Федеральное государственное бюджетное
образовательное учреждение высшего образования
«Государственный институт русского языка им. А.С. Пушкина»

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)

Наименование дисциплины (модуля):

Инструменты искусственного интеллекта для анализа языка и текста

код и наименование дисциплины (модуля)

Уровень высшего образования:

магистратура

указывается: бакалавриат, магистратура

Направление подготовки:

45.04.01 "Филология"

(код и название направления подготовки)

Направленность (профиль) образования:

Лингвистическая аналитика и экспертиза текстов

Форма обучения:

очная

очная, очно-заочная, заочная

Рабочая программа дисциплины (модуля) разработана в соответствии с федеральным государственным образовательным стандартом высшего образования по направлению подготовки 45.04.01 Филология, утвержденным приказом Минобрнауки России № 980 от 12.08.2020

В соответствии с учебным планом направления магистратуры «Лингвистическая аналитика и экспертиза текста» дисциплина «Инструменты искусственного интеллекта для анализа языка и текста» входит в состав дисциплин (модулей), формируемых участниками образовательных отношений (Б1.В.02.01).

1. Цель освоения дисциплины:

Систематизация и расширение знаний в области новых информационных технологий, методов статистической и автоматической обработки лингвистических явлений для решения прикладных задач в науке и осуществления грамотных лингвистических исследований.

2. Задачи освоения дисциплины:

Знать: современные технологии сбора, обработки и интерпретации данных.

Уметь: применять современные технологии сбора, обработки и интерпретации данных в целых повышения качества исследования.

Владеть: современными технологиями сбора, обработки и интерпретации данных в целых повышения качества исследования.

3. Планируемые результаты обучения по дисциплине (модулю), соотнесенные с требуемыми компетенциями выпускников

Компетенции выпускников (коды)	**Индикаторы достижения компетенций**	**Планируемые результаты обучения по дисциплине (модулю), соотнесенные с индикаторами достижения компетенций**
УК-3: Способен организовывать и руководить работой команды, вырабатывая командную стратегию для достижения поставленной цели	УК-3.1 Вырабатывает стратегию сотрудничества и на ее основе организует работу команды для достижения поставленной цели	Знать: стратегии сотрудничества и принципы организации команды. Уметь: организовывать работу команды для достижения поставленной цели. Владеть: стратегиями сотрудничества и принципами организации работы команды.
	УК-3.2 Учитывает в своей социальной и профессиональной деятельности интересы, особенности поведения и мнения (включая критические) людей, с которыми	Знать: особенности поведения и мнения (включая критические) людей разных культур, социальных, этнических и религиозных групп.

	работает/взаимодействует, в том числе посредством корректировки своих действий	Уметь: учитывать в своей социальной и профессиональной деятельности интересы, особенности поведения и мнения (включая критические) людей, с которыми предстоит работать/взаимодействовать. Владеть: принципами учета интересов людей, относящихся к разным этническим, социальным и религиозным группам.
	УК-3.3 Разрешает конфликты и противоречия при деловом общении на основе учета интересов всех сторон. Создает рабочую атмосферу, позитивный эмоциональный климат в команде	Знать: пути возникновения конфликтов и противоречий при деловом общении на основе учета интересов всех сторон. Уметь: создавать рабочую атмосферу, позитивный эмоциональный климат в команде. Владеть: методами разрешения конфликтов и принципами создания позитивного эмоционального климата в команде.
	УК-3.4 Планирует командную работу, распределяет поручения и делегирует полномочия членам команды, организует обсуждение разных идей и мнений, дает обратную связь по результатам	Знать: принципы планирования командной работы, распределения поручений и организации обсуждения разных идей и мнений. Уметь: планировать командную работу, распределять поручения и делегировать полномочия членам команды, организовывать обсуждение разных идей и мнений. Владеть: принципами планирования командной работы, организации обсуждения разных идей и мнений.

