этапами интеллектуального анализа текста. | разработка проекта (РП) | Вопрос на зачете, Итоговое проектное задание |
| | | <i>владеет</i> навыками составления методологического инструментария с применением методов интеллектуального анализа текста. | разработка проекта (РП) | Вопрос на зачете, Итоговое проектное задание |
| 4 | ОПК-3.4 Решает организационные и методические вопросы сбора информации в соответствии с поставленными задачами и методической стратегией исследования; контролирует сбор социологических данных. | <i>знает</i> процедуру использования текстовой информации с применением цифровых инструментов (приложений компьютера, программных алгоритмов). | контрольный опрос (КО); тестирование (Т) | Вопрос на зачете, Итоговое тестирование |
| | | <i>умеет</i> составлять программные коды для сбора данных, необходимых для применения интеллектуального анализа текста. | разработка проекта (РП) | Вопрос на зачете, Итоговое проектное задание |

| | | | | |
|---|---|---|--|--|
| | | | | |
| | | <i>владеет</i> инструментами веб-скрайпинга и парсинга, направленного на изъятие и систематизацию текстовой информации из сети «Интернет». | разработка проекта (РП) | Вопрос на зачете, Итоговое проектное задание |
| 5 | ОПК-3.5 Оформляет научно-техническую документацию на всех этапах исследования. | <i>знает</i> особенности представления результатов интеллектуального анализа текста в соответствии с техническим заданием. | контрольный опрос (КО); тестирование (Т) | Вопрос на зачете, Итоговое тестирование |
| | | <i>умеет</i> составлять аналитические справки и аналитические отчеты по результатам интеллектуального анализа текстов в социологических исследованиях. | разработка проекта (РП) | Вопрос на зачете, Итоговое проектное задание |
| | | <i>владеет</i> навыками визуализации научно-исследовательских результатов в аналитической справке и (или) аналитическом отчете. | разработка проекта (РП) | Вопрос на зачете, Итоговое проектное задание |
| 6 | ОПК-4.1 Демонстрирует возможности использования теоретических знаний и результатов социологических исследований для выявления социально значимых проблем. | <i>знает</i> тенденции в современных социологических исследованиях по интеллектуальному анализу текста | контрольный опрос (КО); тестирование (Т) | Вопрос на зачете, Итоговое тестирование |
| | | <i>умеет</i> использовать виды текстового анализа (тематического моделирования, анализа тональности, латентно-семантического анализа) для определения социально значимых проблем. | разработка проекта (РП) | Вопрос на зачете, Итоговое проектное задание |
| | | <i>владеет</i> способами презентации результатов обработанных текстовых данных с помощью цифровых | разработка проекта (РП) | Вопрос на зачете, Итоговое проектное задание |

| | | | | |
|---|--|--|--|--|
| | | инструментов. | | |
| 7 | ОПК-4.2 Выявляет социально значимые проблемы при использовании описательных, объяснительных и прогнозных моделей социальных явлений и процессов. | <i>знает</i> основы тематического моделирования в рамках компьютерных приложений, предназначенных для осуществления интеллектуального анализа текста. | контрольный опрос (КО); тестирование (Т) | Вопрос на зачете, Итоговое тестирование |
| | | <i>умеет</i> применять историко-теоретический аппарат социологической науки для описания, объяснения и прогнозирования полученных результатов. | разработка проекта (РП) | Вопрос на зачете, Итоговое проектное задание |
| | | <i>владеет</i> навыками выявления наиболее репрезентативных моделей интеллектуального анализа текстов, отражающих социальные феномены и явления. | разработка проекта (РП) | Вопрос на зачете, Итоговое проектное задание |
| 8 | ОПК-4.3 Формулирует задачи исследований для определения путей решения социально значимых проблем на основе теоретических знаний и результатов конкретных социологических исследований. | <i>знает</i> возможности применения интеллектуального анализа текста для решения прикладных задач (в маркетинговых исследованиях, в мониторинге СМИ и СМК и т.д.). | контрольный опрос (КО); тестирование (Т) | Вопрос на зачете, Итоговое тестирование |
| | | <i>умеет</i> определять задачи, требуемые для применения интеллектуального анализа текста в рамках прикладных исследованиях. | разработка проекта (РП) | Вопрос на зачете |
| | | <i>владеет</i> цифровыми инструментами применения интеллектуального анализа текста для решения прикладных задач | разработка проекта (РП) | Вопрос на зачете Итоговое проектное задание |

4.1 Фонд оценочных средств для проведения промежуточной аттестации.

