

МИНОБРНАУКИ РОССИИ  
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ  
ВЫСШЕГО ОБРАЗОВАНИЯ  
«ВОРОНЕЖСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»  
(ФГБОУ ВО «ВГУ»)

УТВЕРЖДАЮ  
Заведующий кафедрой  
теоретической и прикладной лингвистики



Шилихина К.М.  
25.04.2025 г.

**РАБОЧАЯ ПРОГРАММА УЧЕБНОЙ ДИСЦИПЛИНЫ**  
Б1.О.13 Инструменты искусственного интеллекта для анализа языка и текста

**1. Код и наименование направления подготовки/специальности:**

45.05.01 Перевод и переводоведение

**2. Профиль подготовки/специализация:**

Перевод и переводоведение. Французский и испанский языки

**3. Квалификация выпускника: лингвист-переводчик**

**4. Форма обучения: очная**

**5. Кафедра, отвечающая за реализацию дисциплины:** кафедра теоретической и прикладной лингвистики

**6. Составители программы:** Подтегжникова Е.Н., кандидат филол. наук, доцент кафедры теоретической и прикладной лингвистики, Стародубцева Ю.А., кандидат филол. наук, Нагорная Е.В., преподаватель.

**7. Рекомендована:** НМС факультета РГФ, протокол № 8 от 18.04.2025 г.

**8. Учебный год:** 2025/2026

**Семестр(ы):** 2

## **9. Цели и задачи учебной дисциплины**

**Целью** изучения дисциплины является овладение с технологиями автоматической обработки естественного языка.

### **Задачи:**

- 1) овладение знаниями в области лингвистического обеспечения различных компьютерных систем и баз данных (автоматическая обработка языка и речи, языковые корпуса, информационно-поисковые, экспертные, лексикографические системы, системы машинного перевода и др.), а также знаниями о различных источниках информации и принципах работы современных информационных технологий при решении задач профессиональной деятельности;
- 2) формирование информационно-поисковых умений (поиск, сбор, хранение, обработка и представление информации при решении задач профессиональной деятельности);
- 3) формирование навыков использования технологии автоматической обработки естественного языка и искусственного интеллекта в различных сферах деятельности для решения познавательных задач.

## **10. Место учебной дисциплины в структуре ООП: блок Б1. Дисциплины (модули). Обязательная часть.**

Для успешного освоения дисциплины требуются базовые навыки работы с компьютером, владение лингвистической терминологией.

## **11. Планируемые результаты обучения по дисциплине/модулю (знания, умения, навыки), соотнесенные с планируемыми результатами освоения образовательной программы (компетенциями) и индикаторами их достижения:**

Код	Название компетенции	Код(ы)	Индикатор(ы)	Планируемые результаты обучения
ОПК –4	Способен работать с электронными словарями, различными источниками информации, осуществлять поиск, хранение, обработку и анализ информации, представлять ее в требуемом формате с использованием информационных, компьютерных и сетевых	ОПК-4.1	владеет знаниями о технологиях автоматической обработки естественного языка и искусственного интеллекта в различных сферах деятельности для решения познавательных задач, в том числе для анализа языка и текста	знать: принципы работы лингвистически ориентированных программных продуктов  уметь: работать с электронными словарями, различными источниками информации, осуществлять поиск, хранение, обработку и анализ информации, представлять ее в требуемом формате с использованием информационных, компьютерных и сетевых технологий  владеть: навыками использования различных справочных ресурсов, печатных и электронных словарей, сетевых, иных электронных и печатных источников в профессиональной деятельности для поиска необходимой лингвистической и экстралингвистической информации.

