

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ  
РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное бюджетное образовательное учреждение  
высшего образования  
«Ростовский государственный экономический университет (РИНХ)»

УТВЕРЖДЕНА  
Ученым советом ФГБОУ ВО «РГЭУ (РИНХ)»  
(протокол № 2 от 29.09.2020)  
Председатель ученого совета – ректор  
Е. Н. Макаренко



ДОПОЛНИТЕЛЬНАЯ ПРОФЕССИОНАЛЬНАЯ ПРОГРАММА -  
ПРОГРАММА ПОВЫШЕНИЯ КВАЛИФИКАЦИИ

***«Квантитативные методы, искусственный интеллект  
и информационные технологии в лингвистических  
исследованиях»***

72 час.

## ОБЩАЯ ХАРАКТЕРИСТИКА ПРОГРАММЫ

### 1. Цель программы:

- формирование лингвоинформационной компетентности у слушателей;
- ознакомление с принципами организации информационно-поисковой деятельности, направленной на совершенствование профессиональных умений в области лингвистического образования, переводческой и консультативно-коммуникативной деятельности.
- освоение средств обработки русскоязычных и иноязычных текстов в производственно-практических целях;
- обеспечение основ экспертного анализа программных продуктов лингвистического профиля;
- формирование умений по разработке лингвистических материалов с использованием современных информационных ресурсов и технологий, баз данных, словарей, словарей.

### 2. Планируемые результаты обучения:

#### 2.1. Знания:

2.1.1. основных понятий квантитативной лингвистики, стилеметрии, атрибуции; принципов построения частотных словарей; принципов работы с экспертными программами в области психолингвистики и корпусными технологиями.

2.1.2. принципов и правил создания и редактирования текстов профессионального назначения.

2.1.3. современных технологий сбора, обработки и интерпретации полученных экспериментальных данных.

2.1.4. принципов и приемы освоения инновационных областей и новых методов исследования

#### 2.2. Умения:

2.2.1. применять новые информационные технологии в процессе изучения иностранного языка и профессиональной переводческой деятельности.

2.2.2. создавать и редактировать тексты профессионального назначения.

2.2.3. применять современные технологии сбора, обработки и интерпретации полученных экспериментальных данных.

2.2.4. самостоятельно приобретать и использовать в исследовательской и практической деятельности новые знания и умения, расширять и углублять собственную научную компетентность.

#### 2.3. Навыки

2.3.1. владение основами применения количественных методов в лингвистической практике в условиях информационной среды.

2.3.2. владение навыками создания и редактирования текстов профессионального назначения.

2.3.3. возможность применять современные технологии сбора, обработки и интерпретации полученных экспериментальных данных.

2.3.4. владение способностью к самостоятельному освоению инновационных областей и новых методов исследования.

### 3. Категория слушателей:

2.3. Образование: высшее.

2.4. Квалификация: ограничения не устанавливаются

2.5. Наличие опыта профессиональной деятельности не требуется.

### 4. Учебный план программы «Квантитативные методы, искусственный интеллект и информационные технологии в лингвистических исследованиях»

№ п/п	Модуль	Всего, час	Виды учебных занятий		
			Лекции	Практические занятия	Самостоятельная работа
	Входной контроль	2	-	-	2 (Тест)
1	Модуль 1. «Квантитативная лингвистика как направление общего языкознания. Основные направления исследований в области квантитативной лингвистики. Ключевые понятия квантитативной лингвистики»	24	6	10	8
2	Модуль 2. «Компьютерные технологии в лингвистических исследованиях: работа с электронными словарями, корпусами текстов»	24	6	2	16

3	Модуль 3. «Статистический анализ текста. Атрибуция. Информационные технологии в обработке текстов»	20	4	2	14
	Итоговая аттестация	2	-	-	2 (Тест)

#### 5. Календарный план-график реализации образовательной программы

№ п/п	Наименование учебных модулей	Трудоёмкость (час)	Сроки обучения
	Входной контроль	2	02.11.2020
1	Модуль 1 «Квантитативная лингвистика как направление общего языкознания. Основные направления исследований в области квантитативной лингвистики. Ключевые понятия квантитативной лингвистики»	24	02-05-07.11.2020
2	Модуль 2 «Компьютерные технологии в лингвистических исследованиях: работа с электронными словарями, корпусами текстов»	24	09-13.11.2020
3	Модуль 3 «Статистический анализ текста. Атрибуция. Информационные технологии в обработке текстов»	20	14-16.11.2020
	Итоговая аттестация	2	16.11.2020
<b>Всего:</b>		<b>72</b>	<b>15 дней</b>

#### 6. Учебно-тематический план программы «Квантитативные методы, искусственный интеллект и информационные технологии в лингвистических исследованиях»

№ п/п	Модуль / Тема	Всего, час	Виды учебных занятий			Формы контроля
			Лекции	Практические занятия	Самостоятельная работа	
	<b>Входной контроль</b>	2	-	-	2	<b>Тестирование</b>
1.	<b>Модуль 1. Квантитативная лингвистика как направление общего языкознания. Основные направления исследований в области квантитативной лингвистики. Ключевые понятия квантитативной лингвистики</b>	24	6	10	8	
1.1	Понятие «квантитативная лингвистика» и взаимосвязь дисциплины с другими науками. Характеристика квантитативных методов в лингвистических исследованиях и их роль	2	2	-	-	
1.2	Понятие, цели и единицы лингвостатистического анализа. Методика сбора информации для лингвостатистического анализа. Минимально-необходимые статистические инструменты: частота, генеральная и выборочная совокупности	10	2	4	4	
1.3	Проблема репрезентативности лингвистической выборки. Выборочная частота, средняя частота и отклонение от средней частоты. Проблема определения рационального объема выборки.	12	2	6	4	
2.	<b>Модуль 2. Компьютерные технологии в лингвистических исследованиях: работа с</b>	24	6	2	16	

	<b>электронными словарями, корпусами текстов</b>					
2.1	Количественные и статистические методы в исследовании лексики. Закон Ципфа-Мандельброта. Закономерности функционирования лексики. Частотные словари, модель «ранг-частота	12	2	2	8	
2.2	Корпусная лингвистика как наука. Понятие «корпус текста. Классификация корпусов	2	2	-	-	
2.3	Национальные корпуса текстов. Особенности применения корпусов текстов.	10	2	-	8	
3.	<b>Модуль 3. Статистический анализ текста. Атрибуция. Информационные технологии в обработке текстов</b>	20	4	2	14	
3.1	Понятие стилеметрии и атрибуции. Область их применения. Аспекты процесса атрибуции текста, идентификация авторства текстов	2	2	-	-	
3.2	Обзор программного обеспечения для идентификации авторства текстов	4	-	-	4	
3.3	Искусственный интеллект и обработка текстов на естественных языках. Системы автоматической обработки текстов	4	-	-	4	
3.4	Автоматическое реферирование и аннотирование текстов. Глубинный анализ текстов (TextMining). Контент-анализ. Области применения структурно-вероятностной модели языка. Программное обеспечение для идентификации авторства текстов	10	2	2	6	
	<b>Итоговая аттестация</b>	2	-	-	2	<b>Тестирование</b>
	<b>Всего</b>	<b>72</b>	<b>16</b>	<b>14</b>	<b>42</b>	

**7. Учебная (рабочая) программа повышения квалификации «Квантитативные методы, искусственный интеллект и информационные технологии в лингвистических исследованиях»**

**Модуль 1. Квантитативная лингвистика как направление общего языкознания. Основные направления исследований в области квантитативной лингвистики. Ключевые понятия квантитативной лингвистики (24 час.)**

**Тема 1.1 Понятие «квантитативная лингвистика и взаимосвязь дисциплины с другими науками. Характеристика квантитативных методов в лингвистических исследованиях и их роль» (2 час.)**

Определение лингвистической дисциплины «квантитативная лингвистика», а также ее взаимосвязь с другими науками. Характеристика качественных и количественных квантитативных методов в лингвистических исследованиях, описывается их роль в проведении лингвистических исследований.

