

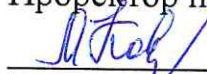

Федеральное государственное бюджетное образовательное учреждение
инклюзивного высшего образования

«МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ ГУМАНИТАРНО-
ЭКОНОМИЧЕСКИЙ УНИВЕРСИТЕТ»

Факультет прикладной математики и информатики
Кафедра Информационных технологий и прикладной математики

УТВЕРЖДАЮ

Проректор по УМР

 Ковалева М.А.
« 01 »  2020 г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

КВАНТИТАТИВНАЯ ЛИНГВИСТИКА И НОВЫЕ
ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

образовательная программа направления подготовки
45.04.02 "Лингвистика"
блок Б1.Б.04 «Дисциплины (модули)», базовая часть

Квалификация (степень) выпускника

Магистр


Форма обучения очная

Курс 1, семестр 1

Москва
2020

Рабочая программа составлена на основании Федерального государственного образовательного стандарта высшего профессионального образования направления подготовки 45.04.02 Лингвистика утвержденного приказом Министерства образования и науки Российской Федерации от 01 июля 2016 г. № 783. Зарегистрировано в Минюсте России 18 июля 2016 г. № 42896.


Составитель рабочей программы: К.т.н., доцент кафедры ИТиПМ
место работы, занимаемая должность

 Никольский А.Е. « 20 » августа 2020 г.
подпись Ф.И.О. Дата


Рецензент: к.т.н., доцент кафедры ИТиПМ
место работы, занимаемая должность

 Белоглазов А.А. « 21 » августа 2020 г.
подпись Ф.И.О. Дата


Рабочая программа утверждена на заседании кафедры Информационных технологий и прикладной математики (протокол № 1 от «24» августа 2020 г.)

Зав. кафедрой ИТиПМ  Петрунина Е.В. «24» августа 2020 г.
подпись Ф.И.О. Дата


СОГЛАСОВАНО
Начальник
Учебного отдела

« 24 » августа 2020 г.  Дмитриева И.Г.
(дата) (подпись) (Ф.И.О.)

СОГЛАСОВАНО
Декан факультета

« 24 » августа 2020 г.  Генин Э.
(дата) (подпись) (Ф.И.О.)

СОГЛАСОВАНО
Заведующий
библиотекой

« 24 » августа 2020 г.  Ахтырская В.А.
(дата) (подпись) (Ф.И.О.)

РАССМОТРЕНО
ОДОБРЕНО
УЧЕБНО-МЕТОДИЧЕСКИМ
СОВЕТОМ
ФАКУЛЬТЕТА
Пр. № 01 от 21 августа 2020

1. Цели и задачи дисциплины, ее место в учебном процессе, требования к уровню освоения содержания дисциплины.

1.1.Цели и задачи изучения дисциплины.

Цель дисциплины – сформировать у студентов знания и навыки использования количественных методов и информационных технологий для осуществления грамотных лингвистических исследований.

Данная цель ставит следующие задачи:

1. сформировать у студентов знания о границах и содержании таких направлений современной лингвистики, как количественная, компьютерная и корпусная лингвистика;
2. выработать у студентов умения и навыки использования терминологического аппарата и методов данных направлений.

1.2. Требования к результатам освоения дисциплины

В результате освоения дисциплины обучающийся должен:

Знать основные сведения о количественной, компьютерной и корпусной лингвистике, ее терминологический аппарат и методы исследования.

Уметь применять полученные знания в области количественной, компьютерной и корпусной лингвистики в научно-исследовательской и других видах деятельности;

Владеть основными методами и приемами исследовательской и практической работы в области количественной, компьютерной и корпусной лингвистики.

Владеть компетенциями:

Код компетенции	Наименование результата обучения
ОПК-17	владением современной информационной и библиографической культурой
ОПК-20	готовностью применять современные технологии сбора, обработки и интерпретации полученных экспериментальных данных
ПК-28	готовностью работать с основными информационно-поисковыми и экспертными системами, системами представления знаний, синтаксического и морфологического анализа, автоматического синтеза, распознавания и понимания речи, обработки лексикографической информации и автоматизированного перевода, автоматизированными системами идентификации и верификации личности
ПК-30	владением современными методиками сбора, хранения и представления баз данных и знаний в интеллектуальных системах различного назначения с учетом достижений корпусной лингвистики владением современными методиками разработки лингвистического обеспечения в автоматизированных системах различного профиля
ПК-31	владением современными методиками разработки лингвистического обеспечения в автоматизированных системах различного профиля

1.3. Место дисциплины в структуре ООП

Дисциплина «Количественная лингвистика и новые информационные технологии» относится к базовой части блока Б.1. Дисциплины (модули) федерального государственного образовательного стандарта по направлениям 45.04.02 Лингвистика, уровень магистратуры.

Учебная программа разрабатывается в тесной взаимосвязи с другими дисциплинами учебного плана специальности, что предупреждает возможное дублирование учебного материала, обеспечивает целостность изучения предметной области и формирование базового уровня знаний для последующего изучения дисциплин,

связанных с данной дисциплиной.