	технологий			
ОПК-5	Способен понимать принципы работы современных информационных технологий и использовать их для решения задач профессиональной деятельности	ОПК-5.1	ОПК-5.1 Критически оценивает современные информационные технологии	знать: основные методы использования современных информационных технологий  уметь: критически оценивать информационные технологии при решении задач профессиональной деятельности  владеть: навыком выбора оптимальных информационных технологий для решения задач профессиональной деятельности
ОПК-5	Способен понимать принципы работы современных информационных технологий и использовать их для решения задач профессиональной деятельности	ОПК-5.2	Подбирает оптимальные информационные технологии при решении задач профессиональной деятельности	знать: инструментарий для поиска, сбора, хранения, обработки, представления информации при решении задач профессиональной деятельности  уметь: подбирать и использовать информационные технологии при решении задач профессиональной деятельности  владеть: навыком выбора оптимальных информационных технологий для решения задач профессиональной деятельности

**12. Объем дисциплины в зачетных единицах/час.** — 2 з.е. / 72 ч.

**Форма промежуточной аттестации:** зачет.

**13. Трудоемкость по видам учебной работы**

Вид учебной работы	Трудоемкость			
	Всего	По семестрам		
		2 семестр	№ семестра	...
Контактная работа	34	34		
в том числе:	лекции	-	-	
	практические	-	-	
	лабораторные	34	34	
	курсовая работа	-	-	

	<i>др. виды(при наличии)</i>				
Самостоятельная работа		38	38		
Итого:		72	72		

### 13.1. Содержание дисциплины

п/п	Наименование раздела дисциплины	Содержание раздела дисциплины	Реализация раздела дисциплины с помощью онлайн-курса, ЭУМК *
1.	Искусственный интеллект в гуманитарной сфере	Проблемы искусственного интеллекта и типы решаемых задач. Введение в цифровую филологию. Направления цифровой филологии. Digital Humanities.	
2.	Лингвистические и филологические ресурсы и программы	Электронные лингвистические ресурсы. Лингвистические базы данных. Построение и применение лингвистических ресурсов. Цифровые библиотеки. Информационные системы в филологических задачах.	
3.	Компьютерная лексикография.	Компьютерная лексикография. Прикладные аспекты лексикографии. Электронные словари: основные типы словарей и принципы их организации. Структура словарной статьи электронного словаря. Электронные словари, доступные в сети. Частотные словари. Иноязычные словари. Тезаурусы.	
4.	Корпусная лингвистика	Корпусная лингвистика. Принципы организации корпуса. Основные свойства корпуса. Классификации корпусов. Современные корпуса текстов. Интернет как корпус. Национальный корпус русского языка и другие русскоязычные корпусы. Корпусы иных языков. Обработка корпусных данных. Работа с параллельными корпусами. Структура и назначение параллельных корпусов в решении задач ИИ. Google books Ngram Viewer.	
5.	Системы автоматической обработки письменного текста.	Автоматический анализ естественного языка. Общие проблемы автоматической обработки естественного языка. Автоматическая обработка письменного текста. Автоматическое распознавание	

		текста. Ввод, редактирование и форматирование текста (в среде Microsoft Word). Распознавание и графематический анализ текста (OCR).	
6.	Аспекты интеллектуального анализа текстов. Text Mining	Текст как объект интеллектуального анализа. Классификация и кластеризация текстов. Самообучающиеся системы: нейронные сети. Машинное обучение в лингвистике.	
7.	Визуализация текстовых данных	Визуализация данных (диаграммы, гистограммы, графики; облака слов). Современные автоматизированные средства визуализации данных.	
8.	Задачи информационного поиска с точки зрения лингвиста	Информационный поиск. Основные понятия информационного поиска. Лингвистические аспекты информационного поиска. Современные информационно-поисковые системы: принципы организации. Поиск научной литературы. Фильтрация и ранжирование.	
9.	Квантиративная лингвистика	Применение статистических методов в лингвистических исследованиях. Анализ данных в Microsoft Excel. Частотность.	
10.			
11.	ИИ в переводческой деятельности. Автоматизированный перевод. Машинный перевод	Автоматический перевод. История создания систем автоматического перевода. Современное состояние отрасли. Машинный перевод. Машинный vs автоматизированный перевод. Методы оценки качества перевода. Современные системы машинного перевода. Перспективы развития машинного перевода. Сравнение систем машинного перевода. Программные средства помощи переводчику.	
12.	ИИ в филологическом образовании	Проблемы компьютерной лингводидактики. Использование элементов ИИ в обучении языку. Компьютерные обучающие программы. Справочные системы для изучающих языки.	
13.	ИИ и социолингвистика.	Методика контент-анализа. Мультимодальная лингвистика. Проведение анкетирования. Работа с видео- и аудиоматериалами. Экспертные системы. Стилеметрия в переводоведении.	