**Тема 1.2 Понятие, цели и единицы лингвостатистического анализа. Методика сбора информации для лингвостатистического анализа. Минимально-необходимые статистические инструменты: частота, генеральная и выборочная совокупности (10 час.)**

Понятие лингвостатистического анализа, разъясняются его цели, описываются единицы лингвостатистического анализа, описывается методика сбора информации для лингвостатистического анализа, перечисляются минимально-необходимые статистические инструменты: частота, генеральная и выборочная совокупности.

**Тема 1.3 Проблема репрезентативности лингвистической выборки. Выборочная частота, средняя частота и отклонение от средней частоты. Проблема определения рационального объема выборки. (12 час.)**

Проблема репрезентативности лингвистической выборки, дается понятие выборочной частоты, средней частоты и отклонение от средней частоты, ставится вопрос об определении рационального объема выборки и рассматриваются пути решения данной задачи.

**Модуль 2. Компьютерные технологии в лингвистических исследованиях: работа с электронными словарями, корпусами текстов (24 час.)**

**Тема 2.1. Количественные и статистические методы в исследовании лексики. Закон Ципфа-Мандельброта. Закономерности функционирования лексики. Частотные словари, модель "ранг-частота" (12 час.)**

Количественные и статистические методы в исследовании лексики, суть закона Ципфа-Мандельброта, закономерности функционирования лексики, понятия "частотный словарь", модель "ранг-частота".

**Тема 2.2 Корпусная лингвистика как наука. Понятие «корпус текста. Классификация корпусов (2 час.)**

Определение понятий "корпусная лингвистика", "корпус текста". Путь становления и развития корпусной лингвистики. Классификация лингвистических корпусов.

**Тема 2.3 Национальные корпусы текстов. Особенности применения корпусов текстов. (10 час.)**

Понятие "национальный корпус" текстов, особенности устройства, назначения и применения корпусов текстов.

**Модуль 3. Статистический анализ текста. Атрибуция. Информационные технологии в обработке текстов (20 час)**

**Тема 3.1 Понятие стилеметрии и атрибуции. Область их применения. Аспекты процесса атрибуции текста, идентификация авторства текстов (2 час)**

Определение понятий "стилеметрия" и "атрибуция", область их применения и аспекты процесса атрибуции текста. Этапы процесса идентификации авторства текстов.

**Тема 3.2 Обзор программного обеспечения для идентификации авторства текстов (самостоятельная работа 4 час.)**

Обзор программного обеспечения для идентификации авторства текстов, рассматриваются особенности работы с программами автоматического анализа текстов.

**Тема 3.3 Искусственный интеллект и обработка текстов на естественных языках. Системы автоматической обработки текстов (самостоятельная работа 4 час.)**

Понятие "искусственный интеллект", возможности обработки текстов на естественных языках. Обзор современных систем автоматической обработки текстов.

**Тема 3.4 Автоматическое реферирование и аннотирование текстов. Глубинный анализ текстов (TextMining). Контент-анализ. Области применения структурно-вероятностной модели языка. Программное обеспечение для идентификации авторства текстов (10 час.)**

Автоматическое реферирование и аннотирование текстов, понятие глубинного анализа текстов (TextMining); основы контент-анализа; обзор областей применения структурно-вероятностной модели языка; возможности программного обеспечения для идентификации авторства текстов.

## Описание практико-ориентированных заданий и кейсов

Номер темы/модуля	Наименование практического занятия	Описание
Модуль 1.	<p><b>Квантитативная лингвистика как направление общего языкознания. Основные направления исследований в области квантитативной лингвистики. Ключевые понятия квантитативной лингвистики</b></p>	<p style="text-align: center;"><b>Описание</b></p>
1.2.	<p>Понятие, цели и единицы лингвостатистического анализа. Методика сбора информации для лингвостатистического анализа. Минимально-необходимые статистические инструменты: частота, генеральная и выборочная совокупности</p>	<p><b>Задание 1</b>  Ознакомьтесь с дефинициями введённых терминов по опорным материалам Практикума. Обратите внимание на необходимые и достаточные признаки определяемых понятий.  Дайте подробный ответ на вопрос: Какие основные направления исследований в области квантитативной лингвистики Вам известны? Дайте краткие ответы на вопросы: Каким сокращением обозначается язык автоматической разметки текста? Какая русскоязычная энциклопедия является наиболее популярной.  Рассмотрите возможности использования методов лингвостатистического анализа для решения Ваших исследовательских задач (в Вашей профессионально сфере).</p> <p><b>Задание 2</b>  Темы для самостоятельного рассмотрения:  1. Модели речевой коммуникации в когнитивной науке: модели порождения речи.  2. Моделирование речевой деятельности в когнитивной науке: коннекционистские модели.  3. Модели восприятия речи.  4. Моделирование коммуникации на естественном языке в мультимодальных и мультимедийных системах.</p>
1.3.	<p>Проблема репрезентативности лингвистической выборки. Выборочная частота, средняя частота и отклонение от средней частоты. Проблема определения рационального объема выборки.</p>	<p><b>Задание 1. Отработка процедуры лингвостатистического анализа.</b> Выполните следующие действия на примере конкретного авторского текста, обоснуйте полученные результаты. Результаты представьте в виде файла программы «MS Word».</p> <p>1. Пусть в нашем распоряжении оказалось какое-то количество произведений одного автора. Для удобства упорядочьте их хронологически (т.е. в порядке написания) и для краткости назовите получившуюся генеральную лингвистическую совокупность текстом «данного автора». Таким образом, текст автора (в данном определении) может состоять из нескольких различных произведений — романов, повестей, рассказов и т.п.</p> <p>2. Выделите из этого текста отдельные фрагменты (выборки одинакового объема или выборочные лингвистические совокупности), состоящие из одного и того же количества слов (фиксированного заранее). Это количество слов естественно назвать объемом выборки. Эти равновеликие (равные по объему) выборки выделяйте из текста через равные интервалы, т.е. таким образом, чтобы каждые две</p>