2. Содержание дисциплины

2.1. Объем дисциплины и виды учебной работы

Семестр -1, вид отчетности – зачет

№ п/п	Наименование раздела дисциплины Содержание раздела	Форма текущ. контроля
1. Квантитативная лингвистика как направление общего языкознания		
1.	Прикладная лингвистика. Математический аппарат в лингвистике. Математическая (комбинаторная и квантитативная) лингвистика. Комбинаторная лингвистика: стохастические процессы в языке, метод лингвистических переменных. Лексикостатистика (глотохронология). Количественные методы исследования в лингвистике.	Опрос Тест
2. Ключевые понятия квантитативной лингвистики		
2.	Возникновение дисциплины. Предмет, объект, цели и задачи квантитативной лингвистики. Статистические модели языка. Закон Ципфа. Определение необходимости проведения лингвостатистического анализа. Параметризация языковых единиц. Выборка. Переменные. Виды переменных. Шкала переменных. Группирующие переменные. Ранжирование. Статистическая значимость. Нормальное распределение.	Тест Опрос
3. Метод статистического анализа текста		
3.	Корреляция. Коэффициент корреляций. Показатели центральной тенденции: средняя арифметическая, мода, медиана. Мера рассеяния признака. Показатели меры рассеяния признака: лимиты, вариационный размах, среднее линейное отклонение, дисперсия, среднее квадратичное отклонение, коэффициент вариации, квадратичная ошибка средней, t-критерий Стьюдента. Оценка достоверности статистических показателей в лингвистическом исследовании: корреляция и критерий знаков. Работа с программным обеспечением (Microsoft Excel, STATISTICA и др.). Интерпретация результатов, полученных методом лингвостатистического анализа.	Тест Опрос
4. Направления лингвистики, использующие статистический анализ текста		
4.	Текстология и автороведение. Атрибуция текста. Определение авторства текста. Статистические характеристики гендера, возраста, социального статуса, происхождения автора текста. Статистическая стилистика. Статистическая семантика. Корпусная лингвистика. Машинный перевод. Лингводидактика.	Тест Опрос
5. Компьютерная и корпусная лингвистика		
5.	Предмет, объект, цели и задачи корпусной лингвистики. Крупномасштабные проекты в рамках корпусной лингвистики: Национальный Корпус Русского Языка (http://**/), WordNet (http://wordnet.princeton.edu/). Работа с системами анализа корпусов. Предмет, объект, цели и задачи компьютерной лингвистики. Возникновение и развитие дисциплины: появление ЭВМ, языки программирования, программы и алгоритмы; разработки в области искусственного интеллекта в 1950-е гг. в США, тест Тьюринга, появление и развитие «Всемирной паутины», лингвистика и новые информационные технологии, Semantic Web, нейронные сети.	Тест Опрос
6. Компьютерный анализ текста. Лингвистические аспекты разработок в области искусственного интеллекта		

6.	Автоматизированные системы обработки устной и письменной речи. Парсинг. Стемминг. Поисковые системы. Автоматическое индексирование, аннотирование и реферирование текстов. Системы управления базами данных. Системы машинного перевода. Системы анализа и синтеза устной речи. Язык и интеллект. Искусственный язык versus естественный язык. Компьютерные модели языка. Компьютерное моделирование речевых актов. Когнитивная лингвистика и модели представления знаний. Базы данных. Базы знаний. Тезаурусы, онтологии. Разработка экспертных систем. Data Mining	Тест Опрос
----	--	---------------

3. Структура дисциплины

Вид работы	Трудоемкость часов	
	1 семестр	Всего
Общая трудоемкость	72	72
Аудиторная работа:	24	24
<i>Лекции (Л)</i>	8	8
<i>Практические занятия (ПЗ)</i>	16	16
<i>Лабораторные работы (ЛР)</i>		
<i>Зачет(З)</i>	+	+
Самостоятельная работа:	48	48
Курсовой проект (КП), курсовая работа (КР)	-	-
Расчетно-графическое задание (РГЗ)	-	-
Реферат (Р)	-	-
Эссе (Э)	-	-
Самостоятельное изучение разделов	20	20
Контрольная работа (К)		
Самоподготовка (проработка и повторение лекционного материала и материала учебников и учебных пособий, подготовка к лабораторным и практическим занятиям рубежному контролю и т.д.)	28	28
Вид итогового контроля (указать вид контроля)	Зачет	Зачет

4. Распределение видов учебной работы и их трудоемкости по разделам

Разделы дисциплины, изучаемые в 1 семестре

№ п/п	Наименование разделов	Количество часов				Внеауд. Работа СР
		Всего	Аудиторная работа			
			Л	ПЗ	ЛР	
1	Квантитативная лингвистика как направление общего языкознания	9	1	-		8
2	Ключевые понятия квантитативной лингвистики	11	1	2		8
3	Метод статистического анализа текста	11	1	2		8
4	Направления лингвистики, использующие статистический анализ текста	11	1	2		8
5	Компьютерная и корпусная лингвистика	14	2	4		8
6	Компьютерный анализ текста. Лингвистические аспекты разработок в области искусственного интеллекта	16	2	4		10
	Зачет			2		
Итого по дисциплине:		72	8	16		48

5. Тематический план учебной дисциплины

Наименование разделов и тем	Содержание учебного материала, лабораторные работы и практические занятия, самостоятельная работа студентов, курсовая работа (проект)	Объем часов/зачетных единиц	Образовательные технологии	Формируемые компетенции и/ уровень освоения*	Формы текущего контроля
1	2	3	4	5	6
1. Квантитативная лингвистика как направление общего языкознания	Лекции	1	Коммуникативная лекция	ПК-21/1,2 ПК-24/1,2	Тест
	Прикладная лингвистика. Математический аппарат в лингвистике.				
	Самостоятельная работа студента	8			
	Математическая (комбинаторная и квантитативная) лингвистика. Комбинаторная лингвистика: стохастические процессы в языке, метод лингвистических переменных. Лексикостатистика (глотохронология). Количественные методы исследования в лингвистике.				
2. Ключевые понятия квантитативной лингвистики	Лекции	1	Коммуникативная лекция Работа за компьютером	ОПК-20/1,2 ПК-28/1,2 ПК-30/1,2 ПК/30,31	Опрос
	Возникновение дисциплины. Предмет, объект, цели и задачи квантитативной лингвистики. Статистические модели языка.				
	Практические занятия	2			Тест
	Закон Ципфа. Определение необходимости проведения лингвостатистического анализа.				
	Самостоятельная работа студента	8			
Параметризация языковых единиц. Выборка. Переменные. Виды переменных. Шкала переменных. Группирующие переменные. Ранжирование. Статистическая значимость. Нормальное распределение.					
3. Метод статистического анализа текста	Лекции	1	Проблемная лекция, ИТ-технологии	ОПК-17/1,2 ОПК-20/1,2 ПК-28/1,2 ПК-30/1,2	Тест
	Корреляция. Коэффициент корреляций. Показатели центральной тенденции: средняя арифметическая, мода, медиана.				