ПК-2	ПК-2.1 Знает современные технологии сбора, обработки и интерпретации данных	Знать: современные технологии сбора, обработки и интерпретации данных. Уметь: применять современные технологии сбора, обработки и интерпретации данных. Владеть: современными технологиями сбора, обработки и интерпретации данных
	ПК-2.2 Применяет современные технологии сбора, обработки и интерпретации данных в целых повышения качества исследования	Знать: современные технологии сбора, обработки и интерпретации данных. Уметь: применять современные технологии сбора, обработки и интерпретации данных в целых повышения качества исследования. Владеть: современными технологиями сбора, обработки и интерпретации данных в целых повышения качества исследования

4. Объем дисциплины (модуля) 3 з.е., в том числе 32 академических часов на контактную работу обучающихся с преподавателем; 72 академических часов на самостоятельную работу обучающихся.

5. Содержание дисциплины (модуля), структурированное по темам (разделам) с указанием отведенного на них количества академических часов

Наименование и краткое содержание разделов и тем дисциплины (модуля). **Форма промежуточной аттестации по дисциплине (модулю)**	**Всего (часы)**	В том числе	
		Контактная работа (работа во взаимодействии с преподавателем) ***Виды контактной работы, часы***	**Самостоятельная работа обучающегося** ***Виды***

							самостоятельной работы, часы		
		Занятия лекционного типа	Занятия семинарского типа (практические/	Групповые консультации	Индивидуальная работа с обучающимися	**Всего**	**Самостоятельная работа**	**Проекты**	**Всего**
1. Критерий хи-квадрат и расчет стандартизированных остатков	8	2	2			4	4		4
2. Список Сводеша как объект лингвостатистического анализа	10	2	2			4		6	4
3. Закон Ципфа и его практическое применение при определении естественности текста	8	2	2			4	4		4
4. Применение корпусных данных в диахронных исследованиях языковых явлений	12	2	2			4		8	4
5. Анализ текстов с использованием программы AntConc	8	2	2			4	4		4

6. Чат-боты и нейросети: история создания, методы тестирования и моделирования речемыслительных способностей искусственного интеллекта	8	2	2			4		4	4
7. Особенности речи, сгенерированной искусственным интеллектом (синтаксический, морфологический, лексический, фонетический и статистический аспекты);	8	2	2			4	4		4
8. Типология галлюцинаций сгенерированных текстов: реминисцентные, свободно-ассоциативные, типовые.	10	2	2			4		4	4
Промежуточная аттестация *зачет*	*Самостоятельные работы, проекты* *2*								
Итого	72								

6. Фонд оценочных материалов по дисциплине (модулю)

6.1. Типовые задания и иные материалы, необходимые для оценки результатов обучения.

Вопросы для самостоятельной работы:

1. Методы лингвостатистического исследования: критерий хи-квадрат, расчет среднеквадратичного отклонения, расчет стандартизированных остатков.

2. Список Сводеша как объект статистического исследования.

3. Закон Ципфа: история открытия, вариации, формулировка, применение.

4. Частотные словари. Понятие об IPM. Понятие о «коэффициенте Жуайена». «Метафора хоббита».

5. Корпусная лингвистика: история и основные понятия (конкорданс, разметка, n-граммы, лемы и т.д.)

6. История сознания чат-ботов и нейросетей.

7. Особенности речи, сгенерированной искусственным интеллектом (синтаксический, морфологический, лексический, фонетический и статистический аспекты).

8. Тест Тьюринга. Методы тестирования и моделирования речемыслительных способностей искусственного интеллекта (метафора «португальского стадиона», метафора «китайской комнаты»).

9. Типология галлюцинаций сгенерированных текстов: реминисцентные, свободно-ассоциативные, типовые.

6.2. Тематика проектов:

1. Проведите статистический анализ тем какого-либо периодического издания с применением «критерия хи-квадрат», с использованием метода стандартизированнных остатков и с использованием метода расчета среднеквадратичного отклонения.

2. Опираясь на закон Ципфа проанализируйте лексический состав какого-либо цельного текста, выявите его статистические характеристики. О чем они свидетельствуют?