		<p>соседние выборки были отделены друг от друга примерно одним и тем же количеством слов. Это «расстояние», интервал между соседними выборками называют шагом. Объем выборок и их шаг можно варьировать в зависимости от поставленных задач.</p> <p>Итак, последовательно двигаясь по тексту одного автора, через каждые, например, 10 страниц стандартного книжного текста делайте выборки одного и того же объема, например, в 2000 слов. Чем длиннее исследуемый текст, тем больше выборок вы сможете сделать. Для коротких произведений число выборок будет невелико, что усложняет анализ, делает результаты неустойчивыми.</p> <p>3. Выберите какой-либо лингвистический параметр, например, частоту употребления писателем предлога «в». Изучите эволюцию этого параметра вдоль всего текста, состоящего, быть может, из нескольких отдельных произведений, выстроенных нами в ряд. Для этого сделайте последовательные выборки и подсчитайте для каждой из них значение интересующего вас лингвистического параметра. В результате для каждой выборки (порции) получим свое число. От выборки к выборке оно будет меняться.</p> <p>4. Постройте график, отложив по горизонтали целые числа 1, 2, 3 и т.д., являющиеся номерами последовательных выборок, а по вертикали — значения изучаемой нами лингвистической характеристики.</p> <p>В результате эволюция данного параметра вдоль всего исследуемого текста изобразится некоторой ломаной линией. Она наглядно показывает поведение исследуемого параметра вдоль произведений данного автора. Такие графики очень удобны при поиске характерных черт данного автора — авторских инвариантов.</p> <p>5. Теперь задача может быть переформулирована так: требуется найти такой лингвистический параметр и такой оптимальный объем выборок, чтобы соответствующие им графики изображались бы для каждого автора практически горизонтальными линиями (прямыми), т.е. слабо колеблющимися ломаными. Другими словами, это будет означать, что числовые значения найденного инварианта мало отклоняются от своего среднего значения вдоль произведений каждого отдельного автора. Это явление — сглаживание ломаной кривой и ее стремление к горизонтальной прямой — назовем стабилизацией лингвистического параметра. Эта черта может быть названа характерной чертой данного автора или авторским инвариантом. Такая исследовательская находка, возможно, ляжет в основу вашего научного проекта.</p>
<p><b>Модуль 2.</b></p>	<p><b>Компьютерные технологии в лингвистических исследованиях: работа с электронными словарями, корпусами текстов</b></p>	

2.1.	<p>Количественные и статистические методы в исследовании лексики.  Закон Ципфа-Мандельброта.  Закономерности функционирования лексики.  Частотные словари, модель «ранг-частота»</p>	<p>Задание 1  а) ознакомьтесь с опорными материалами соответствующего раздела Практикума; б) рассмотрите возможности использования словарей разных видов для решения Ваших исследовательских задач.  Задание 2. Интернет-ралли по теме «Национальные корпусы текстов». Вам предложены ссылки на наиболее известные национальные корпусы различных языков. Ознакомьтесь с предложенными ресурсами и ответьте на следующие вопросы по каждому корпусу.  1. Какое количество слов содержит данный национальный корпус?  2. Какие подкорпусы включены в данный национальный корпус?  3. Каков режим доступа к данному национальному корпусу?  4. Каковы очевидные преимущества данного национального корпуса?  5. Каковы очевидные недостатки данного национального корпуса?  Результаты работы представьте в виде файла программы «MS Word».  Список национальных корпусов:  1. Британский национальный корпус (<a href="http://www.natcorp.ox.ac.uk/">www.natcorp.ox.ac.uk/</a> или <a href="http://sara.natcorp.ox.ac.uk/">http://sara.natcorp.ox.ac.uk/</a>).  2. Венгерский национальный корпус (<a href="http://corpus.nyttud.hu/mnsz">corpus.nyttud.hu/mnsz</a>).  3. Корпус испанского языка (исторический) (<a href="http://www.corpusdelespanol.org">www.corpusdelespanol.org</a>).  4. Корпус современного датского языка (<a href="http://www.korpus2000.dk">www.korpus2000.dk</a>).  5. Корпус современного итальянского языка CORIS/CODIS (<a href="http://www.cilta.unibo.it/ricerca.htm">www.cilta.unibo.it/ricerca.htm</a>).  6. Корпус современного китайского языка (LIVAC Synchronous Corpus) (<a href="http://www.rcl.cityu.edu.hk/livac">www.rcl.cityu.edu.hk/livac</a>).  7. Мангеймский корпус немецкого языка (Institut für Deutsche Sprache, Mannheim, Germany) (<a href="http://corpora.ids-mannheim.de/~cosmas">corpora.ids-mannheim.de/~cosmas</a>).  8. Национальный корпус словенского языка (<a href="http://www.fida.net/eng/">www.fida.net/eng/</a>).</p>
2.3.	<p>Национальные корпусы текстов.  Особенности применения корпусов текстов.</p>	<p>Изучите основные аспекты работы с Национальным корпусом русского языка (НКРЯ). Ознакомьтесь с основными аспектами работы с Национальным корпусом русского языка (НКРЯ). Осуществите пошагово предложенные действия.</p>
Модуль 3	<p>Статистический анализ текста.  Атрибуция. Информационные технологии в обработке текстов</p>	



3.2.	Обзор программного обеспечения для идентификации авторства текстов	<p>Темы для самостоятельного рассмотрения:  <b>Машинный (автоматический) перевод и его программное обеспечение.</b>  Ресурсы Интернета: программы анализа и лингвистической обработки текстов.  Возможности моделирования речевой коммуникации на основе материалов базы данных CHILDES.  Обзор программ обработки звукового сигнала.  Программа PRAAT.  Лингвистическая экспертиза текста.</p>
3.3.	Искусственный интеллект и обработка текстов на естественных языках. Системы автоматической обработки текстов	<p>Работа с системой интеллектуального поиска и анализа научных публикаций «ExactusExpert».  Пройдите по ссылке <a href="http://expert.exactus.ru">http://expert.exactus.ru</a>, зарегистрируйтесь в системе. Ознакомьтесь с основами работы системы, скачав руководство пользователя.</p>
3.4.	<p>Автоматическое реферирование и аннотирование текстов.  Глубинный анализ текстов (TextMining). Контент-анализ.  Области применения структурно-вероятностной модели языка.  Программное обеспечение для идентификации авторства текстов</p>	<p>Задание 1. Ознакомьтесь с опорными материалами соответствующего раздела Практикума; б) рассмотрите возможности использования материалов корпусов текстов русского и иностранного языка для решения Ваших исследовательских задач.  Задание 2  1. Перейдите по ссылке <a href="http://dict.ruslang.ru/freq.php">dict.ruslang.ru/freq.php</a>. Изучите руководство пользователя словаря и дайте ответьте на нижеследующие вопросы (результаты представьте в виде файла в формате «MS Word»):  а) Дайте определение коэффициенту вариации D?  б) Какую информацию можно извлечь об употреблении слова в разные временные периоды?  в) По какой причине имена собственные и аббревиатуры выделяются в отдельный список?  г) Объясните суть словаря значимой лексики, и что в нем означают показатели frq1, frq2 и LL-score?  Отработайте базовые навыки работы с «Новым частотным словарем русской лексики» (результаты представьте в виде файла программы «MS Word»):  а) Найдите в словаре определенное лично слово; слова, начинающиеся на «пере-»; слова, оканчивающиеся на «-енько».  б) Найдите данную информацию о любом слове русского языка: общее количество употреблений леммы, частотный ранг, число текстов, в которых встречалось слово, коэффициент вариации D, распределение употребления слова в текстах.  в) Получите информацию об использовании трех словоформ выбранного слова.  г) Выделите список 100 самых частотных глаголов в русском языке.  2. Знакомство с работой программы «Wordstat».  Программа «Wordstat» предназначена для статистического анализа текстов. Обработать можно любой текст, предварительно сохранив его в формате txt или html. В результате работы программы пользователь получает список слов из заданного текста с указанием частоты их употребления в заданном тексте. Ознакомьтесь с основными принципами работы, о результатах доложите в ходе групповой дискуссии.</p>

8. **Оценочные материалы по образовательной программе**  
**8.1. Вопросы аттестации**