	Практические занятия	2		ПК/30,31	Опрос	
	Работа с программным обеспечением (Microsoft Excel, STATISTICA и др.). Интерпретация результатов, полученных методом лингвостатистического анализа					
	Самостоятельная работа студента	8			Опрос	
	Мера рассеяния признака. Показатели меры рассеяния признака: лимиты, вариационный размах, среднее линейное отклонение, дисперсия, среднее квадратичное отклонение, коэффициент вариации, квадратичная ошибка средней, t-критерий Стьюдента. Оценка достоверности статистических показателей в лингвистическом исследовании: корреляция и критерий знаков.					
4. Направления лингвистики, использующие статистический анализ текста	Лекции	1	Коммуни- кативная лекция IT- технологии	ОПК-17/1,2 ОПК-20/1,2 ПК-28/1,2 ПК-30/1,2 ПК/30,31	Опрос	
	Текстология и автороведение. Атрибуция текста. Определение авторства текста.					
	Практические занятия	2				Тест
	Статистические характеристики гендера, возраста, социального статуса, происхождения автора текста.					
	Самостоятельная работа студента	8			Опрос	
	Статистическая стилистика. Статистическая семантика. Корпусная лингвистика. Машинный перевод. Лингводидактика.					
5. Компьютерная и корпусная лингвистика	Лекции	2	Лекция- беседа IT- технологии	ОПК-17/1,2 ОПК-20/1,2 ПК-28/1,2 ПК-30/1,2 ПК/30,31	Тест	
	Предмет, объект, цели и задачи корпусной лингвистики. Крупномасштабные проекты в рамках корпусной лингвистики: Национальный Корпус Русского Языка (http://**/), WordNet (http://wordnet.princeton.edu/).					
	Практические занятия	4				Тест
	Работа с системами анализа корпусов. Предмет, объект, цели и задачи компьютерной лингвистики.					
	Самостоятельная работа студента.	8			Опрос	

	Возникновение и развитие дисциплины: появление ЭВМ, языки программирования, программы и алгоритмы; разработки в области искусственного интеллекта в 1950-е гг. в США, тест Тьюринга, появление и развитие «Всемирной паутины», лингвистика и новые информационные технологии, Semantic Web, нейронные сети.				
6. Компьютерный анализ текста. Лингвистические аспекты разработок в области искусственного интеллекта	Лекции Автоматизированные системы обработки устной и письменной речи. Парсинг. Стемминг. Поисковые системы. Автоматическое индексирование, аннотирование и реферирование текстов.	2	Проблемные лекции, ИТ-технологии	ОПК-17/1,2 ОПК-20/1,2 ПК-28/1,2 ПК-30/1,2 ПК/30,31	Тест
	Практические занятия Системы управления базами данных. Системы машинного перевода. Системы анализа и синтеза устной речи. Язык и интеллект. Искусственный язык versus естественный язык. Компьютерные модели языка.	4			Опрос
	Самостоятельная работа студента Компьютерное моделирование речевых актов. Когнитивная лингвистика и модели представления знаний. Базы данных. Базы знаний. Тезаурусы, онтологии. Разработка экспертных систем. Data Mining	10			Опрос
	Зачет	2			
	Всего:	72/2			

* В таблице уровень усвоения учебного материала обозначен цифрами:

1. – репродуктивный (освоение знаний, выполнение деятельности по образцу, инструкции или под руководством);
2. – продуктивный (планирование и самостоятельное выполнение деятельности, решение проблемных задач; применение умений в новых условиях);
3. – творческий (самостоятельное проектирование экспериментальной деятельности; оценка и самооценка инновационной деятельности)

6. Образовательные технологии

Интерактивные образовательные технологии, используемые в аудиторных занятиях

Семестр	Вид занятия (Л, ПР, ЛР)	Используемые интерактивные образовательные технологии	Количество часов
1	Л	Проблемная лекция, лекция-беседа	4
	ПР	IT-технологии	10
Итого:			14

7. Оценочные средства для текущего контроля успеваемости и промежуточной аттестации

7.1. Организация текущего и промежуточного контроля обучения

Текущий контроль: опрос, тест.

Промежуточная аттестация: зачет

7.2. Организация контроля:

Практические работы выполняются в форме индивидуальных заданий по разделам тем, выполняемых на ЭВМ.

Тематика опросов по дисциплине:

1. Операционные системы семейства Windows
2. Загрузка Windows.
3. Выход из Windows
4. Организация работы в среде Windows.
5. Windows-окно
6. Справочная система
7. Основы работы текстового редактора MS Word-2010,2013
8. Создание нового документа
9. Создание и форматирование таблиц
10. Создание списков.
11. Организация печати документа
12. Сохранение текстового документа
13. Регуляция вида экрана
14. Экономия времени, эффективная работа
15. Авто коррекция ошибок, расшифровка сокращений и поиск в словарях
16. Современные способы организации презентаций
17. Системы оптического распознавания информации
18. СУБД MS Access-2000
19. Компьютерные сети
20. Основы информационной и компьютерной безопасности

7.3. Тематика рефератов, проектов, творческих заданий, эссе и т.п.