3.Пользуясь частотным словарем, проанализируйте динамику в изменениях нескольких любых слов, относящихся к разным пластам лексики (абстрактные, конкретные, историзмы, архаизмы и т.д.).

4. Пользуясь Национальным корпусом русского языка проанализируйте трансформацию в значении нескольких слов.

5. Составьте с помощью нейросети текст, содержащий большой объем фактической информации, сгруппируйте галлюцинации, представленные в этом тексте.

Шкала и критерии оценивания

Критерии и шкала оценки проекта в баллах:

Критерии	Показатели
Выполнены все требования к написанию проекта: тема раскрыта полностью,	Отлично (зачтено)

сформулированы выводы, выдержан объем, соблюдены требования к внешнему оформлению, даны правильные ответы на дополнительные вопросы	
Основные требования к проекту выполнены, но при этом имеются недочеты. В частности, имеются неточности в изложении материала, отсутствует логическая последовательность в суждениях, не выдержан объем доклада, имеются упущения в оформлении, на дополнительные вопросы даны неполные ответы	Хорошо (зачтено)
Имеются существенные отступления от требований к оформлению проекта: тема освещена лишь частично, допущены фактические ошибки в содержании или при ответе на дополнительные вопросы (или не смог ответить на дополнительные вопросы), отсутствует вывод	Удовлетворительно (зачтено)
Обучающийся не смог раскрыть тему проекта, обнаруживается существенное непонимание проблемы, не ответил на дополнительные вопросы либо проект не представлен	Неудовлетворительно (не зачтено)

7. Методические материалы

7.1. Перечень основной и дополнительной учебной литературы и иные материалы

Основная литература:

1. Воронов, М. В. Системы искусственного интеллекта: Учебник и практикум для вузов / М. В. Воронов, В. И. Пименов, И. А. Небаев. — 2-е изд., перераб. и доп. — Москва: Издательство Юрайт, 2024. — 268 с. — (Высшее образование). — ISBN 978-5-534-17032-0. — Текст: электронный // Образовательная платформа Юрайт [сайт]. — URL: https://urait.ru/bcode/544161 (дата обращения: 06.12.2024).

Дополнительная литература:

1. Айдагулова А.Р. Особенности текстов, сгенерированных искусственным интеллектом // Вестник Башкирского государственного педагогического университета им. М. Акмуллы. 2023. № 4 (72). С. 154–156.

2. Безуглый Т.А., Ершова М.Е. Использование текстовых нейросетей и искусственного интеллекта в учебных работах студентов // Проблемы современного образования. 2023. № 5. С. 206–216.

3. Искусственный интеллект: теоретические аспекты, практическое применение: материалы Донец. междунар. науч. круглого стола. Донецк: ИПИИ, 2020. 252 с.

4. Киселева И.А., Борисова Ю.В., Маевская А.Ю. Анализ возможности использования нейросети для генерирования заданий по английскому языку // Перспективы науки и образования. 2025. № 1. С. 319–335. DOI: 10.32744/pse.2025.1.21. (На англ. яз.).

5. Оломская Н.Н., Зиньковская А.В. Механизмы генерирования фейковой информации искусственным интеллектом в современном медиадискурсе // Российский социально-гуманитарный журнал. 2024. № 2. С. 207–223. DOI: 10.18384/2224-0209-2024-2-1456.

6. Соколов А.В., Габайдулин Р.В., Сухонина Н.А. Ответы нейросетей на запросы пользователей: интепретации и особенности восприятия // Вестник Пермского университета. Политология. 2025. Т. 19. № 1. С. 100–111. DOI:10.17072/2218-1067-2025-1-100-111.

7. Станкевич, Л. А. Интеллектуальные системы и технологии: Учебник и практикум для вузов / Л. А. Станкевич. — 2-е изд., перераб. и доп. — Москва: Издательство Юрайт, 2024. — 478 с. — (Высшее образование). — ISBN 978-5-534-20363-9. — Текст: электронный // Образовательная платформа Юрайт [сайт]. — URL: https://urait.ru/bcode/558009 (дата обращения: 06.12.2024).