Вопросы входного тестирования	Вопросы промежуточного тестирования	Вопросы итогового тестирования
<p>1. _____ - это раздел общей лингвистики и, в частности, математической лингвистики, исследующей язык при помощи статистических методов и имеющей своей конечной целью формулирование законов функционирования языка и построение общей теории языка.            А) Когнитивная лингвистика            В) Квантитативная лингвистика            С) Компьютерная лингвистика            D) Нейролингвистика</p> <p>2. _____ - это массив текстов, собранных в единую систему по определенным признакам (языку, жанру, времени создания текста, автору и т.д.) и снабженных поисковой системой.            А) Энциклопедический словарь            В) Лингвистический корпус            С) Тезаурус            D) Электронная библиотека</p> <p>3. Методы _____ используются не только в цифровой гуманитарке, но и в компьютерных системах, помогающих переводчикам.            А) Психоллингвистики            В) Квантитативной лингвистики            С) Корпусной лингвистики            D) Контрастивной лингвистики</p> <p>4. _____ - это набор слов языка с указанием информации о частоте их встречаемости.            А) Толковый словарь            В) Частотный словарь            С) Этимологический словарь            D) Фразеологический словарь</p> <p>5. Как называется программа для формирования конкорданса?            А) Текстовый редактор            В) Издательская система            D) Текстовый процессор            С) Конкордансер</p> <p>6. _____ - метод качественно-количественного анализа содержания документов с целью выявления и оценки специфических характеристик текстов.</p>	<p>Промежуточная аттестация по модулям не предусмотрена</p>	<p>1. _____ — несплошное наблюдение, при котором исследуется не вся совокупность языковых единиц, а лишь определенная часть.            А) Выборочное наблюдение            В) Случайное наблюдение</p> <p>2. _____ — это функция, приближающая плотность вероятности некоторого распределения, построенная на основе выборки из него.            А) Граф            В) Гистограмма</p> <p>3. _____ — вероятность, с которой доверительный интервал «захватит» истинное значение параметра генеральной совокупности.            А) Кумулята            В) Доверительная вероятность</p> <p>4. _____ — интервал вокруг выборочного значения определенного параметра, который с заданной (достаточно высокой) вероятностью покрывает истинное значение параметра. Величина доверительного интервала зависит как от доверительной вероятности (т.е. надежности оценивания), так и от объема выборки.            А) Доверительный интервал            В) Выборочный интервал</p> <p>5. Закон _____ — закономерность функционирования лексики, сутью которой является то, что чем шире сфера употребления слова, тем беднее его содержание.            А) Крушевского-Куриловича            В) Мандельброта</p> <p>6. Закон _____ — эмпирическая закономерность распределения частоты слов естественного языка, суть которого: если к какому-либо достаточно большому тексту составить список всех встретившихся в нем слов, расположить эти слова в порядке убывания частоты их встречаемости в данном тексте и пронумеровать в порядке от 1 (номер наиболее часто встречающегося слова) до <math>g</math>, то для любого слова произведение его порядкового номера-ранга (<math>g</math>) в таком списке и частоты встречаемости слова в тексте (<math>f</math>) будет величиной постоянной,</p>

<p>A) Семантический анализ  B) Контент-анализ  C) Лингвостилистический анализ  D) Контекстный анализ</p> <p>7. _____ - информационная модель предметной области, имеющая вид ориентированного графа, вершины которого соответствуют объектам предметной области, а дуги (рёбра) задают отношения между ними. Объектами могут быть: понятия, события, свойства, процессы.  A) Фрейм  B) Контентная сеть  C) Социальная сеть  D) Семантическая сеть</p> <p>8. _____ - это список всех употреблений заданного языкового выражения (например, слова) в контексте, возможно, со ссылками на источник.  A) Словарная статья  B) Конкорданс  D) Толковый словарь  C) Тезаурус</p> <p>9. _____ - процесс определения подлинности или подложности рукописного текста и установление его автора.  A) Стилистический анализ  B) Атрибуция текста  C) Контент-анализ  D) Герменевтический анализ</p> <p>10. _____ — процесс перевода текстов с одного естественного языка на другой при помощи специализированной компьютерной программы, направление научных исследований, связанных с построением подобных систем.  A) Последовательный перевод  B) Машинный перевод  C) Коммунальный перевод  D) Авторизированный перевод</p>		<p>имеющей примерно одинаковое значение для любого словарного списка.  A) Ципфа  B) Крушевского-Куриловича</p> <p>7. _____ — отнесение объектов базы данных к заранее определенным категориям.  A) Классификация  B) Кластеризация</p> <p>8. _____ — выделение из базы данных компактных подгрупп объектов с близкими свойствами.  A) Кластеризация  B) Классификация</p> <p>9. Метод _____ анализа квазисинонимичных лексем — метод атрибуции произведений, который заключается в выявлении авторских предпочтений в выборе из группы квазисинонимов — близких по значению слов или устойчивых словосочетаний (фразеологизмов).  A) количественного  B) качественного</p> <p>10. _____ (систематическая) выборка — разновидность случайной выборки, упорядоченная по какому-либо признаку. Первый элемент отбирается случайно, затем, с шагом <math>n</math> отбирается каждый <math>k</math> элемент. Размер генеральной совокупности, при <math>m \ N = nk</math>.  A) Механическая  B) Автоматическая</p> <p>11. Модель _____ слова А.А. Поликарпова — тенденция к увеличению степени абстрактности значений у многозначного слова — чем позже возникает значение, тем оно более абстрактно.  A) жизненного цикла  B) изменения смысла</p> <p>12. _____ (степень точности) — величина, которая характеризует ширину доверительного интервала, в который попадает относительная частота исследуемого свойства лингвистической единицы.  A) Относительная величина  B) Относительная ошибка</p> <p>13. _____ — характеристика какого-либо параметра в выборке; является не точным, а приближенным значением этого же параметра в генеральной совокупности.</p>
--	--	--

		<p>A) Оценка B) Определение</p> <p>14. _____ — один из способов графического представления плотности вероятности случайной величины. Представляет собой ломаную, соединяющую точки, соответствующие средним значениям интервалов группировки и частотам этих интервалов. A) Гистограмма B) Полигон частот</p> <p>15. _____ — выборка, при которой вся генеральная совокупность разбивается на серии, гнезда. Затем производят случайный или механический отбор, единицами которого выступают не сами объекты, а группы (кластеры или гнезда). Объекты внутри групп обследуются в очередности расположения. A) Кластерная выборка B) Случайная выборка</p> <p>16. _____ — двусторонняя единица, представленная внешне (цепочка фонем, ударение) и внутренне (значение слова). A) Словарная статья B) Словоформа</p> <p>17. _____ — такая выборка, которая предполагает однородность генеральной совокупности, одинаковую вероятность доступности всех элементов, наличие полного списка всех элементов. A) Случайная выборка B) Сплошная выборка</p> <p>18. _____ - вид словаря, в котором приводятся числовые характеристики употребительности слов (словоформ, словосочетаний) какого-либо языка, в том числе языка писателя, какого-либо произведения. A) Частотный словарь B) TranslationMemory</p> <p>19. _____ собрание текстов в электронной форме, представляющих данный язык на определенном этапе его существования. A) Национальный корпус B) Массив лингвистических данных</p> <p>20. _____ - большой, представленный в машиночитаемом виде, унифицированный, структурированный, размеченный, филологически компетентный массив языковых данных, предназначенный для решения конкретных лингвистических задач.</p>
--	--	---