Не предусмотрена программой

7.4. Курсовая работа – Не предусмотрена программой

7.5. Вопросы к зачету

1. Прикладная лингвистика.
2. Математический аппарат в лингвистике.
3. Математическая (комбинаторная и квантитативная) лингвистика.
4. Комбинаторная лингвистика: стохастические процессы в языке, метод лингвистических переменных.
5. Лексикостатистика (глотохронология). Количественные методы исследования в лингвистике.
6. Возникновение дисциплины. Предмет, объект, цели и задачи квантитативной лингвистики.
7. Статистические модели языка.
8. Закон Ципфа. Определение необходимости проведения лингвостатистического анализа.
9. Параметризация языковых единиц. Выборка.
10. Переменные. Виды переменных. Шкала переменных. Группирующие переменные.
11. Ранжирование. Статистическая значимость. Нормальное распределение.
12. Корреляция. Коэффициент корреляций.
13. Показатели центральной тенденции: средняя арифметическая, мода, медиана.
14. Мера рассеяния признака. Показатели меры рассеяния признака: лимиты, вариационный размах, среднее линейное отклонение, дисперсия, среднее квадратичное отклонение, коэффициент вариации, квадратичная ошибка средней, t-критерий Стьюдента.
15. Оценка достоверности статистических показателей в лингвистическом исследовании: корреляция и критерий знаков.
16. Текстология и автороведение. Атрибуция текста. Определение авторства текста.
17. Статистические характеристики гендера, возраста, социального статуса, происхождения автора текста.
18. Статистическая стилистика. Статистическая семантика.
19. Корпусная лингвистика. Машинный перевод. Лингводидактика.
20. Предмет, объект, цели и задачи корпусной лингвистики. Крупномасштабные проекты в рамках корпусной лингвистики: Национальный Корпус Русского Языка (http://**/), WordNet (<http://wordnet.princeton.edu/>).
21. Предмет, объект, цели и задачи компьютерной лингвистики. Возникновение и развитие дисциплины.
22. Автоматизированные системы обработки устной и письменной речи.
23. Парсинг. Стемминг. Поисковые системы.
24. Автоматическое индексирование, аннотирование и реферирование текстов.
25. Системы управления базами данных.
26. Системы машинного перевода.
27. Системы анализа и синтеза устной речи.
28. Язык и интеллект. Искусственный язык versus естественный язык.
29. Компьютерные модели языка. Компьютерное моделирование речевых актов.
30. Когнитивная лингвистика и модели представления знаний.
31. Базы данных. Базы знаний.
32. Тезаурусы, онтологии. Разработка экспертных систем. Data Mining

7.6. Критерии оценки

«Зачтено» ставится в случае, когда студент демонстрирует отличное, хорошее или удовлетворительное владение теоретическим материалом, сформированность необходимых практических навыков работы с освоенным материалом; все или большинство предусмотренных программой обучения учебных заданий выполнены, предоставлены отчеты по выполнению практических работ; студент ориентируется в

пройденном материале, демонстрирует способность к аналитической деятельности и самостоятельность мышления.

«**Не зачтено**» ставится в случае, когда теоретическое содержание курса не освоено или ответ содержит существенные пробелы в знании основного содержания учебной программы дисциплины и не умеющего использовать полученные знания при решении практических задач (в том числе при выполнении самостоятельной работы); необходимые практические навыки работы не сформированы, не предоставлены или частично предоставлены отчеты по выполнению практических работ, либо отчеты содержат грубые ошибки, дополнительная самостоятельная работа над материалом курса не приведет к какому-либо значимому повышению качества выполнения учебных заданий.

8. Сведения о материально-техническом обеспечении дисциплины

№ п/п	Наименование оборудованных учебных кабинетов, лабораторий	Перечень оборудования и технических средств обучения
1	Лекционная аудитория	Мультимедийный проектор.
2	Компьютерный класс	Компьютерный класс (компьютеры МХР Pentium, мониторы LG), принтеры, мультимедиа проектор –1. Терминалы к сети Internet.

9. Учебно-методическое обеспечение дисциплины

9.1. Основная литература:

1. Станкевич, Л. А. Интеллектуальные системы и технологии : учебник и практи-кум для бакалавриата и магистратуры / Л. А. Станкевич. — Москва : Издательство Юрайт, 2019. — 397 с. — (Бакалавр и магистр. Академический курс). — ISBN 978-5-534-02126-4. — Текст : электрон-ный // ЭБС Юрайт [сайт]. — URL: <https://biblio-online.ru/bcode/433370> .

2. Замятина, О. М. Вычислительные системы, сети и телекоммуникации. Моделирование сетей : учебное пособие для магистратуры / О. М. Замятина. — Москва : Издательство Юрайт, 2019. — 159 с. — (Университеты России). — ISBN 978-5-534-00335-2. — Текст : электронный // ЭБС Юрайт [сайт]. — URL: <https://biblio-online.ru/bcode/433938>.

9.2. Дополнительная литература:

1. Базовые и прикладные информационные технологии : учебник / В.А. Гвоздева. — Москва : ИД «ФОРУМ» : ИНФРА-М, 2019. — 383 с. — (Высшее образование). - Текст : электронный. - URL: <https://new.znaniium.com/catalog/product/1019243> .

2. Информационные технологии в науке и образовании : учеб. пособие / Е.Л. Федотова, А.А. Федотов. — Москва : ИД «ФОРУМ» : ИНФРА-М, 2019. — 335 с. — (Высшее образование). - Текст : электронный. - URL: <https://new.znaniium.com/catalog/product/1018730> .

3. Интеллектуальные интерактивные системы и технологии управления удаленным доступом: методы и модели управления процессами защиты и сопровождения интеллектуальной собственности в сети Internet/Intrane: Учебное пособие / Ботуз С.П., - 3-е изд., доп - Москва :СОЛОН-Пр., 2014. - 340 с.: ISBN 978-5-91359-132-6 - Текст : электронный. - URL: <https://new.znaniium.com/catalog/product/884094> .