8. Тельпов Р.Е. Особенности представления фактической информации в сгененрированных текстах / Коммуникативные исследования. 2025. Т. 12. № 4. С. 805-818. Айдагулова А.Р. Особенности текстов, сгенерированных искусственным интеллектом // Вестник Башкирского государственного педагогического университета им. М. Акмуллы. 2023. № 4 (72). С. 154–156.

9. Тельпов Р.Е. Об одушевленности слова чат-бот / Магия ИННО: перспективы развития лингвистики и лингводидактики в современных условиях. Электронный сборник научных трудов. Москва, 2024. С. 106-113.

10. Тельпов Р.Е. Космос, эфир и Вселенная – названия безграничных пространств в диахроническом аспекте / Корпус Современных Моделей и Кодов русского языка. сборник научных статей. Москва, 2024. С. 61-68.

11. Тельпов Р.Е., Ларцина С.В. Типовые различия естественных и сгенерированных сетью текстов в квантитативном аспекте // Научный диалог. 2023. Т. 12. № 7. С. 47–65. DOI: 10.24224/2227-1295-2023-12-7-47-65.

12. Тернопол Т.В., Мухина А.М. Использование нейросетей в дополнительном образовании детей (проектная деятельность на занятиях по английскому языку) // Актуальные проблемы филологии и методики преподавания иностранных языков. 2025. Т. 19. № 1. С. 76–82.

13. Туркулец И.А. Композиционные особенности текстов, сгенерированных ChatGPT, как маркер несамостоятельности выполнения работ студентами // Правовая реальность в условиях цифровизации общества: материалы Всерос. науч.- практ. конф. (Хабаровск, 9–11 нояб. 2023 г.). Хабаровск: Дальневост. гос. ун-т путей сообщения, 2023. С. 59–68.

14. Черкасова М.Н., Тактарова А.В. Признаки сгенерированного текста в академическом дискурсе: проблема идентификации // Филологические науки. Вопросы теории и практики. 2024. Т. 17. № 7. С. 2226–2232. DOI: 10.30853/phil20240307.

15. Angerschmid A., Theuermann K., Holzinger A., Chen F., Zhou J. Effects of Fairness and Explanation on Trust in Ethical AI // Machine Learning and Knowledge Extraction: 6th IFIP TC 5, TC 12, WG 8.4, WG 8.9, WG 12.9 International CrossDomain Conference (CD-MAKE 2022), Vienna, Austria, August 23–26, 2022: Proceedings / eds. A. Holzinger, P. Kieseberg, A.M. Tjoa, E. Weippl. Cham: Springer, 2022. P. 51–67. (Lecture Notes in Computer Science; vol. 13480).

DOI: 10.1007/978-3-031-14463-9_4.

16. Ferrara E. Fairness and Bias in Artificial Intelligence: A Brief Survey of Sources, Impacts, and Mitigation Strategies // Sci. 2024. Vol. 6. Iss. 1. Art. 3. DOI: 10.3390/sci6010003.

17. Mylrea M., Robinson N. Artificial Intelligence (AI) Trust Framework and Maturity Model: Applying an Entropy Lens to Improve Security, Privacy, and Ethical AI // Entropy. 2023. Vol. 25. Iss. 10. Art. 1429. DOI: 10.3390/e25101429.

6.2. Словари и справочники:

1. Большая Российская энциклопедия. - https://old.bigenc.ru/ (Электронная версия издания с открытым доступом в Интернет)

7.2. Перечень лицензионного программного обеспечения

Институт обеспечен следующим необходимым комплектом лицензионного и свободно распространяемого программного обеспечения, в том числе отечественного производства:

1. Microsoft Windows 10,
2. Microsoft Office 2016 стандартная,
3. Big Blue Button.

8. Ресурсное обеспечение

8.1. Перечень профессиональных баз данных и информационных справочных систем:

Наименование	Ссылка на ресурс	Доступность (свободный доступ / ограниченный доступ)
Лингвистические корпуса и электронные собрания текстов		
Национальный корпус русского языка	www.ruscorpora.ru	Свободный
Электронная библиотека ИРЛИ РАН	http://lib.pushkinskijdom.ru/?ysclid=lyctg57lca218433412	Свободный
The Corpus of Contemporary American English (COCA)	www.english-corpora.org	Свободный
Электронные словари		

О.Н. Ляшевская, С.А. Шаров Новый частотный словарь русской лексики	http://dict.ruslang.ru/freq.php	Свободный
Словари русского языка – Грамота.ру	https://gramota.ru/biblioteka/slovari	Свободный
Электронно-библиотечные системы		
ЭБС Юрайт	https://urait.ru	Требуется регистрация
Ресурсы для анализа и обработки текстов		
Ваал-мини	http://vaal.ru/prog/free.php	Свободный
Лингвоанализатор	https://www.rusf.ru	Свободный
Онлайн-калькулятор	https://math.semestr.ru/index.php	Свободный
Текстометр	https://textometr.ru	Свободный
AntConc	https://laurenceanthony.net/software/antconc	Свободный
SEO-анализ текста	https://pr-cy.ru/tools/	Свободный
WebScript.Ru	https://webscript.ru	Свободный

8.2. Перечень ресурсов информационно-телекоммуникационной сети «Интернет»:

Наименование	Ссылка на ресурс	Доступность (свободный доступ / ограниченный доступ)
Лингвистические корпуса и электронные собрания текстов		
WordProject	https://www.wordproject.org/	Свободный
Электронные словари		
GIGA CHAT	https://developers.sber.ru	Свободный

9. Язык преподавания: русский

10. Разработчики программы: кандидат филологических наук, доцент Тельпов Роман Евгеньевич

Приложение к рабочей
программе дисциплины
(модуля)

Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное
образовательное учреждение высшего образования
«Государственный институт русского языка им. А.С. Пушкина»

ФОНД ОЦЕНОЧНЫХ МАТЕРИАЛОВ ПО ДИСЦИПЛИНЕ (МОДУЛЮ)

Инструменты искусственного интеллекта для анализа языка и текста (название дисциплины (модуля))

Код и наименование направления подготовки: 45.04.01 "Филология"

Уровень высшего образования: магистратура

Направленность (профиль) образования: Лингвистическая аналитика и экспертиза текстов
Форма (формы) обучения: очная

Квалификация: магистр

г. Москва

Об утверждении Положения о порядке разработки и утверждения образовательных программы высшего образования (бакалавриат, магистратура)

1. ОЦЕНИВАНИЕ РЕЗУЛЬТАТОВ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Цель фонда оценочных материалов – контроль и оценка образовательных достижений обучающихся в ходе освоения рабочей программы дисциплины «Информационные технологии в профессиональной сфере».

Цель текущей аттестации – обеспечение своевременной обратной связи, для коррекции обучения, активизации самостоятельной работы обучающихся.

Цель промежуточной аттестации – оценивание промежуточных и окончательных результатов обучения по дисциплине, обеспечивающих достижение планируемых результатов освоения ОПОП магистратуры «Международные отношения».

2. ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ДЛЯ ПРОВЕРКИ СФОРМИРОВАННОСТИ КОМПЕТЕНЦИЙ

№ п/п	Содержание задания	Шифр индикатора достижения компетенции (ИДК)					
		УК-3.1	УК-3.2	УК-3.3	УК-3.4	ПК-2.1	ПК-2.2
Раздел/ Тема Методы статистической обработки данных в лингвистических исследованиях							
1.	Что изучает квантитативная лингвистика? - взаимодействие звуков в речевом потоке - статистические закономерности в языке - взаимодействие высказывания и смысла **Ключ:** статистические закономерности в языке	+	+	+	+	+	+
2.	Зачем применяется критерий хи-квадрат в лингвистике? - для правильного подсчета слов - для выведения процентов из языковых фактов - для выявления статистической значимости в наблюдаемых языковых фактах **Ключ:** для выявления статистической значимости в наблюдаемых языковых фактах	+	+	+	+	+	+
4.	Что является фактором в таблице сопряженности? - содержание колонок	+	+	+	+	+	+