		<p>А) Лингвистический, языковой корпус текстов В) База данных</p> <p>21. _____ — направление лингвистики, в рамках которого изучаются и эксплицируются лингвистические явления с помощью методов количественной математики (теория вероятностей, математическая статистика, теория информации, математический анализ и др.). А) Квантитативная лингвистика В) Когнитивная лингвистика</p> <p>22. _____ — численная мера возможности наступления некоторого события. А) вероятность или вероятностная мера В) лингвистическая частота</p> <p>23. _____ - часть генеральной совокупности, объединенная общим признаком совокупность языковых единиц, выбранная из генеральной совокупности при проведении выборочного наблюдения. А) выборочная совокупность В) выборочная лингвистическая совокупность</p> <p>24. _____ — часть объектов генеральной совокупности, отобранная с помощью специальных приемов для получения надежной информации обо всей генеральной совокупности. А) выборочная совокупность В) выборочная лингвистическая совокупность</p> <p>25. _____ - несплошное наблюдение, при котором исследуется не вся совокупность языковых единиц, а лишь определенная часть. А) контент-анализ В) выборочное наблюдение</p> <p>26. _____ - частоты, показанные отдельными лингвистическими выборками. А) выборочные частоты В) ключевые слова</p> <p>27. _____ - множество всех возможных однородных объектов, обладающих признаками (признаком), составляющими предмет анализа. А) Генеральная совокупность В) Генеральная лингвистическая совокупность</p> <p>28. _____ - совокупность однородных лингвистических объектов (лингвистических единиц), обладающих</p>
--	--	--

		<p>признаком/признаками, составляющим/и предмет лингвистического анализа.  А) Генеральная совокупность  В) Генеральная лингвистическая совокупность</p> <p>29. _____ -вероятность, с которой доверительный интервал «захватит» истинное значение параметра генеральной совокупности.  А) доверительная вероятность  В) доверительный интервал</p> <p>30. _____ — интервал вокруг выборочного значения определенного параметра, который с заданной (достаточно высокой) вероятностью покрывает истинное значение параметра. Величина доверительного интервала зависит как от доверительной вероятности (т.е. надежности оценивания), так и от объема выборки.  А) доверительная вероятность  В) доверительный интервал</p> <p>31. _____ — раздел количественной лингвистики, целью которого является обнаружение факторов порядка в лексике, на основе ранжирования ее по системной и функциональной значимости.  А) Количественная лексикология  В) Когнитивная лингвистика</p> <p>32. _____ — термин, относящийся к основному содержанию текста и повторяющийся в нем несколько раз (с учетом всех возможных синонимов).  А) Ключевое (опорное) слово  В) Кумулята</p> <p>33. _____ — закономерность функционирования лексики, сутью которой является то, что слова, часто встречающиеся вместе в определенном отрезке текста, теснее связаны между собой по смыслу, чем слова, реже появляющиеся рядом в том же отрезке текста.  А) Закон семантических связей  В) Закон Ципфа</p> <p>34. _____ - индивидуальный авторский стиль писателя или поэта, специфика речи художественных произведений, предмет изучения стилистики.  А) Идиостиль  В) Стилиевая характеристика</p> <p>35. _____ - наука и технология создания интеллектуальных машин,</p>
--	--	---

		<p>особенно интеллектуальных компьютерных программ. Искусственный интеллект связан со сходной задачей использования компьютеров для понимания человеческого интеллекта.</p> <p>А) Компьютерная лингвистика В) Искусственный интеллект</p> <p>36. _____ - качественно-количественный метод изучения документов, которое характеризуется объективностью выводов и строгостью процедуры и состоит из квантификационной обработки текста с дальнейшей интерпретацией результатов.</p> <p>А) контент-анализ В) стилистический анализ</p> <p>37. _____ — раздел компьютерной лингвистики, занимающийся разработкой общих принципов построения и использования лингвистических корпусов (корпусов текстов) с применением компьютерных технологий.</p> <p>А) Корпусная лингвистика В) Квантитативная лингвистика</p> <p>38. _____ - графическое представление ряда статистических данных, базы данных</p> <p>А) Лемма В) Кумулята</p> <p>39. _____ - слово в основной, исходной форме, в качестве которой в русском языке выступает обычно именительный падеж единственного числа (для именных форм) и инфинитив (для глагольных форм).</p> <p>А) Лингвистическая единица В) Лемма</p> <p>40. _____ - большой, представленный в машиночитаемом виде, унифицированный, структурированный, размеченный, филологически компетентный массив языковых данных, предназначенный для решения конкретных лингвистических задач.</p> <p>А) Лингвистический, языковой корпус текстов В) База данных</p> <p>41. _____ - выполняемые на компьютере действия по преобразованию речи одного естественного языка в эквивалентную по содержанию речь другого, а также результаты такого действия.</p> <p>А) Машинный перевод В) Автоматический анализ текста</p>
--	--	---

		<p>42. _____ собрание текстов в электронной форме, представляющих данный язык на определенном этапе его существования. А) Национальный корпус В) Массив лингвистических данных</p> <p>43. _____ - приписывание текстам и их компонентам специальных меток. А) Разметка В) Семантическая сеть</p> <p>44. _____ — порядковый номер слова в частотном словаре (списке). А) Ранг В) Частота</p> <p>45. _____ — связный текст, который кратко выражает не только центральную тему или предмет какого-либо документа, но и цель, применяемые методы, основные результаты описанного исследования или разработки. А) Реферат В) Аннотация</p> <p>46. _____ - краткое изложение содержания документа, дающее общее представление об его теме. А) Реферат В) Аннотация</p> <p>47. _____ - соотнесение произведению соответствующих атрибутов, к которым причисляются имя создателя, жанр, время и место создания. А) Атрибуция (авторизация) В) Идиостиль</p> <p>48. _____ - перевод с использованием как традиционной, так и машинной переработки текста, где большую часть стереотипной деятельности по отысканию эквивалентов и соответствий на основании формализованных методик фреймов-сценариев осуществляет искусственный интеллект А) Смешанный перевод В) Машинный перевод</p> <p>49. _____ - прикладная филологическая дисциплина, занимающаяся измерением стилевых характеристик с целью систематизации и упорядочения (типологии, атрибуции, датировки, диагностики, реконструкции и т.д.) текстов и их частей. А) Стилеметрия В) Атрибуция</p>
--	--	---



		50. _____ - вид словаря, в котором приводятся числовые характеристики употребительности слов (словоформ, словосочетаний) какого-либо языка, в том числе языка писателя, какого-либо произведения. А) Частотный словарь В) TranslationMemory
--	--	---

### 8.2. Описание показателей и критериев оценивания, шкалы оценивания

На входном этапе контроля сформированности компетенций по тестовым заданиям (вариант состоит из 10 заданий, решение каждого задания оценивается в 1 балл) применяется аналитическая шкала оценивания:

Балл	Критерии оценивания
1	тестовое задание решено верно
0	тестовое задание решено неверно

На входном этапе контроля сформированности компетенций применяется аналитическая шкала оценивания:

Балл	Критерии оценивания
5	выставляется обучающемуся, если все задания решены верно, выводы обоснованы;
4	выставляется обучающемуся, если есть несущественные ошибки, но в целом задания решены верно, вывод в достаточной степени обоснован;
3	выставляется обучающемуся, если продемонстрировано частичное понимание проблемы. Большинство требований, предъявляемых к заданию, выполнены. При выполнении задания требовалась значительная помощь преподавателя.
2	выставляется обучающемуся, если задания решены с ошибками, исправленными после дополнительных вопросов, вывод недостаточно полон;
1	выставляется обучающемуся, если отсутствует понимание проблемы. Попытки выполнения задания были неверными.
0	выставляется обучающемуся, если нет ответа. Не было попытки решить поставленную практическую задачу.