4. Кудрявцев, В. Б. Интеллектуальные си-стемы : учебник и практикум для вузов / В. Б. Кудрявцев, Э. Э. Гасанов, А. С. Подколзин. — 2-е изд., испр. и доп. — Москва : Издательство Юрайт, 2019. — 165 с. — (Высшее образование). — ISBN 978-5-534-07779-7. — Текст : электрон-ный // ЭБС Юрайт [сайт]. — URL: <https://biblio-online.ru/bcode/433370> .

online.ru/bcode/423761.

5. Математическая лингвистика и компьютерный перевод : раб.учеб.программа / Никольский Анатолий Евгеньевич ; Моск.гос.соц.-гум.ин-т. - М. : МГГЭИ, 2009. - 26с. - 60.00.

9.3. Интернет-ресурсы:

1. Открытый ПП SiLab.
2. Национальный открытый Университет «ИНТУИТ» www.intuit.ru
3. Энциклопедия Кругосвет. Универсальная научно-популярная онлайн-энциклопедия. www.krugosvet.ru
4. Электронная библиотека: <https://znanium.com/>
5. Электронная библиотека <https://biblio-online.ru/>

9.4. Программное обеспечение

1. Сетевой компьютерный класс, оснащенный современной техникой
2. Офисный программный пакет (например, Microsoft Office 2003 или более поздних версий).
3. Web-браузер Mozilla Firefox или Google Chrome
4. Экран для проектора

1. Цели и задачи дисциплины, ее место в учебном процессе, требования к уровню освоения содержания дисциплины.

1.1. Цели и задачи изучения дисциплины.

Цель дисциплины – сформировать у студентов знания и навыки использования количественных методов и информационных технологий для осуществления грамотных лингвистических исследований.

Данная цель ставит следующие задачи:

1. сформировать у студентов знания о границах и содержании таких направлений современной лингвистики, как количественная, компьютерная и корпусная лингвистика;
2. выработать у студентов умения и навыки использования терминологического аппарата и методов данных направлений.

1.2. Требования к результатам освоения дисциплины

В результате освоения дисциплины обучающийся должен:

Знать основные сведения о количественной, компьютерной и корпусной лингвистике, ее терминологический аппарат и методы исследования.

Уметь применять полученные знания в области количественной, компьютерной и корпусной лингвистики в научно-исследовательской и других видах деятельности;

Владеть основными методами и приемами исследовательской и практической работы в области количественной, компьютерной и корпусной лингвистики.

Владеть компетенциями:

Код компетенции	Наименование результата обучения
ОПК-17	владением современной информационной и библиографической культурой
ОПК-20	готовностью применять современные технологии сбора, обработки и интерпретации полученных экспериментальных данных
ПК-28	готовностью работать с основными информационно-поисковыми и экспертными системами, системами представления знаний, синтаксического и морфологического анализа, автоматического синтеза, распознавания и понимания речи, обработки лексикографической информации и автоматизированного перевода, автоматизированными системами идентификации и верификации личности
ПК-30	владением современными методиками сбора, хранения и представления баз данных и знаний в интеллектуальных системах различного назначения с учетом достижений корпусной лингвистики владением современными методиками разработки лингвистического обеспечения в автоматизированных системах различного профиля
ПК-31	владением современными методиками разработки лингвистического обеспечения в автоматизированных системах различного профиля

1.3. Место дисциплины в структуре ООП

Дисциплина «Количественная лингвистика и новые информационные технологии» относится к базовой части блока Б.1. Дисциплины (модули) федерального государственного образовательного стандарта по направлениям 45.04.02 Лингвистика, уровень магистратуры.

Учебная программа разрабатывается в тесной взаимосвязи с другими дисциплинами учебного плана специальности, что предупреждает возможное дублирование учебного материала, обеспечивает целостность изучения предметной области и формирование базового уровня знаний для последующего изучения дисциплин,

связанных с данной дисциплиной.

2. Содержание дисциплины

2.1. Объем дисциплины и виды учебной работы

Семестр -1, вид отчетности – зачет

№ п/п	Наименование раздела дисциплины Содержание раздела	Форма текущ. контроля
1. Квантитативная лингвистика как направление общего языкознания		
1.	Прикладная лингвистика. Математический аппарат в лингвистике. Математическая (комбинаторная и квантитативная) лингвистика. Комбинаторная лингвистика: стохастические процессы в языке, метод лингвистических переменных. Лексикостатистика (глотохронология). Количественные методы исследования в лингвистике.	Опрос Тест
2. Ключевые понятия квантитативной лингвистики		
2.	Возникновение дисциплины. Предмет, объект, цели и задачи квантитативной лингвистики. Статистические модели языка. Закон Ципфа. Определение необходимости проведения лингвостатистического анализа. Параметризация языковых единиц. Выборка. Переменные. Виды переменных. Шкала переменных. Группирующие переменные. Ранжирование. Статистическая значимость. Нормальное распределение.	Тест Опрос
3. Метод статистического анализа текста		
3.	Корреляция. Коэффициент корреляций. Показатели центральной тенденции: средняя арифметическая, мода, медиана. Мера рассеяния признака. Показатели меры рассеяния признака: лимиты, вариационный размах, среднее линейное отклонение, дисперсия, среднее квадратичное отклонение, коэффициент вариации, квадратичная ошибка средней, t-критерий Стьюдента. Оценка достоверности статистических показателей в лингвистическом исследовании: корреляция и критерий знаков. Работа с программным обеспечением (Microsoft Excel, STATISTICA и др.). Интерпретация результатов, полученных методом лингвостатистического анализа.	Тест Опрос
4. Направления лингвистики, использующие статистический анализ текста		
4.	Текстология и автороведение. Атрибуция текста. Определение авторства текста. Статистические характеристики гендера, возраста, социального статуса, происхождения автора текста. Статистическая стилистика. Статистическая семантика. Корпусная лингвистика. Машинный перевод. Лингводидактика.	Тест Опрос
5. Компьютерная и корпусная лингвистика		
5.	Предмет, объект, цели и задачи корпусной лингвистики. Крупномасштабные проекты в рамках корпусной лингвистики: Национальный Корпус Русского Языка (http://**/), WordNet (http://wordnet.princeton.edu/). Работа с системами анализа корпусов. Предмет, объект, цели и задачи компьютерной лингвистики. Возникновение и развитие дисциплины: появление ЭВМ, языки программирования, программы и алгоритмы; разработки в области искусственного интеллекта в 1950-е гг. в США, тест Тьюринга, появление и развитие «Всемирной паутины», лингвистика и новые информационные технологии, Semantic Web, нейронные сети.	Тест Опрос
6. Компьютерный анализ текста. Лингвистические аспекты разработок в области искусственного интеллекта		