Примерные вопросы к итоговому тесту

Вопрос 1. К методу тематического моделирования относится:

- 1) Латентное размещение Дирихле
- 2) «мешок слов»
- 3) поиск корреляций слов
- 4) иерархическая кластеризация

Вопрос 2. Последовательность элементов одного типа на языке R:

- 1) вектор
- 2) матрица
- 3) список
- 4) фрейм данных

Вопрос 3. Веб-скрайпинг – это:

- 1) метод тематического моделирования
- 2) автоматический или полуавтоматический процесс извлечения текстовой информации из веб-страниц и ее сохранения
- 3) инструмент предварительной обработки текста в интернете
- 4) совокупность действий по составлению словаря тональности на основе интернет-коммуникации

Примерное проектное задание

Задание 1. Сделайте веб-скрапинг последних 10 новостей из рубрики «Наука» на сайте Russia Today (<https://russian.rt.com/science>). Сохраните результат в переменной science в формате RData.

Задание 2. Проведите веб-скрайпинг свежих 200 постов на официальной странице информационного агентства «РИА Новости» в социальной интернет-сети «ВКонтакте» (<https://vk.com/ria>). Полученный результат сохраните в переменной ria в формате RData.

Задание 3. Осуществите веб-скрайпинг новостей Telegram-канала «КубГУ | kubsunews» (<https://t.me/kubsunews>) за последнюю неделю. Полученный результат сохраните в переменной kubsu в формате RData.

Вопросы на зачет

1. Контент-анализ: история метода, виды и способы проведения.
2. Нарративный анализ: история метода, виды и способы проведения.
3. Дискурс-анализ: история метода, виды и способы проведения.
4. Интеллектуальный анализ текста: понятие, алгоритм и программные решения.
5. Базовое кодирование в RStudio: имена переменных, векторы, матрицы, списки, фреймы данных.
6. Пакеты для интеллектуального анализа текста в RStudio.
7. Особенности препроцессинга.
8. Метод «мешка слов» («bag of words»).
9. Поиск ассоциаций терминов.
10. Анализ совпадений (co-occurrence): построение сетей терминов.
11. Латентное размещение Дирихле (LDA).
12. Коррелированное тематическое моделирование (CTM).
13. Латентно-семантический анализ (LSA).
14. Характеристика русскоязычных словарей тональностей.
15. Алгоритм построения авторского словаря тональностей.
16. Постпроцессинг.

Критерии оценивания по зачету:

«зачтено»: студент владеет теоретическими знаниями по данному разделу, знает формы и специфику работы с языком программирования R, допускает незначительные ошибки; студент умеет правильно объяснять теоретический материал, иллюстрируя его примерами различных социальных ситуаций из жизни коллектива и / или организации.

«не зачтено»: материал не усвоен или усвоен частично, студент затрудняется привести примеры по использованию языка программирования R, имеет довольно ограниченный объем знаний программного теоретического материала.

Оценочные средства для инвалидов и лиц с ограниченными возможностями здоровья выбираются с учетом их индивидуальных психофизических особенностей.

- при необходимости инвалидам и лицам с ограниченными возможностями здоровья предоставляется дополнительное время для подготовки ответа на экзамене;

- при проведении процедуры оценивания результатов обучения инвалидов и лиц с ограниченными возможностями здоровья предусматривается использование технических средств, необходимых им в связи с их индивидуальными особенностями;

- при необходимости для обучающихся с ограниченными возможностями здоровья и инвалидов процедура оценивания результатов обучения по дисциплине может проводиться в несколько этапов.