На этапе итоговой аттестации применяется измерительная шкала оценивания сформированности компетенций.

Критерий оценки	Показатели оценки					
	0	1	2	3	4	5
% правильных ответов	0-19	≥20	≥40	≥60	≥70	≥85

### 8.3. Примеры контрольных заданий по модулям или всей образовательной программе.

*Пример аналитического задания по модулю 1.*

Задание 1. Дайте подробный ответ на вопрос: Что входит в содержание общего понятия «прикладная лингвистика»? Дайте краткие ответы на вопросы: Как называется программа автоматического поиска правого и левого контекстов слова? Какая модель базы данных включает таблицы?

Конкретизация понятия «квантитативная лингвистика» и сфер ее применения:

Используя поисковые системы Интернет, проанализируйте, какие существуют лаборатории, кафедры и институты прикладной, компьютерной, математической лингвистики в России.

Представьте результаты в виде таблицы.

Название кафедры/института/Лаборатории	Место расположения	Сфера деятельности	Достижения

## **Задание 2.**

А) Найдите средние частоты в текстах. Проведите их сравнительный анализ согласно алгоритму.

Текст 1 (Т<sub>а</sub>) имеет частоты: 95, 98, 89, 105, 102, 85, 111, 115, 93, 107.

Текст 2 (Т<sub>б</sub>) имеет частоты: 98, 112, 114, 108, 106, 122, 95, 87, 125, 133.

Определите относительную частоту употребления неисчисляемых абстрактных существительных в данной выборке.

В) Найдите средние частоты в текстах. Проведите их сравнительный анализ согласно алгоритму.

Текст 1 (Т<sub>а</sub>) имеет частоты: 96, 99, 100, 104, 103, 87, 110, 114, 95, 106.

Текст 2 (Т<sub>б</sub>) имеет частоты: 97, 111, 112, 107, 106, 124, 94, 88, 124, 131.

Определите относительную частоту употребления неисчисляемых абстрактных существительных в данной выборке:

### ***Пример аналитического задания по модулю 2.***

**Задание 1.** Создание и работа с частотным словарем к заданному тексту.

С помощью программы «Wordstat» или «Webscript» создайте частотный словарь по заданному тексту.

**Задание 2.** На основе частотного словаря проведите лексико-грамматический анализ текста. По результатам с помощью программы «MS Excel» создайте описание текста в виде таблицы, содержащей:

общее количество слов;

количество разных слов;

количество слов, встречающихся в тексте лишь один раз;

количество слов, встречающихся в тексте несколько раз;

количество повторений самого популярного слова;

количество существительных;

количество глаголов;

количество прилагательных.

**Задание 3.** Создание и работа с частотным словарем к заданному тексту. С помощью программы «Wordstat» или «Webscript» создайте частотный словарь по заданному тексту.

**Задание 4.** На основе частотного словаря проведите лексико-грамматический анализ текста. По результатам с помощью программы «MS Excel» создайте описание текста в виде таблицы, содержащей:

общее количество слов;

количество разных слов;

количество слов, встречающихся в тексте лишь один раз;

количество слов, встречающихся в тексте несколько раз;

количество повторений самого популярного слова;

количество существительных;

количество глаголов;

количество прилагательных.

### ***Пример аналитического задания по модулю 3.***

**Задание 1.** Основы работы с программой определения авторства

«Лингвоанализатор». Ознакомьтесь с возможностями программы «Лингвоанализатор», изучив статью Д. Хмелева «Как определить писателя?» (Хмелев Д. Как определить писателя? // Компьютерра. № 9. 2000. С.87-94.)

**Задание 2.** Автоматическая авторизация текста.

Вам предлагается анонимный текст, состоящий из двух частей. Заведомо известно, что эти части написаны разными авторами. Используя возможности программы «Лингвоанализатор» ([www.rusf.ru/books/analysis/](http://www.rusf.ru/books/analysis/)), определите, кто написал каждую из частей анализируемого текста. Сделайте запрос на интерпретацию результатов в научном и литературном виде.

**Задание 3.** Работа с программами автоматического определения плагиата.

Ознакомьтесь с программным продуктом «Антиплагиат», изучив ресурс «Что такое «Антиплагиат»?». Данный материал также можно найти, пройдя по ссылке <http://antiplagiat.ru/index.aspx>

**Задание 4.** Практическое использование программы «Антиплагиат». На кафедре английского языка сдан реферат. Проверьте его на оригинальность с помощью программы «Антиплагиат» ([www. antiplagiat.ru/QuickCheck.aspx](http://www.antiplagiat.ru/QuickCheck.aspx)).

**Задание 5.** Базовые аспекты технологии автоматического анализа текстовой информации. Ознакомьтесь с ресурсом «Технология автоматического анализа текстовой информации», который освещает основные аспекты работы в системе «TextAnalyst». Данный материал вы можете найти, пройдя по ссылке <http://megaputer.ru/textanalyst.php>

На основе изученного материала ответьте на следующие вопросы:

1. Из каких этапов состоит анализ текста в системе «TextAnalyst»?
2. Что такое семантическая сеть?
3. Как можно ускорить обработку большого объема информации, используя автоматическое реферирование текста?
4. Для чего нужна подсистема смыслового поиска в программе «TextAnalyst?»

Ответы оформите в виде файла программы «MS Word».

#### **8.4. Тесты и обучающие задачи (кейсы), иные практикоориентированные формы заданий**

##### **Модуль 1.**

*Пример тестового задания:*

\_\_\_\_\_ — несплошное наблюдение, при котором исследуется не вся совокупность языковых единиц, а лишь определенная часть.

- А) Выборочное наблюдение
- В) Случайное наблюдение

\_\_\_\_\_ — это функция, приближающая плотность вероятности некоторого распределения, построенная на основе выборки из него.

- А) Граф
- В) Гистограмма

\_\_\_\_\_ — вероятность, с которой доверительный интервал «захватит» истинное значение параметра генеральной совокупности.

- А) Кумулята
- В) Доверительная вероятность

#### **8.5. Описание процедуры оценивания результатов обучения**

Процедура оценивания результатов обучения зависит от типа оценочных материалов. Для тестов с множественным выбором предусмотрено автоматическое оценивание результатов. Для оценивания практических заданий применяется метод ручного оценивания.

**Входное и итоговое аттестационные испытания** включают тестовые задания с вопросами только закрытого типа.

## 9. Организационно-педагогические условия реализации программы

### 9.1. Кадровое обеспечение программы

№ п/п	Фамилия, имя, отчество (при наличии)	Место основной работы и должность, ученая степень и ученое звание (при наличии)	Ссылки на веб-страницы с портфолио (при наличии)	Фото в формате jpeg	Отметка о полученном согласии на обработку персональных данных
1	Евсюкова Татьяна Всеволодовна	РГЭУ (РИНХ), декан факультета Лингвистики и журналистики, профессор кафедры лингвистики и межкультурной коммуникации д.филол.н., профессор	<a href="https://rsue.ru/prepodavateli.php?ELEMENT_ID=292">https://rsue.ru/prepodavateli.php?ELEMENT_ID=292</a>		Согласен на обработку персональных данных
2	Казанская Евгения Владимировна	РГЭУ (РИНХ), зав. кафедрой иностранных языков для экономических специальностей, к.филол.н., доцент	<a href="https://rsue.ru/prepodavateli.php?ELEMENT_ID=699">https://rsue.ru/prepodavateli.php?ELEMENT_ID=699</a>		Согласен на обработку персональных данных