6.	Автоматизированные системы обработки устной и письменной речи. Парсинг. Стемминг. Поисковые системы. Автоматическое индексирование, аннотирование и реферирование текстов. Системы управления базами данных. Системы машинного перевода. Системы анализа и синтеза устной речи. Язык и интеллект. Искусственный язык versus естественный язык. Компьютерные модели языка. Компьютерное моделирование речевых актов. Когнитивная лингвистика и модели представления знаний. Базы данных. Базы знаний. Тезаурусы, онтологии. Разработка экспертных систем. Data Mining	Тест Опрос
----	--	---------------

3. Структура дисциплины

Вид работы	Трудоемкость часов	
	1 семестр	Всего
Общая трудоемкость	72	72
Аудиторная работа:	24	24
<i>Лекции (Л)</i>	8	8
<i>Практические занятия (ПЗ)</i>	16	16
<i>Лабораторные работы (ЛР)</i>		
<i>Зачет(З)</i>	+	+
Самостоятельная работа:	48	48
Курсовой проект (КП), курсовая работа (КР)	-	-
Расчетно-графическое задание (РГЗ)	-	-
Реферат (Р)	-	-
Эссе (Э)	-	-
Самостоятельное изучение разделов	20	20
Контрольная работа (К)		
Самоподготовка (проработка и повторение лекционного материала и материала учебников и учебных пособий, подготовка к лабораторным и практическим занятиям рубежному контролю и т.д.)	28	28
Вид итогового контроля (указать вид контроля)	Зачет	Зачет

4. Распределение видов учебной работы и их трудоемкости по разделам

Разделы дисциплины, изучаемые в 1 семестре

№ п/п	Наименование разделов	Количество часов				Внеауд. Работа СР
		Всего	Аудиторная работа			
			Л	ПЗ	ЛР	
1	Квантитативная лингвистика как направление общего языкознания	9	1	-		8
2	Ключевые понятия квантитативной лингвистики	11	1	2		8
3	Метод статистического анализа текста	11	1	2		8
4	Направления лингвистики, использующие статистический анализ текста	11	1	2		8
5	Компьютерная и корпусная лингвистика	14	2	4		8
6	Компьютерный анализ текста. Лингвистические аспекты разработок в области искусственного интеллекта	16	2	4		10
	Зачет			2		
Итого по дисциплине:		72	8	16		48

5. Тематический план учебной дисциплины

Наименование разделов и тем	Содержание учебного материала, лабораторные работы и практические занятия, самостоятельная работа студентов, курсовая работа (проект)	Объем часов/зачетных единиц	Образовательные технологии	Формируемые компетенции и/ уровень освоения*	Формы текущего контроля
1	2	3	4	5	6
1. Квантитативная лингвистика как направление общего языкознания	Лекции	1	Коммуникативная лекция	ПК-21/1,2 ПК-24/1,2	Тест
	Прикладная лингвистика. Математический аппарат в лингвистике.				
	Самостоятельная работа студента	8			
	Математическая (комбинаторная и квантитативная) лингвистика. Комбинаторная лингвистика: стохастические процессы в языке, метод лингвистических переменных. Лексикостатистика (глотохронология). Количественные методы исследования в лингвистике.				
2. Ключевые понятия квантитативной лингвистики	Лекции	1	Коммуникативная лекция Работа за компьютером	ОПК-20/1,2 ПК-28/1,2 ПК-30/1,2 ПК/30,31	Опрос
	Возникновение дисциплины. Предмет, объект, цели и задачи квантитативной лингвистики. Статистические модели языка.				
	Практические занятия	2			Опрос
	Закон Ципфа. Определение необходимости проведения лингвостатистического анализа.				
	Самостоятельная работа студента	8			
	Параметризация языковых единиц. Выборка. Переменные. Виды переменных. Шкала переменных. Группирующие переменные. Ранжирование. Статистическая значимость. Нормальное распределение.				
3. Метод статистического анализа текста	Лекции	1	Проблемная лекция, ИТ-технологии	ОПК-17/1,2 ОПК-20/1,2 ПК-28/1,2 ПК-30/1,2	Тест
	Корреляция. Коэффициент корреляций. Показатели центральной тенденции: средняя арифметическая, мода, медиана.				