	- содержание столбцов - числовые значения в ячейках **Ключ:** содержание колонок						
5.	Что является уровнем фактора в таблице сопряженности? - содержание колонок - числовые значения в ячейках - содержание столбцов **Ключ:** содержание столбцов	+	+	+	+	+	+
6.	Отклик в статистических исследованиях? - содержание колонок - числовые значения в ячейках - содержание столбцов **Ключ:** числовые значения в ячейках	+	+	+	+	+	+
7.	Поясните термины «фактор», «уровень фактора» и «отклик» на примере конкретного исследования, составив таблицу сопряженности.	+	+	+	+	+	+
8.	Выполните собственный расчет среднеквадратичного отклонения в статистическом исследовании.	+	+	+	+	+	+
9.	Как расчет среднеквадратичного отклонения соотносится с критерием хи-квадрат? - зона значимости по критерию хи-квадрат соотносится с высокими значениями абсолютной и относительной ошибки - зона значимости по критерию хи-квадрат соотносится с низкими значениями абсолютной и относительной ошибки - зона значимостьи по критерию хи-квадрат соотносится со средними значениями абсолютной и относительной ошибки **Ключ:** зона значимости по критерию хи-квадрат соотносится с высокими значениями абсолютной и относительной ошибки	+	+	+	+	+	+

10.	Что такое генеральная лингвистическая совокупность? - совокупность всех объектов (единиц), относительно которых предполагается делать выводы при решении конкретной задачи. - совокупность элементов, которая охватывается экспериментом (наблюдением, опросом). **Ключ:** совокупность всех объектов (единиц), относительно которых предполагается делать выводы при решении конкретной задачи	+	+	+	+	+	+
11.	Что такое выборочная лингвистическая совокупность? - совокупность всех объектов (единиц), относительно которых предполагается делать выводы при решении конкретной задачи. - совокупность элементов, которая охватывается экспериментом (наблюдением, опросом). **Ключ:** совокупность элементов, которая охватывается экспериментом (наблюдением, опросом)	+	+	+	+	+	+
12.	Что дает для понимания статистических закономерностей метод расчета стандартизированных остатков? - выявить статистическую значимость данных в таблице сопряженности - выявить склонность / отсутствие склонности отдельных элементов к проявлению признака **Ключ:** выявить склонность / отсутствие склонности отдельных элементов к проявлению признака	+	+	+	+	+	+
13.	Закон Ципфа. Почему именно Джорджа Ципфа называют отцом квантитативной лингвистики? - потому что он изобрел критерий хи-квадрат - потому что он изобрел среднеквадратичное отклонение - потому что он доказал возможность исследования лингвистических данных с использованием статистических методов **Ключ:** потому что он доказал возможность исследования	+	+	+	+	+	+

	лингвистических данных с использованием статистических методов						
14.	Атрибуция текста. В какой сфере она применяется?	+	+	+	+	+	+
15.	Какие виды атрибуции текста вы знаете?	+	+	+	+	+	+
16.	Чем приемы атрибуции древних рукописных памятников отличаются от атрибуции современных текстов?	+	+	+	+	+	+
17.	Какие вы знаете современные электронные ресурсы, помогающие установить авторство текста.	+	+	+	+	+	+
18.	Лингвистический корпус: его предназначение, виды.	+	+	+	+	+	+
19.	Что такое лингвистический корпус? - коллекция текстов, хранящихся на компьютере - коллекция слов, хранящихся на компьютере - база данных по литературе, языкознанию и истории, хранящаяся на компьютере **Ключ:** коллекция текстов, хранящихся на компьютере	+	+	+	+	+	+
20.	Зачем нужны лингвистические корпуса? - потому что данные словарей не являются достоверными - потому что обращение к корпусу помогает узнать информацию об употреблениях слов в их динамике - потому что с помощью корпуса можно ответить на все вопросы о языке **Ключ:** потому что обращение к корпусу помогает узнать информацию об употреблениях слов в их динамике	+	+	+	+	+	+
21.	Что такое лемма? - аналог термина *фонема*, используемый в компьютерной лингвистике - аналог термина *лексема*, используемый в компьютерной лингвистике	+	+	+	+	+	+