### 13.2. Темы (разделы) дисциплины и виды занятий

№ п/п	Наименование темы (раздела) дисциплины	Виды занятий (количество часов)				
		Лекции	Практические	Лабораторные	Самостоятельная работа	Всего

1.1	Искусственный интеллект в гуманитарной сфере		2	2	4
1.2	Системы автоматической обработки письменного текста.		4	4	8
1.3	Аспекты интеллектуального анализа текстов. Text Mining		2	2	4
1.4	Визуализация текстовых данных		2	3	5
1.5	Задачи информационного поиска с точки зрения лингвиста		4	4	8
1.6	Квантитативная лингвистика		2	2	4
1.7	Лингвистические и филологические ресурсы и программы. Big Data		2	3	5
1.8	Корпусная лингвистика		6	6	12
1.9	Компьютерная лексикография.		2	3	5
1.10	ИИ в переводческой деятельности. Автоматизированный перевод. Машинный перевод		4	5	9
1.11	ИИ в филологическом образовании		2	2	4
1.12	ИИ и социолингвистика.		2	2	4
Итого:		34	38	72	

#### **14. Методические указания для обучающихся по освоению дисциплины**

Приступая к изучению дисциплины, обучающемуся необходимо внимательно ознакомиться с тематическим планом занятий, списком рекомендованной учебной литературы, уяснить последовательность выполнения индивидуальных (самостоятельных) учебных заданий.

Лабораторные занятия сопровождаются демонстрацией преподавателем презентационных материалов, раскрывающих основные вопросы темы. Обучающимся рекомендуется вести записи для лучшего усвоения материала.

При изучении учебной дисциплины особое внимание следует уделить приобретению навыков решения профессионально-ориентированных задач. Для этого, изучив материал данной темы, необходимо разобраться в решениях соответствующих задач, которые рассматривались на практических занятиях, приведены в учебно-методических материалах, пособиях, учебниках, обратив особое внимание на методические указания по их решению.

В качестве текущей аттестации обучающимся предлагается выполнение практических работ, обобщающих полученные знания, умения и навыки.

#### **15. Перечень литературы, ресурсов интернет, необходимых для освоения дисциплины**

№ п/п	Источник
1	Гуслякова А. В. Информационные технологии и лингвистика XXI века : учебное пособие / А. В. Гуслякова. - Москва : МПГУ, 2016. - 96 с. - URL: <a href="http://biblioclub.ru/index.php?page=book&amp;id=469675">http://biblioclub.ru/index.php?page=book&amp;id=469675</a> (05.08.2019).
2	Теория и практика машинного перевода : учебное пособие / авт.-сост. Э.В. Пиванова - Ставрополь : СКФУ, 2014. - 115 с. - URL: <a href="http://biblioclub.ru/index.php?page=book&amp;id=457763">http://biblioclub.ru/index.php?page=book&amp;id=457763</a>

3	Захаров, В.П. Корпусная лингвистика. Учебник для студентов гуманитарных вузов / В.П. Захаров ; Богданова С. Ю. – Иркутск : Иркутский государственный лингвистический университет, 2011. – 161 с. // URL: <a href="http://biblioclub.ru/index.php?page=book&amp;id=89753">http://biblioclub.ru/index.php?page=book&amp;id=89753</a>
4	Захарова Т. В. Практические основы компьютерных технологий в переводе : учебное пособие / Т. В. Захарова, Е. В. Турлова. - Оренбург : Оренбургский государственный университет, 2017. - 109 с. - URL: <a href="http://biblioclub.ru/index.php?page=book&amp;id=481823">http://biblioclub.ru/index.php?page=book&amp;id=481823</a> .
5	Копотев М. Введение в корпусную лингвистику / М. Копотев. - Прага : Animedia Company, 2014. - 195 с. - URL: <a href="http://biblioclub.ru/index.php?page=book&amp;id=375463">http://biblioclub.ru/index.php?page=book&amp;id=375463</a> (05.08.2019).
6	Моисеева И. Ю. Квантитативная лингвистика и новые информационные технологии : учебное пособие / И. Ю. Моисеева. - Оренбург : Оренбургский государственный университет, 2017. - 103 с. - URL: <a href="http://biblioclub.ru/index.php?page=book&amp;id=481797">http://biblioclub.ru/index.php?page=book&amp;id=481797</a> (07.08.2019).
7	Онтологии и тезаурусы: модели, инструменты, приложения : учебное пособие / Б. В. Добров, В. В. Иванов, Н. В. Лукашевич, В. Д. Соловьев. - Москва : Интернет-Университет Информационных Технологий, 2009. - 173 с. - URL: <a href="http://biblioclub.ru/index.php?page=book&amp;id=233056">http://biblioclub.ru/index.php?page=book&amp;id=233056</a> (07.08.2019).