### 9.2. Учебно-методическое обеспечение и информационное сопровождение

Учебно-методические материалы	
Методы, формы и технологии	Методические разработки, материалы курса, учебная литература
- лекции в форме вебинаров - тест	Гребенщикова А.В. Основы количественной лингвистики и новых информационных технологий [Электронный ресурс]: учеб. пособие / А.В. Гребенщикова. — 2-е изд., стер. — М.: ФЛИНТА, 2015. — 152 с. Залевская А.А. Количественная лингвистика и новые информационные технологии: практикум. – Тверь: Твер. гос. ун-т, 2015. – 64 с. Потапова Р. К. Новые информационные технологии и лингвистика. 5-е изд. М.: УРСС, 2012 (или другое издание). Словарь культуры как проблема лингвокультурологии / Т. В. Евсюкова; М-во образования Рос. Федерации. Рост. гос. экон. ун-т (РИНХ). - Ростов н/Д : РИНХ, 2001. - 255 с. Котельникова Е.В. «Учебно-методическое пособие по дисциплине «Количественная лингвистика» для магистров направления «Зарубежное регионоведение». Очная форма обучения, 2 курс». – Учебно-методическое пособие – Ростов-на-Дону: Издательско-полиграфический комплекс РГЭУ (РИНХ), 2018. – 96 с.

Информационное сопровождение	
Электронные образовательные ресурсы	Электронные информационные ресурсы
Электронный учебно-методический комплекс образовательной программы размещен на портале электронного обучения РГЭУ (РИНХ)– Режим доступа: <a href="https://do.rsue.ru">https://do.rsue.ru</a>	Герд, А.С. Предмет и основные направления прикладной лингвистики [Электронный ресурс] / URL: <a href="http://project.phil.spbu.ru/lib/data/ru/heard/prikling.html">http://project.phil.spbu.ru/lib/data/ru/heard/prikling.html</a> Викизнанич <a href="http://www.wikiznaniic.ru">http://www.wikiznaniic.ru</a> ; Корпусная лингвистика // Фонд знаний «Ломоносов» [Электронный ресурс] / URL: <a href="http://www.epistemeana.ru/enc/ru/encyclopedia:01210">www.epistemeana.ru/enc/ru/encyclopedia:01210</a> Основы социолингвистики <a href="http://sociolingvistika.narod.ru">http://sociolingvistika.narod.ru</a> ;

	<p>Филология и лингвистика <a href="http://filologia.su">http://filologia.su</a>;  Языкознание.py <a href="http://yazykoznanie.ru">http://yazykoznanie.ru</a>  Электронная библиотека «Все для студента» <a href="http://www.twirpx.com/">http://www.twirpx.com/</a>;  GlossaryCommander. Служба тематических толковых словарей  <a href="http://www.glossary.ru">http://www.glossary.ru</a> ;  Рубрикон. Энциклопедии, словари, справочники  <a href="http://www.rubricon.com">http://www.rubricon.com</a> ;  Университетская библиотека online <a href="http://biblioclub.ru">http://biblioclub.ru</a> ;  Национальный корпус русского языка <a href="http://www.ruscorpora.ru">http://www.ruscorpora.ru</a> ;  Computer Science Alliance <a href="http://www.csa.ru/BAN">http://www.csa.ru/BAN</a>;  Библиотека (интернет-издательство)  <a href="http://www.magister.msk.ru/library/">http://www.magister.msk.ru/library/</a>  Селегей В. Электронные словари и компьютерная лексикография  [Электронный ресурс] / URL:  <a href="http://www.lingvoda.ru/transforum/articles/selegey_al.asp">http://www.lingvoda.ru/transforum/articles/selegey_al.asp</a>  Эпштейн В.Л. Введение в гипертекст и гипертекстовые системы  [Электронный ресурс] / URL: <a href="http://www.ipu.ru/publ/cpstn.htm">http://www.ipu.ru/publ/cpstn.htm</a>  Wikipedia <a href="http://www.wikipedia.org">http://www.wikipedia.org</a> .</p>
--	---

### 9.3. Материально-технические условия реализации программы

Вид занятий	Наименование оборудования, программного обеспечения
Лекции	ПЭВМ под управлением операционной системы Microsoft Windows, программа для видеоконференций Zoom, Wordstat, Webscript, TextAnalyst, Лингвоанализатор, Антиплагиат, ExactusExpert, PROMT
Практические занятия	
Самостоятельная работа	

## II. Паспорт компетенций (Приложение 2)

**ФГБОУ ВО «Ростовский государственный экономический университет (РИНХ)»****ПАСПОРТ КОМПЕТЕНЦИИ**

**Дополнительная профессиональная программа – программа повышения квалификации «Квантитативные методы, искусственный интеллект и информационные технологии в лингвистических исследованиях»**

1.	Наименование компетенции	<b>владение современными программными средствами статистического анализа и математического моделирования</b>	
2.	Указание типа компетенции	Профессиональная	
3.	Определение, содержание и основные сущностные характеристики компетенции	умение применять специализированное программное обеспечение, позволяющее проводить статистический анализ текстовой информации, использовать методики работы с лингвистическим материалом в профессиональных целях.	
4.	Дескриптор знаний, умений и навыков по уровням	<b>Уровни сформированности компетенции обучающегося</b>	<b>Индикаторы</b>
		Начальный уровень	<b>Знает:</b> основные понятия лингвистики и статистического анализа текста; <b>Умеет:</b> систематизировать полученные знания с целью их дальнейшего использования в профессиональной деятельности; <b>Владеет:</b> теоретическими и практическими навыками работы с текстовой информацией, предполагающим решение поставленных целей и задач.
		Базовый уровень	<b>Знает:</b> основные понятия квантитативной лингвистики, стилиметрии, атрибуции; принципы построения частотных словарей; <b>Умеет:</b> применять средства автоматизации в процессе работы с лингвистическим массивом информации; <b>Владеет:</b> основами применения количественных методов в лингвистической практике в условиях информационной среды
		Продвинутый	<b>Знает:</b> принципы работы с экспертными программами в области лингвистики и корпусных технологий; <b>Умеет:</b> использовать специализированное программное обеспечение, позволяющее проводить статистический анализ текстовой информации для достижения профессиональных целей; <b>Владеет:</b> навыками комплексного анализа процессов модернизации и трансформации в сфере профессиональных интересов.
		Профессиональный	<b>Знает:</b> особенности организации информационно-поисковой деятельности, направленной на совершенствование профессиональных умений в области лингвистического образования, переводческой и консультативно-коммуникативной деятельности. <b>Умеет:</b> эффективно применять современные программные средства статистического анализа и математического моделирования в своей профессиональной и научной деятельности;



			<b>Владеет:</b> сформированными умениями по разработке лингвистических материалов, применять современные информационные ресурсы и технологии.
5.	Характеристика взаимосвязи данной компетенции с другими компетенциями/ необходимость владения другими компетенциями для формирования данной компетенции	Владение данной компетенцией является необходимым условием для овладения компетенцией «Готовность применять современные технологии сбора, обработки и интерпретации полученных экспериментальных данных»	
6.	Средства и технологии оценки	Тесты	

## ПАСПОРТ КОМПЕТЕНЦИИ

### Дополнительная профессиональная программа – программа повышения квалификации «Квантитативные методы, искусственный интеллект и информационные технологии в лингвистических исследованиях»