	Практические занятия	2		ПК/30,31	Опрос	
	Работа с программным обеспечением (Microsoft Excel, STATISTICA и др.). Интерпретация результатов, полученных методом лингвостатистического анализа					
	Самостоятельная работа студента	8			Опрос	
	Мера рассеяния признака. Показатели меры рассеяния признака: лимиты, вариационный размах, среднее линейное отклонение, дисперсия, среднее квадратичное отклонение, коэффициент вариации, квадратичная ошибка средней, t-критерий Стьюдента. Оценка достоверности статистических показателей в лингвистическом исследовании: корреляция и критерий знаков.					
4. Направления лингвистики, использующие статистический анализ текста	Лекции	1	Коммуни- кативная лекция IT- технологии	ОПК-17/1,2 ОПК-20/1,2 ПК-28/1,2 ПК-30/1,2 ПК/30,31	Опрос	
	Текстология и автороведение. Атрибуция текста. Определение авторства текста.					
	Практические занятия	2				Тест
	Статистические характеристики гендера, возраста, социального статуса, происхождения автора текста.					
	Самостоятельная работа студента	8			Опрос	
	Статистическая стилистика. Статистическая семантика. Корпусная лингвистика. Машинный перевод. Лингводидактика.					
5. Компьютерная и корпусная лингвистика	Лекции	2	Лекция- беседа IT- технологии	ОПК-17/1,2 ОПК-20/1,2 ПК-28/1,2 ПК-30/1,2 ПК/30,31	Тест	
	Предмет, объект, цели и задачи корпусной лингвистики. Крупномасштабные проекты в рамках корпусной лингвистики: Национальный Корпус Русского Языка (http://**/), WordNet (http://wordnet.princeton.edu/).					
	Практические занятия	4				Тест
	Работа с системами анализа корпусов. Предмет, объект, цели и задачи компьютерной лингвистики.					
	Самостоятельная работа студента.	8			Опрос	

	Возникновение и развитие дисциплины: появление ЭВМ, языки программирования, программы и алгоритмы; разработки в области искусственного интеллекта в 1950-е гг. в США, тест Тьюринга, появление и развитие «Всемирной паутины», лингвистика и новые информационные технологии, Semantic Web, нейронные сети.				
6. Компьютерный анализ текста. Лингвистические аспекты разработок в области искусственного интеллекта	Лекции	2	Проблемные лекции, ИТ-технологии	ОПК-17/1,2 ОПК-20/1,2 ПК-28/1,2 ПК-30/1,2 ПК/30,31	Тест
	Автоматизированные системы обработки устной и письменной речи. Парсинг. Стемминг. Поисковые системы. Автоматическое индексирование, аннотирование и реферирование текстов.				
	Практические занятия	4			Опрос
	Системы управления базами данных. Системы машинного перевода. Системы анализа и синтеза устной речи. Язык и интеллект. Искусственный язык versus естественный язык. Компьютерные модели языка.				
	Самостоятельная работа студента	10			Опрос
	Компьютерное моделирование речевых актов. Когнитивная лингвистика и модели представления знаний. Базы данных. Базы знаний. Тезаурусы, онтологии. Разработка экспертных систем. Data Mining				
	Зачет	2			
	Всего:	72/2			

* В таблице уровень усвоения учебного материала обозначен цифрами:

1. – репродуктивный (освоение знаний, выполнение деятельности по образцу, инструкции или под руководством);
2. – продуктивный (планирование и самостоятельное выполнение деятельности, решение проблемных задач; применение умений в новых условиях);
3. – творческий (самостоятельное проектирование экспериментальной деятельности; оценка и самооценка инновационной деятельности)

6. Образовательные технологии

Интерактивные образовательные технологии, используемые в аудиторных занятиях

Семестр	Вид занятия (Л, ПР, ЛР)	Используемые интерактивные образовательные технологии	Количество часов
1	Л	Проблемная лекция, лекция-беседа	4
	ПР	IT-технологии	10
Итого:			14

7. Оценочные средства для текущего контроля успеваемости и промежуточной аттестации

7.1. Организация текущего и промежуточного контроля обучения

Текущий контроль: опрос, тест.

Промежуточная аттестация: зачет

7.2. Организация контроля:

Практические работы выполняются в форме индивидуальных заданий по разделам тем, выполняемых на ЭВМ.

Тематика опросов по дисциплине:

1. Операционные системы семейства Windows
2. Загрузка Windows.
3. Выход из Windows
4. Организация работы в среде Windows.
5. Windows-окно
6. Справочная система
7. Основы работы текстового редактора MS Word-2010,2013
8. Создание нового документа
9. Создание и форматирование таблиц
10. Создание списков.
11. Организация печати документа
12. Сохранение текстового документа
13. Регуляция вида экрана
14. Экономия времени, эффективная работа
15. Авто коррекция ошибок, расшифровка сокращений и поиск в словарях
16. Современные способы организации презентаций
17. Системы оптического распознавания информации
18. СУБД MS Access-2000
19. Компьютерные сети
20. Основы информационной и компьютерной безопасности

7.3. Тематика рефератов, проектов, творческих заданий, эссе и т.п.

Не предусмотрена программой

7.4. Курсовая работа – Не предусмотрена программой

7.5. Вопросы к зачету

1. Прикладная лингвистика.
2. Математический аппарат в лингвистике.
3. Математическая (комбинаторная и квантитативная) лингвистика.
4. Комбинаторная лингвистика: стохастические процессы в языке, метод лингвистических переменных.
5. Лексикостатистика (глотохронология). Количественные методы исследования в лингвистике.
6. Возникновение дисциплины. Предмет, объект, цели и задачи квантитативной лингвистики.
7. Статистические модели языка.
8. Закон Ципфа. Определение необходимости проведения лингвостатистического анализа.
9. Параметризация языковых единиц. Выборка.
10. Переменные. Виды переменных. Шкала переменных. Группирующие переменные.
11. Ранжирование. Статистическая значимость. Нормальное распределение.
12. Корреляция. Коэффициент корреляций.
13. Показатели центральной тенденции: средняя арифметическая, мода, медиана.
14. Мера рассеяния признака. Показатели меры рассеяния признака: лимиты, вариационный размах, среднее линейное отклонение, дисперсия, среднее квадратичное отклонение, коэффициент вариации, квадратичная ошибка средней, t-критерий Стьюдента.
15. Оценка достоверности статистических показателей в лингвистическом исследовании: корреляция и критерий знаков.
16. Текстология и автороведение. Атрибуция текста. Определение авторства текста.
17. Статистические характеристики гендера, возраста, социального статуса, происхождения автора текста.
18. Статистическая стилистика. Статистическая семантика.
19. Корпусная лингвистика. Машинный перевод. Лингводидактика.
20. Предмет, объект, цели и задачи корпусной лингвистики. Крупномасштабные проекты в рамках корпусной лингвистики: Национальный Корпус Русского Языка (http://**/), WordNet (<http://wordnet.princeton.edu/>).
21. Предмет, объект, цели и задачи компьютерной лингвистики. Возникновение и развитие дисциплины.
22. Автоматизированные системы обработки устной и письменной речи.
23. Парсинг. Стемминг. Поисковые системы.
24. Автоматическое индексирование, аннотирование и реферирование текстов.
25. Системы управления базами данных.
26. Системы машинного перевода.
27. Системы анализа и синтеза устной речи.
28. Язык и интеллект. Искусственный язык versus естественный язык.
29. Компьютерные модели языка. Компьютерное моделирование речевых актов.
30. Когнитивная лингвистика и модели представления знаний.
31. Базы данных. Базы знаний.
32. Тезаурусы, онтологии. Разработка экспертных систем. Data Mining