в) информационные электронно-образовательные ресурсы (официальные ресурсы интернет)\*:

№ п/п	Ресурс
1.	ЭБС «Университетская библиотека онлайн». – Режим доступа: по подписке. – URL: ЭБС "Университетская библиотека онлайн" читать электронные книги ( <a href="http://biblioclub.ru">biblioclub.ru</a> )
2.	
3.	Филологический портал <a href="http://www.philology.ru">www.philology.ru</a>
4.	Международный лингвистический портал The LinguistList. URL: <a href="http://www.linguistlist.org">www.linguistlist.org</a>
5.	База данных языков мира SIL International. URL: <a href="http://www.sil.org/">http://www.sil.org/</a>
6.	Электронные словари Мультитран. URL: <a href="http://www.multitran.ru">www.multitran.ru</a>
7.	Национальный корпус русского языка. URL: <a href="http://www.ruscorpora.ru">www.ruscorpora.ru</a>
8.	The Corpus of Contemporary American English. URL: <a href="http://corpus.byu.edu">http://corpus.byu.edu</a>
9.	Лингвистика Интернета: формирование дисциплинарной парадигмы ( <a href="http://www.textology.ru/article.aspx?aid=76">http://www.textology.ru/article.aspx?aid=76</a> )
10.	Новые возможности лингвистических исследований по исторической семантике с применением электронных ресурсов ( <a href="http://textualheritage.org/content/view/74/68/lang,ru..">http://textualheritage.org/content/view/74/68/lang,ru..</a> )

11.	Языковые ресурсы: традиции и инновации ( <a href="http://elib.grsu.by/katalog/161659-346552.pdf">http://elib.grsu.by/katalog/161659-346552.pdf</a> )
12.	Контроль использования интернет-ресурсов ( <a href="http://alexott.net/ru/writings/cf/JI200502.pdf">http://alexott.net/ru/writings/cf/JI200502.pdf</a> )
13.	Интернет в работе переводчика ( <a href="http://samlib.ru/w/wagapow_a_s/tissen.shtml">http://samlib.ru/w/wagapow_a_s/tissen.shtml</a> )
14.	Компьютерная лингвистика ( <a href="http://www.krugosvet.ru/enc/gumanitarnye_nauki/lingvi..">http://www.krugosvet.ru/enc/gumanitarnye_nauki/lingvi..</a> )
15.	ICT ( <a href="http://iite.unesco.org/pics/publications/en/files/321..">http://iite.unesco.org/pics/publications/en/files/321..</a> )
16.	Создание автоматического оглавления в MS Word 2007. Режим доступа: <a href="http://www.mbooom.ru/news/urok_23_sozdanie_avtomaticheskogo_oglavljenija_v_ms_word_2007/2011-07-25-104">http://www.mbooom.ru/news/urok_23_sozdanie_avtomaticheskogo_oglavljenija_v_ms_word_2007/2011-07-25-104</a>
17.	Текстовое поле. Как убрать или добавить. Режим доступа: <a href="http://informatio.ru/news/programma_ms_word_2007/tekstovoe_pole_kak_ubrat_ili_dobavit_podlozhku_v_word_2007/">http://informatio.ru/news/programma_ms_word_2007/tekstovoe_pole_kak_ubrat_ili_dobavit_podlozhku_v_word_2007/</a>
18.	Инструкция по использованию программ конкорданс. Режим доступа: <a href="https://eflnotes.wordpress.com/2013/03/06/building-your-own-corpus-textstat-antconc/">https://eflnotes.wordpress.com/2013/03/06/building-your-own-corpus-textstat-antconc/</a>
19.	Использование ABBY Lingvo. Режим доступа: <a href="http://www.ciao-italy.ru/materialy-i-sredstva/abbyy-lingvo-tutor-poshagovaya-instrukciya.html">http://www.ciao-italy.ru/materialy-i-sredstva/abbyy-lingvo-tutor-poshagovaya-instrukciya.html</a>