	- аналог термина *морфема*, используемый в компьютерной лингвистике **Ключ:** аналог термина *лексема*, используемый в компьютерной лингвистике						
22.	Что такое токен в компьютерной лингвистике? - совокупность символов от пробела до пробела - совокупность морфем от пробела до пробела - совокупность предложений от точки до точки **Ключ:** совокупность символов от пробела до пробела	+	+	+	+	+	+
23.	Что НЕ входит в понятие клауза? - простое предложение - словосочетание, состоящее из прилагательного и существительного - причастный оборот - деепричастный оборот - вводное слово **Ключ:** словосочетание, состоящее из прилагательного и существительного, вводное слово	+	+	+	+	+	+
24.	Сколько клауз в выражении *Развеселившись, я попросил стоявшего у пульта человека сделать музыку погромче*? a) 1 b) 2 c) 3 d) 4 e) 0 **Ключ:** d	+	+	+	+	+	+
25.	Что такое *разметка* в лингвистическом корпусе? - приписывание лингвистическим явлениям определенных свойств,	+	+	+	+	+	+

	служащих параметром поиска. - деление лингвистических явлений на составляющие их части. - подчеркивание лингвистических явлений определенным способом или их выделение определенным цветом. **Ключ:** приписывание лингвистическим явлениям определенных свойств, служащих параметром поиска						
26.	Что такое коллокация? - сочетание морфем внутри слова - сочетания параллельных текстов, наделенных сходными чертами - сочетание двух слов **Ключ:** сочетание двух слов	+	+	+	+	+	+
27.	Что такое *конкорданс*? Список слов, расположенных по алфавиту. Список слов, расположенных в зависимости от их частотности. Список контекстов, объединенных одним словом. Список правил, облегчающих поиск нужных слов в компьютерной программе. **Ключ:** Список контекстов, объединенных одним словом	+	+	+	+	+	+
28.	История сознания чат-ботов и нейросетей. Чем чат-бот отличается от нейросети? - тем, что чат-бот олицетворяет некую личность - тем, что чат-бот не может создавать голосовые сообщения - тем, что чат-бот не может поддерживать общение **Ключ:** тем, что чат-бот олицетворяет некую личность	+	+	+	+	+	+
29.	Чат-бот VS нейросеть. Особенности речи, сгенерированной искусственным интеллектом (синтаксический, морфологический, лексический, фонетический и статистический аспекты).	+	+	+	+	+	+
30.	Какое место занимает слово «чат-бот» относительно категории одушевленности?	+	+	+	+	+	+

	- оно является одушевленным - оно является неодушевленным - оно занимает промежуточное положение относительно категории одушевленности **Ключ:** оно занимает промежуточное положение относительно категории одушевленности						
31.	Методы распознавания и обработки текстов, расположите их по порядку а) токенизация б) «мешок слов» в) регулярные выражения г) автоматическая разметка семантических ролей **Ключ:** абвг	+	+	+	+	+	+
32.	Зачем нужен тест Тьюринга? - для того, чтобы отличить чат-бота от нейросети - для того, чтобы проверить способность чат-бота генерировать речь, схожую с речью человека - для того, чтобы проверить способность чат-бота решать сложные математические задачи **Ключ:** для того, чтобы проверить способность чат-бота генерировать речь, схожую с речью человека	+	+	+	+	+	+
33.	Доклад: Элиза и Пэрри – первые чат боты.	+	+	+	+	+	+
34.	Методы тестирования и моделирования речемыслительных способностей искусственного интеллекта. Что иллюстрируют метафора «португальского стадиона» и метафора «китайской комнаты»? - способность чат-бота говорить - способность чат-бота решать задачи - способность чат-бота мыслить	+	+	+	+	+	+

	Ключ: способность чат-бота мыслить						

3. ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ

Промежуточная аттестация осуществляется в форме (формах) зачета.