Процедура оценивания результатов обучения инвалидов и лиц с ограниченными возможностями здоровья по дисциплине (модулю) предусматривает предоставление информации в формах, адаптированных к ограничениям их здоровья и восприятия информации:

Для лиц с нарушениями зрения:

- в печатной форме увеличенным шрифтом,
- в форме электронного документа.

Для лиц с нарушениями слуха:

- в печатной форме,
- в форме электронного документа.

Для лиц с нарушениями опорно-двигательного аппарата:

- в печатной форме,
- в форме электронного документа.

Данный перечень может быть конкретизирован в зависимости от контингента обучающихся.

5. Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины (модуля).

5.1 Основная литература:

1. Рунаев Т.А. Интеллектуальный анализ текста в социальных науках : учебное пособие / Т. А. Рунаев ; Министерство науки и высшего образования Российской Федерации, Кубанский государственный университет. - Краснодар : Кубанский государственный университет, 2024. - 127 с. : ил. - Библиогр.: с. 125. - ISBN 978-5-8209-2404-0. - Режим доступа: <http://212.192.134.46/MegaPro/Web/SearchResult/ToPage/1>.
2. Тарасов М.И. Теория текста и дискурса. Дискурс рассуждения : учебник и практикум для вузов / М.И. Тарасов. — Москва : Издательство Юрайт, 2023. — 284 с. — (Высшее образование). — ISBN 978-5-534-14559-5. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — Режим доступа: <https://urait.ru/bcode/519926>.
3. Миркин Б.Г. Введение в анализ данных : учебник и практикум / Б.Г. Миркин. — Москва : Издательство Юрайт, 2023. — 174 с. — (Высшее образование). — ISBN 978-5-9916-5009-0. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — Режим доступа: <https://urait.ru/bcode/511121>.

5.2 Дополнительная литература:

1. Практики анализа качественных данных в социальных науках: учебное пособие. – Москва: Изд. дом Высшей школы экономики, 2023. – 383 с. – ISBN: 978-5-7598-2542-5.
2. Йоргенсен М.В., Филлипс Л.Дж. Дискурс-анализ. Теория и метод. – Харьков: Издательство «Гуманитарный Центр», 2008. – 352 с. – ISBN 966-8324-06-4.
3. Храмов Д.А. Сбор данных в Интернете на языке R. – Москва: ДМК Пресс, 2017. – 280 с. – ISBN: 978-5-97060-459-5.
4. Уикэм Х., Гроулмунд Г. Язык R в задачах науки о данных: импорт, подготовка, обработка, визуализация и моделирование данных. – Санкт-Петербург: ООО «Диалектика», 2018. – 592 с. – ISBN: 978-5-9909446-8-8.
5. Kwartler T. Text Mining in Practice with R. – Hoboken: John Wiley & Sons, 2017. – 307 p. – ISBN: 9781119282082.
6. Silge J., Robinson D. Text mining with R. A Tidy Approach. Sebastopol: O'Reilly Media, 2017. – 184 p. – 978-1-491-98165-8.
7. Wiedemann G. Text Mining for Qualitative Data Analysis in the Social Sciences. A Study on Democratic Discourse in Germany. Wiesbaden: Springer, 2016. – 305 p. – ISBN 978-3-658-15308-3.

7.3 Периодические издания:

- 1) Вестник СПбГУ. Серия: Психология, социология, педагогика
- 2) Вестник МГУ. Серия: Социология и политология
- 3) Журнал практического психолога
- 4) Журналист. Социальные коммуникации
- 5) Общественные науки и современность
- 6) Социально-гуманитарные знания
- 7) СОЦИС / Социологические исследования

6 Перечень ресурсов информационно-телекоммуникационной сети «интернет», необходимых для освоения дисциплины (модуля).