## 16. Перечень учебно-методического обеспечения для самостоятельной работы

№ п/п	Источник
1.	Баранов А. Н. Лингвистическая экспертиза текста. Теоретические основания и практика : учебное пособие / А. Н. - 2-е изд. - М. : Флинта : Наука, 2009. - 591 с.
2.	Автоматическая обработка текстов на естественном языке и компьютерная лингвистика : учеб. пособие / Большакова Е.И., Клышинский Э.С., Ландэ Д.В., Носков А.А., Пескова О.В., Ягунова Е.В. — М.: МИЭМ, 2011. — 272 с.
3.	Всеволодова А.В. Компьютерная обработка лингвистических баз данных: учебное пособие: для студентов, аспирантов, преподавателей-филологов / А.В. Всеволодова .— 2-е изд., испр. — М. : Флинта : Наука, 2007 .— 90 с.
4.	Ляшевская О. Н. Корпусные инструменты в грамматических исследованиях русского языка. / О.Н. Ляшевская. - Москва : Издательский Дом ЯСК : Рукописные памятники Древней Руси, 2016. - 520 с.
5.	Маннинг Кристофер Д. Введение в информационный поиск = Introduction to Information retrieval / Кристофер Д. Маннинг, ПрабхакарРагхаван, ХайнрихШютце ; [пер. с англ. Д.А. Клюшина] .— М. ; СПб. ; Киев : Вильямс, 2011. - 520 с.
6.	Марчук Ю. Н. Компьютерная лингвистика : учебное пособие для студентов вузов, специализирующихся по направлению и специальности "Филология" / Ю.Н. Марчук. - М. : Восток-Запад, 2007. - 317 с.
7.	Новожилова, А.А. Информационные технологии в переводе. - Учебно-метод. пособие /, А.А. Новожилова, Е.В. Степанова, Е.А. Шовгенина - Волгоград: Изд-во ВолГУ, 2012. – 162 с.
8.	Потапова Р. К. Новые информационные технологии и лингвистика: учебное пособие для студ. вузов / Р.К. Потапова; Моск. гос. лингв. ун-т .— Изд. 2-е .— М.: Едиториал УРСС, 2004 .— 317 с.
9.	Потапова Р. К. Новые информационные технологии и лингвистика: учебное пособие для студ. вузов / Р.К. Потапова; Моск. гос. лингв. ун-т.— Изд. 2-е .—

**17. Образовательные технологии, используемые при реализации учебной дисциплины, включая дистанционные образовательные технологии (ДОТ), электронное обучение (ЭО), смешанное обучение):**

При реализации дисциплины могут применяться дистанционные образовательные технологии в части освоения материала, проведения текущей аттестации, самостоятельной работы по дисциплине или отдельным ее разделам и т.д. Для освоения материала на практических занятиях используются задания, предназначенные как для индивидуального решения задач, так и для коллективного обсуждения стратегии решения той или иной задачи.