Способ проведения промежуточной аттестации – устный зачет.

Перечень видов оценочных средств, используемых для промежуточной аттестации по дисциплине: вопросы к зачету, задание к зачету.

3.1. Вопросы к зачету

№ п/п	Вопросы к зачету	Шифр индикатора достижения компетенции (ИДК)					
		УК-3.1	УК-3.2	УК-3.3	УК-3.4	ПК-2.1	ПК-2.2
1	Что изучает квантитативная лингвистика?	+	+	+	+	+	+
2	Критерий хи-квадрат? Зачем используется данный критерий в статистических исследованиях?	+	+	+	+	+	+
3	Фактор в статистических исследованиях? Уровень фактора в статистических исследованиях? Отклик в статистических исследованиях? Поясните эти термины на примере конкретного исследования, составив таблицу сопряженности.	+	+	+	+	+	+
4	Расчет среднеквадратичного отклонения в статистических исследованиях. Как расчет среднеквадратичного отклонения соотносится с критерием хи-квадрат? Что такое генеральная лингвистическая совокупность? Что такое выборочная лингвистическая совокупность?	+	+	+	+	+	+
5	Что дает для понимания статистических закономерностей метод расчета стандартизированных остатков?	+	+	+	+	+	+
6	Закон Ципфа. Почему именно Джорджа Ципфа называют отцом квантитативной лингвистики? «Метафора хоббита».	+	+	+	+	+	+
7	Атрибуция текста. В какой сфере она применяется? Какие виды атрибуции текста вы знаете? Чем приемы атрибуции древних рукописных памятников отличаются от атрибуции современных текстов? Какие вы	+	+	+	+	+	+

	знаете современные электронные ресурсы, помогающие установить авторство текста						
8	Основные вехи в развитии стилеметрии как науки.	+	+	+	+	+	+
9	Лингвистический корпус: его предназначение, виды.	+	+	+	+	+	+
10	Основные вехи в истории корпусной лингвистики.	+	+	+	+	+	+
11	Направления автоматизированной обработки текста на естественных языках. Искусственный интеллект. Основные этапы развития идеи искусственного интеллекта.	+	+	+	+	+	+
12	Реферат и аннотация. Основные характерные черты и виды этих документов. Основные этапы реферирования и аннотирования текста.	+	+	+	+	+	+

3.2. Задание к зачету:

Заданием к зачету служит работа над ошибками, допущенными в проектах по разделам № 1, 2 или 3 (либо письменное выполнение данной самостоятельной работе в случае ее отсутствия у студента).

№ п/п	Задания к зачету	Шифр индикатора достижения компетенции (ИДК)					
		УК-3.1	УК-3.2	УК-3.3	УК-3.4	ПК-2.1	ПК-2.2
1	Провести статистическое исследование однокоренных и разнокорневых, входящих в список Сводеша и заимствованных из трех родственных языков,	+	+	+	+	+	+
2	Основываясь на материалах Национального корпуса русского языка, проанализируйте поведение трех слов, объединенных синонимичными отношениями (Тельпов Р.Е. Космос, эфир и Вселенная – названия безграничных пространств в диахроническом аспекте / Корпус Современных Моделей и Кодов русского языка. сборник научных статей. Москва, 2024. С. 61-68.)	+	+	+	+	+	+
3	Проанализируйте и сопоставьте две любые группы текстов, одни	+	+	+	+	+	+

	из которых сгенерированы нейросетью, другие – созданы людьми. Сравните и обработайте статистические различия между ними (см. подробнее в качестве образца: Тельпов Р.Е. Особенности представления фактической информации в сгененрированных текстах / Коммуникативные исследования. 2025. Т. 12. № 4. С. 805-818.)						