7.6. Критерии оценки

«Зачтено» ставится в случае, когда студент демонстрирует отличное, хорошее или удовлетворительное владение теоретическим материалом, сформированность необходимых практических навыков работы с освоенным материалом; все или большинство предусмотренных программой обучения учебных заданий выполнены, предоставлены отчеты по выполнению практических работ; студент ориентируется в

пройденном материале, демонстрирует способность к аналитической деятельности и самостоятельность мышления.

«**Не зачтено**» ставится в случае, когда теоретическое содержание курса не освоено или ответ содержит существенные пробелы в знании основного содержания учебной программы дисциплины и не умеющего использовать полученные знания при решении практических задач (в том числе при выполнении самостоятельной работы); необходимые практические навыки работы не сформированы, не предоставлены или частично предоставлены отчеты по выполнению практических работ, либо отчеты содержат грубые ошибки, дополнительная самостоятельная работа над материалом курса не приведет к какому-либо значимому повышению качества выполнения учебных заданий.

8. Сведения о материально-техническом обеспечении дисциплины

№ п/п	Наименование оборудованных учебных кабинетов, лабораторий	Перечень оборудования и технических средств обучения
1	Лекционная аудитория	Мультимедийный проектор.
2	Компьютерный класс	Компьютерный класс (компьютеры МХР Pentium, мониторы LG), принтеры, мультимедиа проектор –1. Терминалы к сети Internet.

9. Учебно-методическое обеспечение дисциплины

9.1. Основная литература:

1. Станкевич, Л. А. Интеллектуальные системы и технологии : учебник и практи-кум для бакалавриата и магистратуры / Л. А. Станкевич. — Москва : Издательство Юрайт, 2019. — 397 с. — (Бакалавр и ма-гистр. Академический курс). — ISBN 978-5-534-02126-4. — Текст : электрон-ный // ЭБС Юрайт [сайт]. — URL: <https://biblio-online.ru/bcode/433370> .

2. Замятина, О. М. Вычислительные системы, сети и телекоммуникации. Моделирование сетей : учебное пособие для магистратуры / О. М. Замятина. — Москва : Издательство Юрайт, 2019. — 159 с. — (Университеты России). — ISBN 978-5-534-00335-2. — Текст : электронный // ЭБС Юрайт [сайт]. — URL: <https://biblio-online.ru/bcode/433938>.

9.2. Дополнительная литература:

1. Базовые и прикладные информационные технологии : учебник / В.А. Гвоздева. — Москва : ИД «ФОРУМ» : ИНФРА-М, 2019. — 383 с. — (Высшее образование). - Текст : электронный. - URL: <https://new.znanium.com/catalog/product/1019243> .

2. Информационные технологии в науке и образовании : учеб. пособие / Е.Л. Федотова, А.А. Федотов. — Москва : ИД «ФОРУМ» : ИНФРА-М, 2019. — 335 с. — (Высшее образование). - Текст : электронный. - URL: <https://new.znanium.com/catalog/product/1018730> .

3. Интеллектуальные интерактивные системы и технологии управления удаленным доступом: методы и модели управления процессами защиты и сопровождения интеллектуальной собственности в сети Internet/Intrane: Учебное пособие / Ботуз С.П., - 3-е изд., доп - Москва :СОЛОН-Пр., 2014. - 340 с.: ISBN 978-5-91359-132-6 - Текст : электронный. - URL: <https://new.znanium.com/catalog/product/884094> .

4. Кудрявцев, В. Б. Интеллектуальные си-стемы : учебник и практикум для вузов / В. Б. Кудрявцев, Э. Э. Гасанов, А. С. Подколзин. — 2-е изд., испр. и доп. — Москва : Издательство Юрайт, 2019. — 165 с. — (Высшее образование). — ISBN 978-5-534-07779-7. — Текст : электрон-ный // ЭБС Юрайт [сайт]. — URL: <https://biblio-online.ru/bcode/433370> .

online.ru/bcode/423761.

5. Математическая лингвистика и компьютерный перевод : раб.учеб.программа / Никольский Анатолий Евгеньевич ; Моск.гос.соц.-гум.ин-т. - М. : МГГЭИ, 2009. - 26с. - 60.00.

9.3. Интернет-ресурсы:

1. Открытый ПП SiLab.
2. Национальный открытый Университет «ИНТУИТ» www.intuit.ru
3. Энциклопедия Кругосвет. Универсальная научно-популярная онлайн-энциклопедия. www.krugosvet.ru
4. Электронная библиотека: <https://znanium.com/>
5. Электронная библиотека <https://biblio-online.ru/>

9.4. Программное обеспечение

1. Сетевой компьютерный класс, оснащенный современной техникой
2. Офисный программный пакет (например, Microsoft Office 2003 или более поздних версий).
3. Web-браузер Mozilla Firefox или Google Chrome
4. Экран для проектора