Также используется следующее программное обеспечение:

1. Неисключительные права на ПО Dr. Web Enterprise Security Suite  
Комплексная защита Dr. Web Desktop Security Suite.
2. Программная система для обнаружения текстовых заимствований в учебных и научных работах Антиплагиат.ВУЗ
3. Программное обеспечение Microsoft Windows

**18. Материально-техническое обеспечение дисциплины:**

/ауд. 12/ - компьютерный класс: Компьютер Arbyte Темпо/AOC (12 шт.), Проектор Benq MW523 (1 шт.), Экран проекционный (1 шт.)	г.Воронеж, пл.Ленина 10, ауд.12
---	------------------------------------

**19. Оценочные средства для проведения текущего контроля успеваемости и промежуточной аттестации**

Порядок оценки освоения обучающимися учебного материала определяется содержанием следующих разделов дисциплины:

N п / п	Наименование раздела дисциплины (модуля)	Компетенция(и)	Индикатор(ы) достижения компетенции	Оценочные средства
------------------	---	----------------	---	--------------------

<p>Искусственный интеллект в гуманитарной сфере Системы автоматической обработки письменного текста. Аспекты интеллектуального анализа текстов. Text Mining Визуализация текстовых данных Задачи информационного поиска с точки зрения лингвиста Квантиративная лингвистика Лингвистические и филологические ресурсы и программы. Big Data Корпусная лингвистика Компьютерная лексикография. ИИ в переводческой деятельности. Автоматизированный перевод. Машинный перевод ИИ в филологическом образовании ИИ и социолингвистика.</p>	<p>ОПК-4 ОПК-5</p>	<p>ОПК-4.1 ОПК-5.1 ОПК-5.2</p>	<p>Практикоориентированные задания</p>
<p>Промежуточная аттестация форма контроля – зачет</p>			<p>Теоретический вопрос, Практикоориентированное задание</p>

## 20 Типовые оценочные средства и методические материалы, определяющие процедуры оценивания

### 20.1 Текущий контроль успеваемости

Контроль успеваемости по дисциплине осуществляется с помощью следующих оценочных средств: практикоориентированные задания

**Примеры практикоориентированных заданий:**

#### 1. Лингвистические корпусы текстов. Освоение навигации по НКРЯ

- Найдите примеры с выбранной вами лексической единицей в параллельном подкорпусе НКРЯ. Определите, как можно применить полученную информацию при обучении переводу.
- Автоматически выгрузите из НКРЯ в Excel корпус параллельных предложений с лексической единицей объемом 50 пар предложений. Проанализируйте переводные эквиваленты, выделите переводческие трансформации. Классифицируйте лексический материал, отсортируйте по группам. Проведите статистический анализ полученных групп, результаты представьте в виде диаграммы.

## **2. Машинный перевод**

Сравните системы машинного перевода. Результаты анализа представьте в виде реферата в текстовом процессоре Microsoft Word, оформив его в соответствии с требованиями:

- a. Создайте титульный лист, содержание и введение в соответствии с темой вашей работы.
- b. Текст работы должен соответствовать следующим требованиям:
  - шрифт Times New Roman, размер 14, через 1,5 интервал, стиль "Обычный" (весь текст);
  - абзац - 0,5 см; выравнивание по ширине;
  - все инициалы в тексте набираются через неразрывный пробел ([Ctrl + Shift + Пробел] или команда "Вставка"- "Символ" - "Специальные символы" - "Неразрывный пробел");
  - используйте список
  - вставьте номера страниц
- c. Список литературы дается по алфавиту в конце статьи с обязательной нумерацией и озаглавливается ЛИТЕРАТУРА (выравнивание по центру);
- d. Ссылки на литературу оформляются следующим образом: Л.В. Щерба отмечал: "... [Щерба 1990, с. 1-5].

### **Требования к выполнению заданий (или шкалы и критерии оценивания):**

Для оценивания результатов обучения на зачете используется 2-балльная шкала: «зачтено», «не зачтено».

Соотношение показателей, критериев и шкалы оценивания результатов обучения.

Критерии оценивания компетенций	Уровень сформированности компетенций	Шкала оценок
Обучающийся в полной мере владеет понятийным аппаратом данной области науки (теоретическими основами дисциплины), способен иллюстрировать ответ примерами, фактами, данными научных исследований, применять теоретические знания для решения практических задач в области прикладной лингвистики.	Повышенный уровень	Зачтено
Обучающийся владеет понятийным аппаратом данной области науки (теоретическими основами дисциплины), способен иллюстрировать ответ примерами, но допускает ошибки при применении теоретических знаний для решения практических задач в области прикладной лингвистики.	Базовый уровень	Зачтено
Обучающийся владеет частично теоретическими основами дисциплины, фрагментарно способен иллюстрировать ответ примерами, фактами, данными научных исследований, не умеет применять теоретические знания для решения практических задач в области прикладной лингвистики.	Пороговый уровень	Зачтено
Ответ на контрольно-измерительный материал не соответствует любым четырем из перечисленных	-	Не зачтено