<http://lib.socio.msu.ru/l/library> - Электронная библиотека социологического факультета МГУ имени М.В. Ломоносова

http://window.edu.ru/window_catalog - Единое окно доступа к образовательным ресурсам

<http://www.hh.ru> - Хэд Хантер

<http://www.isras.ru/> - Институт социологии РАН.

<http://www.i-u.ru/biblio> - Русский гуманитарный интернет-университет <http://www.job.ru> - Джоб ру

<http://www.kadrovichka.ru> - Кадровичка

www.ecsoman.edu.ru - Федеральный образовательный портал по социологии, экономике и менеджменту

www.soc.ru.ru - электронный ресурс социологического факультета Санкт-Петербургского государственного университета

www.socionet.ru - портал по общественным наукам

www.wciom.ru - официальный сайт ВЦИОМ

7. Методические указания для обучающихся по освоению дисциплины (модуля). Рекомендации для самостоятельной работы.

Подготовку к *практическим занятиям* рекомендуется осуществлять по следующему алгоритму: работа с планами семинарских занятий. При подготовке к семинарскому занятию необходимо найти ответы на поставленные вопросы. Рекомендуется делать конспекты в форме тезисов на каждый вопрос.

Для более глубокого понимания и лучшего усвоения экономических категорий и терминов рекомендуется обращаться к основной и дополнительной литературе, работать с

информационными ресурсами, справочными материалами и периодическими изданиями. Целесообразно вести собственный словарь терминов и использовать его для повторения.

После изучения материала необходимо построить логическую схему знаний, сформулировать вопросы по тем моментам, которые вызвали затруднения, с целью последующего их вынесения на семинарское занятие для обсуждения.

Важным видом работы студентов при изучении дисциплины является *самостоятельная работа*. Самостоятельная работа должна носить творческий и планомерный характер. В процессе организации самостоятельной работы большое значение имеют консультации преподавателя. Они могут быть как индивидуальными, так и в составе учебной группы.

В освоении дисциплины инвалидами и лицами с ограниченными возможностями здоровья большое значение имеет индивидуальная учебная работа (консультации) - дополнительное разъяснение учебного материала. Индивидуальные консультации по предмету являются важным фактором, способствующим индивидуализации обучения и установлению воспитательного контакта между преподавателем и обучающимся инвалидом или лицом с ограниченными возможностями здоровья.

8. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю).

8.1 Перечень информационных технологий.

- Использование электронных презентаций при проведении практических занятий.

8.2 Перечень необходимого программного обеспечения.

При проведении занятий используется пакет PowerPoint Microsoft Office, ОС Microsoft Windows 10, RStudio.

8.3 Перечень информационных справочных систем:

1. Справочно-правовая система «Консультант Плюс» (<http://www.consultant.ru>)
2. Электронная библиотечная система eLIBRARY.RU (<http://www.elibrary.ru>)
3. Электронная библиотечная система "Университетская библиотека ONLINE" (www.biblioclub.ru)
4. Электронная библиотечная система издательства "Лань" (<http://e.lanbook.com/>)
5. Электронная библиотечная система "Юрайт" (<http://www.biblio-online.ru>)
6. Электронная библиотека "Издательского дома "Гребенников" (www.grebennikon.ru)

9. Материально-техническая база, необходимая для осуществления образовательного процесса по дисциплине (модулю).

| № | Вид работ | Материально-техническое обеспечение дисциплины (модуля) и оснащенность |
|----|--|---|
| 1. | Семинарские занятия | Аудитория, оснащенная презентационной техникой (проектор, экран, компьютер/ноутбук) и соответствующим программным обеспечением (ПО). |
| 2. | Лекционные занятия | Аудитория 244, 246, 249, 250, 258. |
| 3. | Текущий контроль, промежуточная аттестация | Аудитория 244, 246, 249, 250, 258. |
| 4. | Самостоятельная работа | Кабинет для самостоятельной работы, оснащенный компьютерной техникой с возможностью подключения к сети «Интернет», обеспеченный доступом в электронную информационно-образовательную среду университета (библиотека). |