1.	Наименование компетенции	<b>Способность создавать и редактировать тексты профессионального назначения</b>	
2.	Указание типа компетенции	Профессиональная	
3.	Определение, содержание и основные сущностные характеристики компетенции	Данная компетенция соотносится с обобщенной трудовой функцией и/или трудовой функцией из профессионального стандарта «Лингвистическая и текстуальная компетенции в исходном языке и языке перевода».	
4.	Дескриптор знаний, умений и навыков по уровням	<b>Уровни сформированности компетенции обучающегося</b>	<b>Индикаторы</b>
		Начальный уровень	<b>Знает:</b> основные понятия лингвистики и статистического анализа текста; <b>Умеет:</b> накапливать и применять в практической деятельности новые знания; <b>Владеет:</b> основами применения теоретических и практических навыков в профессиональной деятельности
		Базовый уровень	<b>Знает:</b> принципы и правила создания и редактирования текстов профессионального назначения; <b>Умеет:</b> создавать и редактировать тексты профессионального назначения; <b>Владеет:</b> навыками создания и редактирования текстов профессионального назначения.
		Продвинутый	<b>Знает:</b> принципы работы с профессиональными программами обработки текстов; <b>Умеет:</b> применять на практике сценарии работы с текстовым материалом; <b>Владеет:</b> развитыми навыками работы со средствами обработки русскоязычных и иноязычных текстов в производственно-практических целях.
		Профессиональный	<b>Знает:</b> сценарии морфологического, синтаксического и семантического анализа текстового материала;

			<p><b>Умеет:</b> выстраивать эффективную модель обработки больших массивов текстов профессиональной направленности;</p> <p><b>Владеет:</b> навыками выполнения ряда прототипических сценариев настройки и управления системы искусственного интеллекта в работе с тактовыми массивами.</p>
5.	Характеристика взаимосвязи данной компетенции с другими компетенциями/ необходимость владения другими компетенциями для формирования данной компетенции	Успешное освоение компетенции связано с компетенцией «Готовность применять современные технологии сбора, обработки и интерпретации полученных экспериментальных данных»	
6.	Средства и технологии оценки	Тесты	

### ПАСПОРТ КОМПЕТЕНЦИИ

**Дополнительная профессиональная программа – программа повышения квалификации «Квантитативные методы, искусственный интеллект и информационные технологии в лингвистических исследованиях»**

1.	Наименование компетенции	<b>Готовность применять современные технологии сбора, обработки и интерпретации полученных экспериментальных данных</b>	
2.	Указание типа компетенции	Профессиональная	
3.	Определение, содержание и основные сущностные характеристики компетенции	Данная компетенция соотносится с обобщенной трудовой функцией и/или трудовой функцией из профессионального стандарта «Компетенция, касающаяся сбора и обработки информации»	
4.	Дескриптор знаний, умений и навыков по уровням	<b>Уровни сформированности компетенции обучающегося</b>	<b>Индикаторы</b>
		Начальный уровень	<b>Знает:</b> основные способы сбора и обработки экспериментальных данных; <b>Умеет:</b> идентифицировать методы обработки экспериментальных данных; <b>Владеет:</b> навыками интерпретации экспериментальных данных.
		Базовый уровень	<b>Знает:</b> специфику ведения экспериментальных работ; <b>Умеет:</b> применять современные технологии сбора и обработки экспериментальных данных; <b>Владеет:</b> методикой анализа данных.
		Продвинутый	<b>Знает:</b> современные технологии сбора, обработки и интерпретации полученных экспериментальных данных. <b>Умеет:</b> применять современные технологии сбора, обработки и интерпретации полученных экспериментальных данных. <b>Владеет:</b> навыками применения современных технологий сбора, обработки и интерпретации полученных экспериментальных данных.
		Профессиональный	<b>Знает:</b> специфику работы с источниками данных; <b>Умеет:</b> осуществлять выбор технологии обработки и интерпретации полученных



			экспериментальных данных в соответствии с целью исследования; <b>Владеет:</b> способностью принимать решения, навыком осуществления научного исследования в соответствии с современными технологиями обработки и интерпретации полученных экспериментальных данных.
5.	Характеристика взаимосвязи данной компетенции с другими компетенциями/ необходимость владения другими компетенциями для формирования данной компетенции	Компетенция сопряжена с компетенцией «Способность создавать и редактировать тексты профессионального назначения»	
6.	Средства и технологии оценки	Тесты	

## ПАСПОРТ КОМПЕТЕНЦИИ

### Дополнительная профессиональная программа – программа повышения квалификации «Квантитативные методы, искусственный интеллект и информационные технологии в лингвистических исследованиях»

1.	Наименование компетенции	способность к самостоятельному освоению инновационных областей и новых методов исследования	
2.	Указание типа компетенции	Профессиональная	
3.	Определение, содержание и основные сущностные характеристики компетенции	Данная компетенция соотносится с обобщенной трудовой функцией и/или трудовой функцией из профессионального стандарта.	
4.	Дескриптор знаний, умений и навыков по уровням	<b>Уровни сформированности компетенции обучающегося</b>	<b>Индикаторы</b>
		Начальный уровень	<b>Знает:</b> основные исследовательских методов; <b>Умеет:</b> различать методы исследования; <b>Владеет:</b> способностью к освоению новых сфер профессиональной деятельности.
		Базовый уровень	<b>Знает:</b> принципы и приемы освоения инновационных областей и новых методов исследования; <b>Владеет:</b> способностью к самостоятельному освоению инновационных областей и новых методов исследования; <b>Умеет:</b> применять различные инновации и новые методы исследования в своей деятельности.
		Продвинутый	<b>Знает:</b> инновационные области исследования в лингвистике; <b>Умеет:</b> осуществлять выбор наиболее эффективных инновационных областей и новых методов исследования; <b>Владеет:</b> навыками применения инноваций и новых методов исследования в своей деятельности

		Профессиональный	<b>Знает:</b> новые методы исследования в лингвистике; <b>Умеет:</b> оценивать эффективность использования инновационных областей и новых методов исследования; <b>Владеет:</b> навыком самостоятельного освоения инновационных областей и новых методов исследования.
5.	Характеристика взаимосвязи данной компетенции с другими компетенциями/ необходимость владения другими компетенциями для формирования данной компетенции	Успешное освоение компетенции связано с компетенцией «Способность использовать понятийный аппарат философии, теоретической и прикладной лингвистики, переводоведения, лингводидактики и теории межкультурной коммуникации для решения профессиональных задач».	
6.	Средства и технологии оценки	Тесты	

**VI. Иная информация о качестве и востребованности образовательной программы -** программа реализуется впервые.

**V. Рекомендаций к программе от работодателей:** имеются 2 письма-рекомендации: ООО «МебельМаркет24»; ООО «КолдМаркет».

**VI. Указание на возможные сценарии профессиональной траектории граждан по итогам освоения образовательной программы**

Текущий статус	Цель
состоящий на учете в Центре занятости, безработный	трудоустроенный, самозанятый (фриланс), ИП/бизнесмен
освоение смежных профессиональных областей	повышение уровня дохода, расширение профессиональной деятельности
работающий по найму в организации, на предприятии	развитие профессиональных качеств
	повышение заработной платы


**VII. Дополнительная информация –** отсутствует.

**VIII. Приложенные Скан-копии** утвержденной образовательной программы.

**СОГЛАСОВАНО:**

Проректор  
по развитию образовательных программ

Директор Бизнес-школы


Т.В. Горопова

О.Н. Степаненко