показателей. Обучающийся демонстрирует отрывочные, фрагментарные знания, допускает грубые ошибки.		
---	--	--

## 20.2 Промежуточная аттестация

Промежуточная аттестация по дисциплине осуществляется с помощью следующих оценочных средств: Собеседование по билетам к зачету, Практико-ориентированные задания.

### Перечень вопросов к зачету

1. Проблемы искусственного интеллекта и типы решаемых задач. Гуманитарная экспертиза технологических изменений. Этические аспекты функционирования ИИ.
2. Введение в цифровую филологию. Направления цифровой филологии. Digital Humanities.
3. Автоматический анализ естественного языка. Общие проблемы автоматической обработки естественного языка.
4. Автоматическая обработка письменного текста. Автоматическое распознавание текста.
5. Текст как объект интеллектуального анализа. Классификация и кластеризация текстов.
6. Самообучающиеся системы: нейронные сети. Машинное обучение в лингвистике.
7. Визуализация данных. Современные автоматизированные средства визуализации данных.
8. Информационный поиск. Основные понятия информационного поиска. Лингвистические аспекты информационного поиска.
9. Современные информационно-поисковые системы: принципы организации. Фильтрация и ранжирование.
10. Применение статистических методов в лингвистических исследованиях. Частотность.
11. Электронные лингвистические ресурсы. Лингвистические базы данных. Построение и применение лингвистических ресурсов.
12. Цифровые библиотеки. Информационные системы в филологических задачах.
13. Открытые данные и проблема их обработки. Верификация данных. Большие данные в филологии и языкоznании.
14. Корпусная лингвистика. Принципы организации корпуса.
15. Основные свойства корпуса. Классификации корпусов.
16. Современные корпуса текстов. Интернет как корпус. Национальный корпус русского языка и другие русскоязычные корпусы. Корпусы иных языков.
17. Обработка корпусных данных. Работа с параллельными корпусами. Структура и назначение параллельных корпусов в решении задач ИИ.
18. Компьютерная лексикография. Прикладные аспекты лексикографии.
19. Электронные словари: основные типы словарей и принципы их организации. Структура словарной статьи электронного словаря. Электронные словари, доступные в сети.
20. Частотные словари. Иноязычные словари. Лексическая информация в системах ИИ.
21. Автоматический перевод. История создания систем автоматического перевода. Современное состояние отрасли.

22. Машинный перевод. Машинный vs автоматизированный перевод.
23. Современные системы машинного перевода. Перспективы развития машинного перевода.
24. Методы оценки качества перевода. Сравнение систем машинного перевода. Программные средства помощи переводчику.
25. Проблемы компьютерной лингводидактики. Использование элементов ИИ в обучении языку.
26. Компьютерные обучающие программы. Справочные системы для изучающих язык.
27. Методика контент-анализа. Проведение анкетирования.
28. Мультимодальная лингвистика. Работа с видео- и аудиоматериалами.
29. Экспертные системы. Стилеметрия в переводоведении.

**Примеры практикоориентированных заданий:**  
**Лингвистические корпусы текстов. Освоение навигации по НКРЯ**

На примере выбранной вами лексической единицы покажите организацию различных видов разметки в НКРЯ.

Оценка знаний, умений и навыков, характеризующая этапы формирования компетенций в рамках изучения дисциплины осуществляется в ходе текущей и промежуточной аттестаций.

Текущая аттестация проводится в соответствии с Положением о текущей аттестации обучающихся по программам высшего образования Воронежского государственного университета. Текущая аттестация проводится в форме выполнения практикоориентированных заданий. Критерии оценивания приведены выше.

Промежуточная аттестация проводится в соответствии с Положением о промежуточной аттестации обучающихся по программам высшего образования.

