

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ
ФЕДЕРАЦИИ
федеральное государственное автономное образовательное учреждение высшего
образования
"САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
АЭРОКОСМИЧЕСКОГО ПРИБОРОСТРОЕНИЯ"

Кафедра № 63

УТВЕРЖДАЮ
Ответственный за образовательную
программу

доц., к.ф.н., доц.

(должность, уч. степень, звание)

М.А. Чиханова

(инициалы, фамилия)

(подпись)

«20» февраля 2025 г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

«Автоматизированная обработка письменного языка»
(Наименование дисциплины)


| | |
|---|--|
| Код направления подготовки/ специальности | 45.03.02 |
| Наименование направления подготовки/ специальности | Лингвистика |
| Наименование направленности | Теоретическая и прикладная лингвистика |
| Форма обучения | заочная |
| Год приема | 2025 |

Санкт-Петербург– 2025

Лист согласования рабочей программы дисциплины

Программу составил (а)

доц., к.э.н.
(должность, уч. степень, звание)

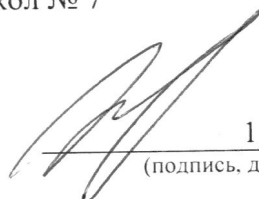
 14.02.2025
(подпись, дата)

О.В. Злобина
(инициалы, фамилия)

Программа одобрена на заседании кафедры № 63
«14» февраля 2025 г, протокол № 7

Заведующий кафедрой № 63

доц., к.ф.н., доц.
(должность, уч. степень, звание)

 14.02.2025
(подпись, дата)

М.А. Чиханова
(инициалы, фамилия)

Заместитель декана факультета №6 по методической работе

проф., д.и.н., доц.
(должность, уч. степень, звание)

 14.02.2025
(подпись, дата)

Л.Ю. Гусман
(инициалы, фамилия)

Аннотация

Дисциплина «Автоматизированная обработка письменного языка» входит в образовательную программу высшего образования – программу бакалавриата по направлению подготовки/ специальности 45.03.02 «Лингвистика» направленности «Теоретическая и прикладная лингвистика». Дисциплина реализуется кафедрой «№63».

Дисциплина нацелена на формирование у выпускника следующих компетенций:

УК-1 «Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач»

ПК-3 «Готовность проводить лингвистические экспертизы»

ПК-7 «Готовность овладеть основными математико-статистическими методами обработки лингвистической информации с учетом элементов программирования и автоматической обработки лингвистических корпусов»

Содержание дисциплины охватывает круг вопросов, связанных с представлениями о современных лингвистических и информационных технологиях, использующихся при разработке систем автоматизированной обработке естественной речи и, в частности, письменного языка, а также с автоматическим морфологическим, синтаксическим и семантическим анализом письменных текстов, в том числе анализ на основе корпусов текстов.

Преподавание дисциплины предусматривает следующие формы организации учебного процесса: *лекции, лабораторные работы, самостоятельная работа обучающегося.*

Программой дисциплины предусмотрены следующие виды контроля: текущий контроль успеваемости, промежуточная аттестация в форме зачета.

Общая трудоемкость освоения дисциплины составляет 2 зачетных единицы, 72 часа.

Язык обучения по дисциплине «русский».

1. Перечень планируемых результатов обучения по дисциплине

1.1. Цели преподавания дисциплины - систематизация знаний по проблемам алгоритмизации и моделирования для решения лингвистических задач, с помощью современного языка программирования для использования в будущей профессиональной деятельности; ознакомление студентов с существующими сегодня возможностями использования систем автоматической обработки письменного языка. В ходе изучения курса решаются следующие задачи: – ввести основные понятия, необходимые для работы с соответствующими программами и вспомогательными средствами; обучить студентов приемам эффективного поиска необходимой терминологии с использованием электронных словарей, глоссариев, баз данных, корпусов.

К задачам дисциплины следует отнести также и следующие: 1) углубить знания о применении информационных технологий для решения прикладных лингвистических задач; 2) изучить специфику письменных и устных текстовых массивов; 3) сформировать умения и навыки обработки письменных и устных текстовых массивов с помощью специализированных компьютерных программ

1.2. Дисциплина входит в состав части, формируемой участниками образовательных отношений, образовательной программы высшего образования (далее – ОП ВО).

1.3. Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения ОП ВО.

В результате изучения дисциплины обучающийся должен обладать следующими компетенциями или их частями. Компетенции и индикаторы их достижения приведены в таблице 1.

Таблица 1 – Перечень компетенций и индикаторов их достижения

| Категория (группа) компетенции | Код и наименование компетенции | Код и наименование индикатора достижения компетенции |
|--------------------------------|---|---|
| Универсальные компетенции | УК-1 Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач | УК-1.3.1 знать методики поиска, сбора и обработки информации, в том числе с использованием информационных технологий, включая интеллектуальные |
| Профессиональные компетенции | ПК-3 Готовность проводить лингвистические экспертизы | ПК-3.3.1 знать методы лингвистических экспертиз, методы фоносемантического и ассоциативного анализа лингвистических объектов ПК-3.У.1 уметь применять различные методы лингвистических экспертиз, оценивая соответствие лингвистического объекта кодифицированным нормам современного русского языка ПК-3.В.1 владеть методами проведения лингвистических экспертиз, навыками анализа макроструктуры и микроструктуры текстов в практике решения экспертных задач |
| Профессиональные компетенции | ПК-7 Готовность овладеть основными | ПК-7.У.1 уметь с учетом элементов программирования и автоматической обработки лингвистических корпусов |

| | | |
|--|--|---|
| | математико-статистическими методами обработки лингвистической информации с учетом элементов программирования и автоматической обработки лингвистических корпусов | обрабатывать лингвистическую информацию ПК-7.В.1 владеть навыками программирования и навыками автоматической обработки корпусов |
|--|--|---|

2. Место дисциплины в структуре ОП

Дисциплина может базироваться на знаниях, ранее приобретенных обучающимися при изучении следующих дисциплин:

- Информационный поиск и извлечение информации
- Компьютерные программы в лингвистических исследованиях

Знания, полученные при изучении материала данной дисциплины, имеют как самостоятельное значение, так и могут использоваться при написании ВКР и изучении других дисциплин:

- Автоматизированный перевод
- Производственная преддипломная практика.

3. Объем и трудоемкость дисциплины

Данные об общем объеме дисциплины, трудоемкости отдельных видов учебной работы по дисциплине (и распределение этой трудоемкости по семестрам) представлены в таблице 2.

Таблица 2 – Объем и трудоемкость дисциплины

| Вид учебной работы | Всего | Трудоемкость по семестрам |
|---|-------|---------------------------|
| | | №5 |
| 1 | 2 | 3 |
| Общая трудоемкость дисциплины, ЗЕ/ (час) | 2/ 72 | 2/ 72 |
| Из них часов практической подготовки | 9 | 9 |
| