

Документ подписан простой электронной подписью

Информация о владельце:

ФИО: Волков В.В.

Должность: Ректор


Дата подписания: 13.09.2024 11:40:38


Уникальный программный ключ:

ed68fd4b85b778e0f0b1bfea5dbc56cf4148f1229917e799a70e513174766571

**Автономная некоммерческая образовательная организация высшего образования  
«Европейский университет в Санкт-Петербурге»**

**Факультет социологии**

УТВЕРЖДАЮ:  
Ректор  В.В. Волков  
« 29 » мая 2024 г.  
Протокол УС № 5 от 29.05.2024 г.



Рабочая программа дисциплины  
**Сбор и анализ текстовых данных**

образовательная программа  
направление подготовки  
**39.04.01 Социология**

направленность (профиль)  
**«Социальные исследования: смыслы и числа»**  
программа подготовки – магистратура

язык обучения – русский  
форма обучения - очная

квалификация выпускника  
**Магистр**

**Санкт-Петербург**

**Автор:**

Тенишева К.А., кандидат социологических наук, доцент факультета социологии АНООВО «ЕУСПб»

**Рецензент:**

Губа К.С., кандидат социологических наук, директор Центра институциональных исследований науки и образования АНООВО «ЕУСПб»

Рабочая программа дисциплины «**Сбор и анализ текстовых данных**», входящей в образовательную программу уровня магистратуры «Социальные исследования: смыслы и числа», утверждена на заседании Совета факультета социологии

Протокол заседания № 11 от 26.04.2024 года

## **АННОТАЦИЯ РАБОЧЕЙ ПРОГРАММЫ ДИСЦИПЛИНЫ.**

### **«Сбор и анализ текстовых данных»**

Дисциплина **«Сбор и анализ текстовых данных»** является факультативной дисциплиной образовательной программы «Социальные исследования: смыслы и числа» по направлению подготовки 39.04.01 Социология.

Дисциплина **«Сбор и анализ текстовых данных»** состоит из двух элементов и представляет собой введение в автоматизированный сбор данных (скрейпинг и парсинг) в сети Интернет и введение в количественный текстовый анализ. Обучение скрейпингу и парсингу, основным продуктом которого часто являются текстовые данные, продолжается логичным рассмотрением основных методов количественного анализа собранных текстов. Мы разберем основы скрейпинга и парсинга – скрейпинг статических и динамических сайтов, парсинг html-кода и таблиц, а также основы количественного текстового анализа – формирование корпусов, разбивка корпусов на токены, стемминг и лемматизацию, работу со стоп-словами, базовые методы количественного текстового анализа и визуализации. Курс не требует предварительной подготовки и углубленных знаний в области программирования или количественного анализа данных.

Общая трудоемкость освоения дисциплины составляет 1 зачетная единица, 36 часов.

## СОДЕРЖАНИЕ

1. НАИМЕНОВАНИЕ, ЦЕЛИ И ЗАДАЧИ ОСВОЕНИЯ ДИСЦИПЛИНЫ	5
2. ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОСВОЕНИЯ ДИСЦИПЛИНЫ	5
3. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОП	7
4. ОБЪЕМ ДИСЦИПЛИНЫ.	7
5. СОДЕРЖАНИЕ И СТРУКТУРА ДИСЦИПЛИНЫ.	8
5.1. Содержание дисциплины	8
5.2. Структура дисциплины	13
5.3. Рекомендации по распределению учебного времени по видам самостоятельной работы и разделам дисциплины:	14
6.1. Общие положения.	14
6.2. Перечень основных вопросов по изучаемым темам для самостоятельной работы обучающихся по дисциплине	14
6.3. Перечень литературы для самостоятельной работы обучающегося:	17
6.4. Перечень учебно-методического обеспечения для самостоятельной работы	17
7. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ ДЛЯ ПРОВЕДЕНИЯ ТЕКУЩЕЙ И ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ ОБУЧАЮЩИХСЯ ПО ДИСЦИПЛИНЕ	20
7.1 Показатели, критерии и оценивание компетенций по уровням их формирования в процессе текущей аттестации	20
7.2 Контрольные задания для текущей аттестации	26
7.3 Показатели, критерии и оценивание компетенций по уровням их формирования в процессе промежуточной аттестации	28
7.4. Типовые задания к промежуточной аттестации	30
8. ОСНОВНАЯ И ДОПОЛНИТЕЛЬНАЯ ЛИТЕРАТУРА	34
8.1. Основная литература:	35
8.2. Дополнительная литература:	35
9. ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ, ИСПОЛЬЗУЕМЫЕ ПРИ ОСУЩЕСТВЛЕНИИ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА	38
9.1 Программное обеспечение	38
9.2 Перечень информационно-справочных систем и профессиональных баз данных информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины:	38
9.3 Лицензионные электронные ресурсы библиотеки Университета	39
9.4 Электронная информационно-образовательная среда Университета	40
10. МАТЕРИАЛЬНО-ТЕХНИЧЕСКАЯ БАЗА, НЕОБХОДИМАЯ ДЛЯ ОСУЩЕСТВЛЕНИЯ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА	40
Приложение 1	42

## 1. НАИМЕНОВАНИЕ, ЦЕЛИ И ЗАДАЧИ ОСВОЕНИЯ ДИСЦИПЛИНЫ

**Цель** освоения дисциплины состоит в том, чтобы познакомить слушателей с автоматизированными методами сбора информации в сети Интернета, ее извлечения и систематизации, а также с основами количественного текстового анализа, привить им практические навыки работы с текстовыми данными в рамках своих собственных исследовательских проектов.

### **Задачи:**

1. Дать понимание основных понятий, связанных с автоматизированным сбором данных в сети Интернет: скрейпинг и парсинг.
2. Рассмотреть основные методы скрейпинга статических и динамических сайтов.
3. Рассмотреть основные методы парсинга html-кода и таблиц.
4. Развить навыки работы с языком программирования Python.
5. Дать понимание основных понятий количественного текстового анализа.
6. Рассмотреть основные методы количественного текстового анализа: подготовка текстовых данных, подсчет частотности и взаимной встречаемости, классификация текстов.
7. Рассмотреть основные методы визуализации результатов количественного текстового анализа
8. Развить навыки работы в программной среде RStudio.
9. Сформировать умение формулировать исследовательские задачи в привязке к методологии количественного текстового анализа.

## 2. ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОСВОЕНИЯ ДИСЦИПЛИНЫ

В результате изучения учебной дисциплины обучающийся должен овладеть следующими результатами обучения по дисциплине, направленными на формирование обучающихся следующих универсальных (УК) компетенций (Таблица 1).

Таблица 1.

### **Планируемые результаты освоения дисциплины, соотнесенные с планируемыми результатами формирования компетенций обучающихся**

<b>Код и название компетенции</b>	<b>Индикаторы компетенции</b>	<b>Планируемые результаты обучения по дисциплине, характеризующие этапы формирования компетенций</b>
УК-1 Способен осуществлять критический анализ проблемных ситуаций на основе системного подхода, вырабатывать стратегию действий	ИД.УК-1.1. Анализирует проблемную ситуацию как систему, выявляя её составляющие и связи между ними	<b>Знать:</b> методы научного познания, в основе которых лежит рассмотрение объекта как системы: целостного комплекса взаимосвязанных элементов, методы и модели стратегического планирования З (УК-1)
	ИД.УК-1.2. Определяет пробелы в информации, необходимой для решения проблемной ситуации, и проектирует процессы по их устранению	<b>Уметь:</b> с использованием методов системного подхода анализировать альтернативные варианты решения исследовательских задач, вырабатывать стратегию действий и оценивать социальную эффективность реализации стратегических планов У (УК-1)
	ИД.УК-1.3. Критически оценивает надёжность источников информации, работает с противоречивой информацией из разных источников ИД.УК-1.4. Разрабатывает и содержательно аргументирует стратегию решения проблемной	<b>Владеть:</b> целостной системой навыков методологического использования системного подхода при решении проблем, возникающих при выполнении исследовательских работ, навыками отстаивания своей точки зрения при выработке стратегических планов выполнения исследовательских работ В (УК-1)

Код и название компетенции	Индикаторы компетенции	Планируемые результаты обучения по дисциплине, характеризующие этапы формирования компетенций
	ситуации на основе системного и междисциплинарного подходов ИД.УК-1.5. Строит сценарии реализации стратегии, определяя возможные риски и предлагая пути их устранения	

В результате освоения дисциплины магистрант должен:

- **знать:** основные понятия, связанные с автоматизированным сбором данных в сети Интернет, основные методы скрейпинга и парсинга статических и динамических сайтов и таблиц; принципы работы с языком программирования Python; основные понятия и методы, связанные с количественным анализом текстов; основные методы работы в программной среде RStudio; принципы формулирования целей и постановки задач научных исследований в фундаментальных и прикладных областях социологии; принципы решения исследовательских задач с использованием новейшего отечественного и зарубежного опыта и с применением современной аппаратуры, оборудования, информационных технологий;

- **уметь:** применять в самостоятельной профессиональной научной деятельности основные концепты автоматизированного сбора данных и количественного текстового анализа; применять в профессиональной деятельности язык программирования Python и программную среду R; формулировать исследовательские задачи в связи с методологией количественного текстового анализа; осваивать новые теории, модели, методы исследования, разрабатывать новые методические подходы с учетом целей и задач исследования; формулировать цели и ставить задачи научных исследований в фундаментальных и прикладных областях социологии; решать исследовательские задачи с использованием новейшего отечественного и зарубежного опыта и с применением современной аппаратуры, оборудования, информационных технологий; соотносить собственные исследовательские задачи с возможностями методологии количественного текстового анализа;

- **владеть:** навыками применения в самостоятельной профессиональной научной деятельности основных концептов автоматизации сбора данных, основных методов количественного текстового анализа в рамках социальных наук; навыками применения в профессиональной деятельности языка программирования Python и программной среды R; навыками постановки исследовательских задач в связи с методологией количественного текстового анализа; навыками освоения новых теорий, моделей, методов исследования, разработки новых методических подходов с учетом целей и задач исследования; навыками постановки цели и задач научных исследований в фундаментальных и прикладных областях социологии; навыками решения исследовательских задач с использованием новейшего отечественного и зарубежного опыта и с применением современной аппаратуры, оборудования, информационных технологий; навыками соотнесения собственных исследовательских задач с возможностями методологии количественного текстового анализа.

### 3. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОП

Дисциплина «Сбор и анализ текстовых данных» является факультативной дисциплиной. Курс читается в первом семестре, форма промежуточной аттестации – зачет с оценкой.

### 4. ОБЪЕМ ДИСЦИПЛИНЫ.

Общая трудоемкость освоения дисциплины составляет 1 зачетная единица, 36 часов.

Таблица 2.

Объем дисциплины

Вид учебных занятий и самостоятельная работа	Объем дисциплины, час.				
	Всего	Семестр			
		1	2	3	4
<b>Контактная работа обучающихся с преподавателем, в том числе:</b>	<b>28</b>	28	-	-	-
Лекции (Л)	<b>14</b>	14	-	-	-
Семинарские занятия (СЗ)	<b>14</b>	14	-	-	-
Самостоятельная работа (СР)	<b>8</b>	8	-	-	-
<b>Промежуточная аттестация</b>	<b>форма</b>	<b>Зачет с оценкой</b>	Зачет с оценкой	-	-
	<b>час.</b>	<b>-</b>	-	-	-
<b>Общая трудоемкость (час. / з.е.)</b>	<b>36/1</b>	36/1	-	-	-

## 5. СОДЕРЖАНИЕ И СТРУКТУРА ДИСЦИПЛИНЫ.

Содержание дисциплины соотносится с планируемыми результатами обучения по дисциплине: через задачи, формируемые компетенции и их компоненты (знания, умения, навыки – далее ЗУВ) по средствам индикаторов достижения компетенций в соответствии с Таблицей 3.

### 1. 5.1. Содержание дисциплины

Таблица 3.

Содержание дисциплины

№ п/п	Наименование тем (разделов)	Содержание тем (разделов)	Коды компетенций	Индикаторы компетенций (в соот. с Таблицей 1)	Коды ЗУВ (в соответствии с Таблицей 1)
<b>1</b>	Информация, данные и автоматизация сбора. Скрейпинг и парсинг. Основы Python.	Что такое информация, что такое данные, и как автоматизировать преобразование информации в данные. Наиболее типичные формы информации/данных, доступные в сети Интернет (сырые данные – текст и числа в тексте, таблицы), типы веб-сайтов.	УК-1	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5.	З (УК-1) У (УК-1) В (УК-1)
<b>2</b>	Скрейпинг и парсинг статического сайта – часть I.	Основы Python: продолжение. HTML теги и структура сайта. HTTP-запросы. Библиотеки requests и BeautifulSoup для скрейпинга и парсинга. Списки и преобразование данных в датафрейм.	УК-1	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5.	З (УК-1) У (УК-1) В (УК-1)
<b>3</b>	Скрейпинг и парсинг статического сайта – часть I. Парсинг таблиц.	Основы Python: продолжение. Различные методы доступа к HTML тегам и их атрибутам. Таблицы как наиболее распространённый формат данных в сети Интернет. Подходы к парсингу таблиц.	УК-1	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5.	З (УК-1) У (УК-1) В (УК-1)
<b>4.</b>	Скрейпинг и парсинг	Библиотека Selenium и доступ к содержимому динамических веб-	УК-1	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3.	З (УК-1) У (УК-1) В (УК-1)

	динамического сайта – часть I.	сайтов.		ИД.УК-1.4. ИД.УК-1.5.	
5	Скрейпинг и парсинг динамического сайта – часть II.	Библиотека Selenium: решение более сложных задач: скроллинг страниц, переключение между вкладками и ожидание появления элементов.	УК-1	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5.	З (УК-1) У (УК-1) В (УК-1)
6	Количественный текстовый анализ: основы и принципы. Основы работы в RStudio.	Основные принципы количественного текстового анализа. Методология bag of words. Знакомство с RStudio и пакетом quanteda	УК-1	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5.	З (УК-1) У (УК-1) В (УК-1)
7	Подготовка текстов к количественному анализу.	Формирование корпуса текстов, разбиение на токены и работа со стоп-словами. Стемминг и лемматизация.	УК-1	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5.	З (УК-1) У (УК-1) В (УК-1)
8	Преобразование корпуса текстов в структуры данных, пригодные для количественного текстового анализа.	Получение DFM (document frequency matrix), feature cooccurrence matrix и расчет метрики term frequency-inverse document frequency.	УК-1	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5.	З (УК-1) У (УК-1) В (УК-1)
9	Базовые методы количественного текстового анализа. Визуализации результатов.	Расчет частотности использования слов. Расчет матрицы взаимного использования слов. Визуализация частотности слов. Визуализация контекста использования отдельных терминов в корпусе текстов.	УК-1	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5.	З (УК-1) У (УК-1) В (УК-1)
10	Применение машинного обучения для классификации текстов.	Байесовский наивный классификатор – применение для классификации по темам, визуализация результатов.	УК-1	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5.	З (УК-1) У (УК-1) В (УК-1)

## 2. 5.2. Структура дисциплины

Таблица 4.

Структура дисциплины

№ п/п	Наименование тем (разделов)	Объем дисциплины, час.				Форма текущего контроля успеваемости*, промежуточной аттестации
		Всего	Контактная работа обучающихся с преподавателем по типам учебных занятий в соответствии с УП		СР	
			Л	СЗ		
Тема 1	Информация, данные и автоматизация сбора. Скрейпинг и парсинг. Основы Python.	2	1	1	-	ДЗ/О
Тема 2	Скрейпинг и парсинг статического сайта – часть I.	2	1	1	-	ДЗ/ПЗ



№ п/п	Наименование тем (разделов)	Объем дисциплины, час.				Форма текущего контроля успеваемости*, промежуточной аттестации
		Всего	Контактная работа обучающихся с преподавателем по типам учебных занятий в соответствии с УП		СР	
			Л	СЗ		
Тема 3	Скрейпинг и парсинг статического сайта – часть I. Парсинг таблиц.	3	1	1	1	ДЗ/ПЗ
Тема 4	Скрейпинг и парсинг динамического сайта – часть I.	3	1	1	1	ДЗ/ПЗ
Тема 5	Скрейпинг и парсинг динамического сайта – часть II.	3	1	1	1	ДЗ/ПЗ
Тема 6	Количественный текстовый анализ: основы и принципы. Основы работы в RStudio.	3	1	1	1	ДЗ/ПЗ
Тема 7	Подготовка текстов к количественному анализу.	5	2	2	1	ДЗ/ПЗ
Тема 8	Преобразование корпуса текстов в структуры данных, пригодные для количественного текстового анализа.	5	2	2	1	ДЗ/ПЗ
Тема 9	Базовые методы количественного текстового анализа. Визуализации результатов.	5	2	2	1	ДЗ/ПЗ
Тема 10	Применение машинного обучения для классификации текстов.	5	2	2	1	ДЗ/ПЗ
<b>Промежуточная аттестация</b>		-	-	-	-	<b>Зачет с оценкой</b>
<b>Всего:</b>		<b>36/1</b>	<b>14</b>	<b>14</b>	<b>8</b>	<b>-</b>

*\*Примечание: формы текущего контроля успеваемости: домашнее задание (ДЗ), практическое задание (ПЗ), опрос (О)*

## **6. УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ САМОСТОЯТЕЛЬНОЙ РАБОТЫ ОБУЧАЮЩИХСЯ ПО ДИСЦИПЛИНЕ.**

### **6.1. Общие положения.**

Знания и навыки, полученные в результате лекций, семинарских занятий закрепляются и развиваются в результате повторения материала, усвоенного в аудитории, путем чтения текстов и исследовательской литературы (из списков основной и дополнительной литературы) и их анализа.

Самостоятельная работа является важнейшей частью процесса высшего образования. Ее следует осознанно организовать, выделив для этого необходимое время и соответственным образом организовав рабочее пространство. Важнейшим элементом самостоятельной работы является проработка материалов прошедших занятий (анализ конспектов, чтение рекомендованной литературы) и подготовка к следующим лекциям/семинарским занятиям. Литературу, рекомендованную в программе курса, следует, по возможности, читать в течение всего семестра, концентрируясь на обусловленных программой курса темах.

Существенную часть самостоятельной работы магистранта представляет самостоятельное изучение вспомогательных учебно-методических изданий, лекционных конспектов, интернет-ресурсов и пр. Подготовка к семинарским занятиям, опросам, диспутам, выполнение практических заданий также является важной формой работы

магистранта. Самостоятельная работа может вестись как индивидуально, так и при содействии преподавателя.

## **6.2 Рекомендации по распределению учебного времени по видам самостоятельной работы и разделам дисциплины**

### **Тема 3. Скрейпинг и парсинг статического сайта – часть I. Парсинг таблиц**

1. Подготовка к занятиям по предложенным темам, самостоятельное изучение рекомендованной литературы, повторение материала лекций – 1 час.

### **Тема 4. Скрейпинг и парсинг динамического сайта – часть I**

1. Подготовка к занятиям по предложенным темам, самостоятельное изучение рекомендованной литературы, повторение материала лекций – 1 час.

### **Тема 5. Скрейпинг и парсинг динамического сайта – часть II**

1. Подготовка к занятиям по предложенным темам, самостоятельное изучение рекомендованной литературы, повторение материала лекций – 1 час.

### **Тема 6. Количественный текстовый анализ: основы и принципы. Основы работы в RStudio**

1. Подготовка к занятиям по предложенным темам, самостоятельное изучение рекомендованной литературы, повторение материала лекций – 1 час.

### **Тема 7. Подготовка текстов к количественному анализу**

1. Подготовка к занятиям по предложенным темам, самостоятельное изучение рекомендованной литературы, повторение материала лекций – 1 час.

### **Тема 8. Преобразование корпуса текстов в структуры данных, пригодные для количественного текстового анализа**

1. Подготовка к занятиям по предложенным темам, самостоятельное изучение рекомендованной литературы, повторение материала лекций – 1 час.

### **Тема 9. Базовые методы количественного текстового анализа. Визуализации результатов**

1. Подготовка к занятиям по предложенным темам, самостоятельное изучение рекомендованной литературы, повторение материала лекций – 1 час.

### **Тема 10. Применение машинного обучения для классификации текстов**

1. Подготовка к занятиям по предложенным темам, самостоятельное изучение рекомендованной литературы, повторение материала лекций – 1 час.

## **3. 6.3 Перечень основных вопросов по изучаемым темам для самостоятельной работы обучающихся по дисциплине**

### **Вопросы для самостоятельной подготовки по темам:**

1. В чём разница между информацией и данными? Как преобразовать информацию в данные для анализа?
2. Какие типичные формы информации/данных существуют в интернете, и чем они отличаются?
3. Что такое HTML-теги и как они описывают структуру веб-сайта?
4. Как работают HTTP-запросы, и для чего они используются при сборе данных?
5. Каковы основные функции библиотек requests и BeautifulSoup в Python?

6. Как преобразовать собранные данные в датафрейм для дальнейшего анализа?
7. Какие существуют методы доступа к HTML-тегам и их атрибутам в Python?
8. Почему таблицы являются распространенным форматом данных в интернете, и как их парсить?
9. Что такое динамические веб-сайты, и как библиотека Selenium помогает получить к ним доступ?
10. Как с помощью Selenium реализовать скроллинг, переключение между вкладками и ожидание загрузки элементов?
11. В чем суть количественного текстового анализа и методологии "мешка слов"?
12. Как сформировать корпус текстов, разбить текст на токены и удалить стоп-слова?
13. Что такое стемминг и лемматизация, и зачем они нужны?
14. Как получить DFM, feature cooccurrence matrix и рассчитать TF-IDF?
15. Как визуализировать результаты анализа текстовых данных (частотность слов, контекст использования)?

#### **6.4. Перечень литературы для самостоятельной работы обучающегося:**

1. Колданов, А. П. Теория вероятностей и математическая статистика : учебник / А. П. Колданов, П. А. Колданов ; Нац. исслед. ун-т «Высшая школа экономики». — М. : Изд. дом Высшей школы экономики, 2023. — 245, [3] с. — (Учебники Высшей школы экономики). — ISBN 978-5-7598-2544-9 (в пер.). — ISBN 978-5-7598-2829-7 (e-book). — URL: <http://176.9.74.196/book.html?currBookId=45113>.

2. Львовский, С. М. Основы математического анализа [Текст] : учебник для вузов / С. М. Львовский ; Нац. исслед. ун-т «Высшая школа экономики». — М. : Изд. дом Высшей школы экономики, 2021. — (Учебники Высшей школы экономики). — 367, [1] с. — 600 экз. — ISBN 978-5-7598-1183-1 (в пер.). — ISBN 978-5-7598-2405-3 (e-book). — URL: <http://176.9.74.196/book.html?currBookId=40183>.

#### **6.5. Перечень учебно-методического обеспечения для самостоятельной работы**

Для обеспечения самостоятельной работы магистрантов по дисциплине «Сбор и анализ текстовых данных» разработано учебно-методическое обеспечение в составе:

1. Контрольные задания для подготовки к процедурам текущего контроля (п. 7.2 Рабочей программы).
2. Типовые задания для подготовки к промежуточной аттестации (п. 7.4 Рабочей программы).
3. Рекомендуемые основная, дополнительная литература, Интернет-ресурсы и справочные системы (п. 8, 9 Рабочей программы).
4. Рабочая программа дисциплины размещена в электронной информационно-образовательной среде Университета на электронном учебно-методическом ресурсе АНООВО «ЕУСПб» — образовательном портале LMS Sakai — Sakai@EU.

### **7. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ ДЛЯ ПРОВЕДЕНИЯ ТЕКУЩЕЙ И ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ ОБУЧАЮЩИХСЯ ПО ДИСЦИПЛИНЕ**

#### **7.1 Показатели, критерии и оценивание компетенций и индикаторов их достижения в процессе текущей аттестации**

Информация о содержании и процедуре текущего контроля успеваемости, методике оценивания знаний, умений и навыков обучающегося в ходе текущего контроля доводятся

научно-педагогическими работниками Университета до сведения обучающегося на первом занятии по данной дисциплине.

Текущий контроль предусматривает подготовку магистрантов к каждому семинарскому занятию, участие в опросах, диспутах, выполнять обзоры литературы, активное слушание на лекциях. Магистрант должен присутствовать на семинарских занятиях, отвечать на поставленные вопросы, показывая, что прочитал разбираемую литературу, представлять содержательные реплики по темам обсуждения.

Текущий контроль проводится в форме устных опросов, оценивания участия магистрантов в проходящих диспутах, выполнении практических заданий демонстрирующих степень знакомства магистрантов с дополнительной литературой.

Таблица 5.

**Показатели, критерии и оценивание компетенций и индикаторов их достижения в процессе текущей аттестации**

<b>Наименование тем (разделов)</b>	<b>Коды компетенций</b>	<b>Индикаторы компетенций (в соот. с Таблицей 1)</b>	<b>Коды ЗУВ (в соответствии с Таблицей 1)</b>	<b>Формы текущего контроля успеваемости и</b>	<b>Результаты текущего контроля</b>
Информация, данные и автоматизация сбора. Скрейпинг и парсинг. Основы Python.	УК-1	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5.	З (УК-1) У (УК-1) В (УК-1)	Опрос Домашнее задание	зачтено/ не зачтено
Скрейпинг и парсинг статического сайта – часть I.	УК-1	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5.	З (УК-1) У (УК-1) В (УК-1)	Практическое задание Домашнее задание	зачтено/ не зачтено
Скрейпинг и парсинг статического сайта – часть I. Парсинг таблиц.	УК-1	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5.	З (УК-1) У (УК-1) В (УК-1)	Практическое задание Домашнее задание	зачтено/ не зачтено
Скрейпинг и парсинг динамического сайта – часть I.	УК-1	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5.	З (УК-1) У (УК-1) В (УК-1)	Практическое задание Домашнее задание	зачтено/ не зачтено
Скрейпинг и парсинг динамического сайта – часть II.	УК-1	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5.	З (УК-1) У (УК-1) В (УК-1)	Практическое задание Домашнее задание	зачтено/ не зачтено
Количественный текстовый анализ: основы и принципы. Основы работы в RStudio.	УК-1	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5.	З (УК-1) У (УК-1) В (УК-1)	Опрос Домашнее задание	зачтено/ не зачтено
Подготовка текстов к количественному анализу.	УК-1	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5.	З (УК-1) У (УК-1) В (УК-1)	Практическое задание Домашнее задание	зачтено/ не зачтено

Преобразование корпуса текстов в структуры данных, пригодные для количественного текстового анализа.	УК-1	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5.	З (УК-1) У (УК-1) В (УК-1)	Практическое задание Домашнее задание	зачтено/ не зачтено
Базовые методы количественного текстового анализа. Визуализации результатов.	УК-1	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5.	З (УК-1) У (УК-1) В (УК-1)	Практическое задание Домашнее задание	зачтено/ не зачтено
Применение машинного обучения для классификации текстов.	УК-1	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5.	З (УК-1) У (УК-1) В (УК-1)	Практическое задание Домашнее задание	зачтено/ не зачтено

Таблица 6

### Критерии оценивания

Формы текущего контроля успеваемости	Критерии оценивания
Домашнее задание	выполнение домашнего задания с существенными ошибками или пропусками заданий – не зачтено, полное и правильное выполнение домашнего задания – зачтено
Практическое занятие	выполнение практического задания с существенными ошибками или пропусками заданий – не зачтено, полное и правильное выполнение этапов практического задания – зачтено
Опрос	ответ отсутствует или является односложным – не зачтено развернутый ответ с доказательствами или обоснованием — зачтено

## 7.2 Контрольные задания для текущей аттестации

**Примерный материал для опросов, практических заданий и домашних заданий.**

**Тема 1. Информация, данные и автоматизация сбора. Скрейпинг и парсинг.**

### Основы Python

#### Опрос 1:

1. Какие существуют основные методы сбора данных с веб-страниц?
2. В чем различие между скрейпингом и парсингом данных?
3. Какие преимущества и недостатки автоматизации сбора данных вы можете назвать?
4. Какие библиотеки Python используются для скрейпинга и парсинга?

#### Домашнее задание 1:

Напишите скрипт на Python, который будет выводить только те элементы списка, длина которых равна 4. для решения задачи примените итерацию по спискам и условные операторы.

### Тема 2. Скрейпинг и парсинг статического сайта – часть I

#### Практическое задание 1:

Создайте Python-скрипт, который будет извлекать все ссылки с заданной веб-страницы. Используйте библиотеку BeautifulSoup и requests.

#### Домашнее задание 2:

Создайте Python-скрипт, который будет скрейпить первую страницу указанного сайта и парсить заголовки, ссылки, даты и описания текстовых элементов. Сохраните все

проанализированные данные (даты, заголовки, ссылки и тексты) в виде датафрейма. Экспортируйте датафрейм в файлы csv, xlsx.

### **Тема 3. Скрейпинг и парсинг статического сайта – часть I. Парсинг таблиц**

#### **Практическое задание 2:**

Используя Python и библиотеку BeautifulSoup, напишите скрипт, который будет извлекать данные из таблицы на веб-странице и сохранять их в формате csv и xlsx.

#### **Домашнее задание 3:**

Создайте скрипт на Python, который собирает данные из таблицы с заданного веб-сайта, а затем преобразует и сохраняет их в формате xlsx и csv. Обратите внимание на правильное форматирование данных в таблице.

### **Тема 4. Скрейпинг и парсинг динамического сайта – часть I**

#### **Практическое задание 3:**

Разработайте Python-скрипт, который сможет извлечь данные с указанного динамического сайта, путем взаимодействия с динамическими объектами: выпадающими меню, полями ввода и радио-кнопками для загрузки контента. Используйте библиотеку Selenium.

#### **Домашнее задание 4:**

Напишите скрипт на Python, который будет взаимодействовать с динамическим сайтом, используя Selenium. В качестве примера соберите информацию о товарах с онлайн-магазина, добавьте фильтры на сайте, и соберите данные о ценах и описаниях товаров.

### **Тема 5. Скрейпинг и парсинг динамического сайта – часть II**

#### **Практическое задание 4:**

Напишите скрипт на Python с использованием Selenium, который откроет браузер, перейдет на динамический сайт, заполнит поля/будет взаимодействовать с указанными интерактивными элементами и извлечет данные с нескольких страниц.

#### **Домашнее задание 5:**

Разработайте более сложный скрипт, который не только взаимодействует с интерактивными элементами, но и обходит несколько страниц сайта для сбора указанных данных.

### **Тема 6. Количественный текстовый анализ: основы и принципы. Основы работы в RStudio**

#### **Опрос 2:**

1. В чем заключается основная идея количественного текстового анализа?
2. Какие задачи решаются с помощью количественного текстового анализа?
3. Какие базовые библиотеки и функции RStudio вы знаете, которые помогают в проведении текстового анализа?

#### **Домашнее задание 6:**

Изучите документацию к библиотеке quanteda в R. Потренируйтесь загружать набор текстов (например, статьи из одного новостного источника) в R.

### **Тема 7. Подготовка текстов к количественному анализу**

#### **Практическое задание 5:**

Используя R и библиотеку quanteda, проведите очистку текста от стоп-слов, пунктуации и других ненужных элементов. Разбейте текст на токены и сохраните результат в формате, пригодном для дальнейшего анализа.

### **Домашнее задание 7:**

Подготовьте текстовые данные (например, статьи из одного новостного источника) для количественного анализа. Очистите данные от ненужных символов, стоп-слов, и проведите лемматизацию. Сохраните очищенные данные в форматах csv и xlsx.

## **Тема 8. Преобразование корпуса текстов в структуры данных, пригодные для количественного текстового анализа**

### **Практическое задание 6:**

Создайте document frequency matrix на основе текстового корпуса в R с помощью библиотеки quanteda. Используйте эту матрицу для проведения простого анализа частоты слов. Создайте frequency cooccurrence matrix с помощью библиотеки quanteda.

### **Домашнее задание 8:**

Потренируйтесь в создании DFM и FCM-матриц для заданного корпуса текстов.

## **Тема 9. Базовые методы количественного текстового анализа. Визуализации результатов**

### **Практическое задание 7:**

Используя библиотеку quanteda в R визуализируйте разными способами частотность распределения слов по текстам и взаимную встречаемость слов с помощью сетевого графика.

### **Домашнее задание 9:**

Примените изученные базовые методы визуализации DFM и FCM-матриц для корпуса текстов. Интерпретируйте результаты визуализаций.

## **Тема 10. Применение машинного обучения для классификации текстов**

### **Практическое задание 8:**

Используя R, разработайте модель машинного обучения для классификации текстов. Проведите тематический анализ текста с использованием LDA (Latent Dirichlet Allocation) в R.

### **Домашнее задание 10:**

Примените метод машинного обучения - LDA - для классификации указанного корпуса текстов.

## **7.3 Показатели, критерии и оценивание компетенций и индикаторов их достижения в процессе промежуточной аттестации**

**Форма промежуточной аттестации** – зачет с оценкой, который проходит в форме написания финального эссе. Темы эссе к промежуточному контролю сообщены магистрантом заранее.

Перед зачетом с оценкой проводится консультация, на которой преподаватель отвечает на вопросы магистрантов.

В результате промежуточного контроля знаний студенты получают аттестацию по дисциплине.

Таблица 7

### **Показатели, критерии и оценивание компетенций и индикаторов их достижения в процессе промежуточной аттестации**

Форма промежуточной аттестации/вид промежуточной аттестации	Коды компетенций	Индикаторы компетенций (в соот. с Таблицей 1)	Коды ЗУВ (в соответствии и с Таблицей 1)	Критерии оценивания	Оценка
Зачет с оценкой / письменная работа (эссе)	УК-1	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5.	З (УК-1) У (УК-1) В (УК-1)	<p>Письменная работа (эссе) соответствует следующим требованиям:</p> <p>вопрос раскрыт развернуто, использована основная и дополнительная литература по курсу, соблюдены структура и научный стиль, сформулированы выводы, аргументация убедительна, правильно оформлен библиографический аппарат и т.д.</p> <p>Магистрант демонстрирует:</p> <p>глубокое усвоение программного материала; изложение данного материала исчерпывающе, последовательно, четко; умение делать обоснованные выводы; соблюдение норм устной и письменной литературной речи.</p>	Зачтено, отлично
				<p>В письменной работе (эссе) не соблюдены некоторые требования: вопрос раскрыт в целом полно, но данные представлены недостаточно убедительно, выводы сформулированы недостаточно четко, аргументация недостаточно убедительна, использована основная и дополнительная литература по курсу.</p> <p>Магистрант демонстрирует:</p> <p>твердое знание материала курса; последовательное изложение материала; знание теоретических положений без обоснованной их аргументации; соблюдение норм устной и письменной литературной речи.</p>	Зачтено, хорошо
				Письменная работа (эссе) содержит существенные оплошности:	Зачтено, удовлетворительно



Форма промежуточной аттестации/вид промежуточной аттестации	Коды компетенций	Индикаторы компетенций (в соот. с Таблицей 1)	Коды ЗУВ (в соответствии и с Таблицей 1)	Критерии оценивания	Оценка
				<p>нарушено сразу несколько требований, например, выводы плохо обоснованы; есть фактические ошибки, вопрос раскрыт частично, использована основная, но не использована дополнительная литература по курсу.</p> <p>Магистрант демонстрирует: знание основного материала, но владение им не в полном объеме;</p> <p>допущение существенных неточностей; допущение недостаточно правильных формулировок; допущение нарушения логической последовательности в изложении материала; наличие нарушений норм литературной устной и письменной речи.</p>	
				<p>Письменная работа (эссе) не представлена или не отвечает предъявляемым требованиям: вопрос не раскрыт, не использована литература по курсу, магистрант допускает нарушение научного стиля, структуры изложения ответа, не аргументирует свои тезисы и идеи.</p> <p>Магистрант демонстрирует: незнание значительной части программного материала: наличие существенных ошибок в определениях, формулировках, понимании теоретических положений; бессистемность при ответе на поставленный вопрос; отсутствие в ответе логически корректного анализа, аргументации, классификации;</p> <p>наличие нарушений норм устной и письменной литературной речи.</p>	Не зачтено, неудовлетворительно

Результаты промежуточного контроля по дисциплине, выраженные в оценках «зачтено, удовлетворительно», «зачтено, хорошо», «зачтено, отлично» показывают уровень сформированности у обучающегося компетенций по дисциплине в соответствии с картами компетенций образовательной программы «Социальные исследования: смыслы и числа» по направлению подготовки 39.04.01 Социология (уровень магистратуры).

Результаты промежуточного контроля по дисциплине, выраженные в оценке «не зачтено, неудовлетворительно», показывают не сформированность у обучающегося компетенций по дисциплине в соответствии с картами компетенций образовательной программы «Социальные исследования: смыслы и числа» по направлению подготовки 39.04.01 Социология (уровень магистратуры).

Таблица 7а

Система оценки знаний обучающихся		
Пятибалльная (стандартная) система	Стобалльная система оценки	Бинарная система оценки
5 (отлично)	100-81	зачтено
4 (хорошо)	80-61	
3 (удовлетворительно)	60-41	
2 (неудовлетворительно)	40 и менее	не зачтено

#### 7.4. Типовые задания к промежуточной аттестации

Работа над итоговым эссе в рамках дисциплины предполагает работу со статическим или динамическим сайтом (выбранным магистрантом или предложенным преподавателем), скрейпинг нескольких страниц этого сайта и последующий парсинг выбранных фрагментов веб-страниц с целью формирования базы текстовых данных для последующей работы с ними в рамках итоговой аттестации. Полученную базу текстовых данных магистрант использует для проведения количественного текстового анализа данных и ответа на поставленный исследовательский вопрос. В ходе подготовки и анализа текстовых данных магистрант должен использовать один или несколько изученных способов подготовки и анализа текстовых данных, а также одну или несколько способов визуализации этих данных. Вместе с текстом эссе магистрант предоставляет преподавателю скрипты, использовавшиеся для сбора и анализа данных.

##### Требование к итоговому эссе:

Тема письменной итоговой работы (эссе) согласовывается с преподавателем, зависит как от научных интересов магистранта, так и от рекомендаций преподавателя. При написании итоговой письменной работы (эссе) необходимо помнить о правильной структуре письменной работы. Она должна содержать введение, основную часть и заключение. Заключение не должно вводить новых мыслей, а должно резюмировать аргумент.

При написании эссе обязательно использование языка программирования Python и прогаммной среды R/Rstudio.

Обязательно использование текстовых данных и их анализа при написании работы и формулировке теоретических выводов.

Предпочтительно использование основной и дополнительной литературы.

Объем работы: 8000-10000 знаков.

Работа должна выполняться самостоятельно, не содержать построчного переложения чужих текстов и перевода иностранных источников. Все цитаты должны быть с указанием их источника.

В письменной работе обязательно должен присутствовать ряд блоков.

1. Вопрос, проблема, цели, задачи гипотезы. Из этого блока должно быть понятно, что хочет узнать автор путем сбора и анализа тех или иных текстовых данных.

2. Описание процедуры сбора, подготовки и анализа данных. Перечень переменных, их (краткое) описание и перечень преобразований, совершенных с ними при подготовке к

анализу. Из этого блока должно быть понятно, какие переменные и почему автор использует для ответа на свой вопрос.

3. Описание выбора процедур анализа – краткая характеристика выбранного метода анализа и объяснение применимости данного метода а) для решения поставленных задач и б) для этих данных. Из этого блока должно быть понятно, что с помощью данного метода можно найти ответ на поставленный вопрос и что данный метод можно использовать, применительно к этим данным.

4. Результаты анализа. Основные таблицы и/или визуализации. Первичная интерпретация результатов. Из этого блока должно быть понятно, к каким выводам приходит автор.

6. Интерпретация. Краткое резюме с ответом на основной вопрос и (по возможности) помещение этих результатов в контекст научной дискуссии. Из этого блока должно быть понятно, к каким выводам приходит автор по поводу поставленных в начале вопросов.

Сдаче итоговой аттестации предшествует работа студента на лекционных, семинарских занятиях и самостоятельная работа. Отсутствие студента на занятиях без уважительных причин и невыполнение заданий самостоятельной работы является основанием для недопущения студента к зачету с оценкой.

#### **Примерные темы зачетных письменных работ (эссе):**

1. Классификация новостных сообщений о Ковид-19;
2. Определение схожести риторики в журналах по международной политике и официальном дискурсе;
3. Анализ различий риторики стран членов СБ ООН по вопросу применения концепции “Обязанность защищать”

### **7.5. Средства оценки индикаторов достижения компетенций**

Таблица 8

Средства оценки индикаторов достижения компетенций		
Коды компетенций	Индикаторы компетенций (в соот.с Таблицей 1)	Средства оценки (в соот. с Таблицами 5, 7)
УК-1	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5.	Домашнее задание, практическое задание, опрос, эссе

Таблица 9

Описание средств оценки индикаторов достижения компетенций	
Средства оценки (в соот. с Таблицами 5, 7)	Рекомендованный план выполнения работы
Опрос	Магистранту рекомендуется в ходе подготовки и участия в опросе по темам дисциплины, совершать следующие действия: - анализируя проблемную ситуацию, определяет пробелы в информации, оценивает надёжность источников информации, разрабатывает стратегию решения проблемной ситуации на основе системного и междисциплинарного подходов, строит сценарии реализации стратегии, определяя возможные риски и предлагая пути их устранения
Домашнее задание	Магистрант в ходе подготовки и выполнения домашнего задания по темам на разных этапах, показывает наличие теоретической и практической базы знаний в рамках дисциплины, необходимой для выполнения следующих действий в области профессиональной деятельности: - анализируя проблемную ситуацию, определяет пробелы в информации,

Средства оценки (в соот. с Таблицами 5, 7)	Рекомендованный план выполнения работы
	оценивает надёжность источников информации, разрабатывает стратегию решения проблемной ситуации на основе системного и междисциплинарного подходов, строит сценарии реализации стратегии, определяя возможные риски и предлагая пути их устранения
Практическое задание	Магистрант в ходе подготовки и представления практического задания по темам на разных этапах, показывает наличие теоретической и практической базы знаний в рамках дисциплины, необходимой для выполнения следующих действий в области профессиональной деятельности: - анализируя проблемную ситуацию, определяет пробелы в информации, оценивает надёжность источников информации, разрабатывает стратегию решения проблемной ситуации на основе системного и междисциплинарного подходов, строит сценарии реализации стратегии, определяя возможные риски и предлагая пути их устранения
Эссе	Магистрант в ходе подготовки эссе, показывает наличие теоретической и практической базы знаний в рамках дисциплины, необходимой для выполнения следующих действий в области профессиональной деятельности: - анализируя проблемную ситуацию, определяет пробелы в информации, оценивает надёжность источников информации, разрабатывает стратегию решения проблемной ситуации на основе системного и междисциплинарного подходов, строит сценарии реализации стратегии, определяя возможные риски и предлагая пути их устранения

## 8. ОСНОВНАЯ И ДОПОЛНИТЕЛЬНАЯ ЛИТЕРАТУРА

### Основная литература:

1. Крыштановский, А. О. Анализ социологических данных с помощью пакета SPSS [Текст] : учеб. пособие для вузов / А. О. Крыштановский ; Гос. ун-т — Высшая школа экономики. — 2-е изд. — М. : Изд. дом ГУ ВШЭ, 2007. — 281, [3] с. — (Учебники Высшей школы экономики). — Прил.: с. 225—281. — ISBN 978-5-7598-0486-4. — URL: <http://176.9.74.196/book.html?currBookId=21059>.
2. Козырьков, В. П. Социология культуры в цифровом обществе : учебное пособие / В. П. Козырьков, М. В. Придатченко, Н. В. Шалютина ; под ред. В. П. Козырькова. — СПб. : Алетей, 2022. — 378 с. — ISBN 978-5-00165-388-2. — URL: <http://176.9.74.196/book.html?currBookId=41707>.
3. Козлов, А. Д. Методы анализа предметных областей [Электронный ресурс] : учебное пособие / А. Д. Козлов, В. А. Лекае, М. С. Шаповалова ; Рос. гос. гуманитарн. ун-т. — 3-е изд. (эл.). — Электрон. текст. дан. (1 файл pdf : 203 с.). — М. : Рос. гос. гуманитарн. ун-т, 2019. — Систем. требования: Adobe Reader XI либо Adobe Digital Editions 4.5 ; экран 10". — ISBN 978-5-7281-2489-4. — ISBN 978-5-7281-2065-0. — URL: <http://176.9.74.196/book.html?currBookId=31920>.

### Дополнительная литература:

1. Практики анализа качественных данных в социальных науках : учеб. пособие / отв. ред. Е. В. Полухина ; Нац. исслед. ун-т «Высшая школа экономики». — М. : Изд. дом Высшей школы экономики, 2023. — 383, [1] с. — ISBN 978-5-7598-2542-5 (в обл.). — ISBN 978-5-7598-2497-8 (e-book). — URL: <http://176.9.74.196/book.html?currBookId=45129>.
2. Львовский, С. М. Основы математического анализа [Текст] : учебник для вузов / С. М. Львовский ; Нац. исслед. ун-т «Высшая школа экономики». — М. : Изд. дом Высшей

школы экономики, 2021. — (Учебники Высшей школы экономики). — 367, [1] с. — 600 экз. — ISBN 978-5-7598-1183-1 (в пер.). — ISBN 978-5-7598-2405-3 (e-book). — URL: <http://176.9.74.196/book.html?currBookId=40183>.

3. Дятлов, А. В. Методы многомерного статистического анализа данных в социологии : учебник / А. В. Дятлов, П. Н. Лукичев ; Южный федеральный университет. - Ростов-на-Дону : Издательство Южного федерального университета, 2023. - 236 с. - ISBN 978-5-9275-4265-9. - Текст : электронный. - URL: <https://znanium.ru/catalog/product/2135836> (дата обращения: 05.06.2024). – Режим доступа: по подписке.

## **9. ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ, ИСПОЛЬЗУЕМЫЕ ПРИ ОСУЩЕСТВЛЕНИИ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА**

### **9.1 Программное обеспечение**

При осуществлении образовательного процесса магистрантами и профессорско-преподавательским составом используется следующее лицензионное программное обеспечение:

4. OS Microsoft Windows (OVS OS Platform)
5. MS Office (OVS Office Platform)
6. Adobe Acrobat Professional 11.0 MLP AOO License RU
7. Adobe CS5.5 Design Standart Win IE EDU CLP
8. ABBYY FineReader 11 Corporate Edition
9. ABBYY Lingvo x5
10. Adobe Photoshop Extended CS6 13.0 MLP AOO License RU
11. Adobe Acrobat Reader DC /Pro – бесплатно
12. Яндекс. Браузер
13. Google Chrome – бесплатно
14. Opera – бесплатно
15. Mozilla – бесплатно
16. VLC – бесплатно

### **9.2 Перечень информационно-справочных систем и профессиональных баз данных информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины:**

#### **Ресурсы системы федеральных образовательных порталов**

1. Гарант.Ру. Информационно-правовой портал: <http://www.garant.ru>
2. Информационная система «Единое окно доступа к образовательным ресурсам»: <http://window.edu.ru/>
3. Открытое образование. Ассоциация «Национальная платформа открытого образования»: <http://npoed.ru>
4. Официальная Россия. Сервер органов государственной власти Российской Федерации: <http://www.gov.ru>
5. Официальный интернет-портал правовой информации. Государственная система правовой информации: <http://pravo.gov.ru>
6. Правовой сайт КонсультантПлюс: <http://www.consultant.ru/sys>
7. Российское образование. Федеральный портал: <http://www.edu.ru>

#### **Профессиональные базы данных информационно-телекоммуникационной сети «Интернет»:**

1. Национальная электронная библиотека НЭБ: <http://www.rusneb.ru>
2. Президентская библиотека: <http://www.prlib.ru>
3. Российская государственная библиотека: <http://www.rsl.ru/>

4. Российская национальная библиотека: <http://www.nlr.ru/poisk/>

### 9.3 Лицензионные электронные ресурсы библиотеки Университета

Полный перечень доступных обучающимся профессиональных баз данных представлен на официальном сайте Университета <https://eusp.org/library/electronic-resources>, включая следующие базы данных:

1. East View – 100 ведущих российских журналов по гуманитарным наукам (архив и текущая подписка): <https://dlib.eastview.com/browse>;
2. eLIBRARY.RU — Российский информационно-аналитический портал в области науки, технологии, медицины и образования, содержащий рефераты и полные тексты научных статей и публикаций, наукометрическая база данных: <http://elibrary.ru>;
3. Университетская информационная система РОССИЯ — база электронных ресурсов для учебных программ и исследовательских проектов в области социально-гуманитарных наук: <http://www.uirussia.msu.ru/>;
4. Электронные журналы по подписке (текущие номера научных зарубежных журналов).

#### Электронные библиотечные системы:

1. **Znanium.com** – Электронная библиотечная система (ЭБС) – <http://znanium.com/>;
2. Университетская библиотека онлайн – Электронная библиотечная система (ЭБС) – <http://biblioclub.ru/>

### 9.4 Электронная информационно-образовательная среда Университета

Образовательный процесс по дисциплине поддерживается средствами электронной информационно-образовательной среды Университета, которая включает в себя электронный учебно-методический ресурс АНООВО «ЕУСПб» — образовательный портал LMS Sakai — Sakai@EU, лицензионные электронные ресурсы библиотеки Университета, официальный сайт Университета (Европейский университет в Санкт-Петербурге [<https://eusp.org>]), локальную сеть и корпоративную электронную почту Университета, и обеспечивает:

доступ к учебным планам, рабочим программам дисциплин (модулей), практик и к изданиям электронных библиотечных систем и электронным образовательным ресурсам, указанным в рабочих программах;

фиксацию хода образовательного процесса, результатов промежуточной аттестации и результатов освоения основной образовательной программы;

формирование электронного портфолио обучающегося, в том числе сохранение работ обучающегося, рецензий и оценок за эти работы со стороны любых участников образовательного процесса;

взаимодействие между участниками образовательного процесса, в том числе синхронное и (или) асинхронное взаимодействие посредством сети «Интернет» (электронной почты и т.д.).

Каждый обучающийся в течение всего периода обучения обеспечен индивидуальным неограниченным доступом к электронным ресурсам библиотеки Университета, содержащей издания учебной, учебно-методической и иной литературы по изучаемой дисциплине.

## 10. МАТЕРИАЛЬНО-ТЕХНИЧЕСКАЯ БАЗА, НЕОБХОДИМАЯ ДЛЯ ОСУЩЕСТВЛЕНИЯ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА

В ходе реализации образовательного процесса используются специализированные многофункциональные аудитории для проведения занятий лекционного типа, занятий семинарского типа (практических занятий, лабораторных работ), групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, укомплектованные специализированной мебелью и техническими средствами обучения, служащими для представления учебной информации большой аудитории.

Проведение занятий лекционного типа обеспечивается демонстрационным оборудованием и учебно-наглядными пособиями.

Помещения для самостоятельной работы оснащены компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду организации.

Для лиц с ограниченными возможностями здоровья и инвалидов предоставляется возможность присутствия в аудитории вместе с ними ассистента (помощника). Для слабовидящих предоставляется возможность увеличения текста на экране ПК. Для самостоятельной работы лиц с ограниченными возможностями здоровья в помещении для самостоятельной работы организовано одно место (ПК) с возможностями бесконтактного ввода информации и управления компьютером (специализированное лицензионное программное обеспечение – Camera Mouse, веб камера). Библиотека университета предоставляет удаленный доступ к электронным ресурсам библиотеки Университета с возможностями для слабовидящих увеличения текста на экране ПК. Лица с ограниченными возможностями здоровья могут при необходимости воспользоваться имеющимся в университете креслом-коляской. В учебном корпусе имеется адаптированный лифт. На первом этаже оборудован специализированный туалет. У входа в здание университета для инвалидов оборудована специальная кнопка, входная среда обеспечена информационной доской о режиме работы университета, выполненной рельефно-точечным тактильным шрифтом (азбука Брайля).