Контрольно-измерительные материалы промежуточной аттестации включают в себя теоретический вопрос, позволяющий оценить уровень полученных знаний, а также практикоориентированное задание. При оценивании используются качественные шкалы оценок. Критерии оценивания приведены выше.

**20.3 Фонд оценочных средств сформированности компетенций студентов, рекомендуемый для проведения диагностических работ**

ОПК-4 Способен работать с электронными словарями, различными источниками информации, осуществлять поиск, хранение, обработку и анализ информации, представлять ее в требуемом формате с использованием информационных, компьютерных и сетевых технологий

ОПК-4.1 владеет знаниями о технологиях автоматической обработки естественного языка и искусственного интеллекта в различных сферах деятельности для решения познавательных задач, в том числе для анализа языка и текста

ОПК-5 Способен понимать принципы работы современных информационных технологий и использовать их для решения задач профессиональной деятельности

ОПК-5.1 критически оценивает современные информационные технологии

ОПК-5.2 подбирает оптимальные информационные технологии при решении задач профессиональной деятельности

**Перечень заданий для оценки сформированности компетенции:**

1. Метод автоматического аннотирования, при котором важные слова выделяются в заголовке, подзаголовке, начале и конце текста, называется ...
  - а) статистическим;
  - б) логико-семантическим;
  - в) позиционным;
  - г) функциональным.
2. Слово, относящееся к основному содержанию текста и повторяющееся в нем несколько раз, в автоматическом реферировании называется ...
  - а) важным словом;
  - б) термином;
  - в) главным словом;
  - г) ключевым словом.
3. Реферат — это...
  - а) связный текст, который кратко выражает тему, предмет, цель, методы и результаты исследования;
  - б) процесс составления содержания документа (книги, статьи, патента на изобретение и др.);
  - с) краткое изложение содержания документа, дающее общее представление о его теме;
  - д) краткий текст, выполняющий сигнальную функцию (информирует о том, что есть публикация на определенную тему).
4. Совокупность специально отобранных текстов, размещенных по различным лингвистическим параметрам и обеспеченных системой поиска, называется ...
  - а) базой данных;
  - б) словарем;
  - в) информационным массивом;
  - г) корпусом.
5. Разметка бывает ...
  - а) морфологической; синтаксической; семантической и просодической;
  - б) полнотекстовой и фрагментной;
  - с) синхронической и диахронической;
  - д) звуковой, письменной, смешанной.
6. Символ \* в поисковом запросе заменяет
  - а) один символ;
  - б) одно слово;
  - в) любое количество символов в словоформе;
  - г) любое количество символов в предложении.
7. С появлением корпусов лингвисты получили возможность исследовать:
  - а) языковую норму;
  - б) реальное употребление языковых единиц;
  - в) ошибки в речи носителей языка;
  - г) частотность грамматических конструкций.
8. Самый первый языковой корпус содержал
  - а) 100 млн. словоупотреблений;
  - б) 1 млн. текстов;
  - в) 1 млн. размеченных предложений;
  - г) 1 млн. словоупотреблений;
9. Что такое метаразметка корпуса?
  - а) грамматический анализ предложения;
  - б) характеристика текста;
  - в) информация о

b) информация о свойствах словоформ; d) информация о частях речи.

10. Морфологическая и синтаксическая разметка обеспечивают

- |   |                              |
|---|------------------------------|
| a) возможность грамматического анализа предложения; | c) поиск точных форм слов;   |
| b) автоматический поиск грамматической информации   | d) поиск начальных форм слов |

Критерии и шкалы оценивания:

Для оценивания выполнения заданий используется балльная шкала:

**1) тестовые задания:**

- средний уровень сложности (в формулировке задания перечислены все варианты ответа 1 балл – указан верный ответ;
- 0 баллов – указан неверный ответ, в том числе частично.
- повышенный уровень сложности (в формулировке задания отсутствуют варианты ответа 2 балла – указан верный ответ;
- 0 баллов – указан неверный ответ, в том числе частично.

**Задания раздела 20.3 рекомендуются к использованию при проведении диагностических работ с целью оценки остаточных результатов освоения данной дисциплины (знаний, умений, навыков).**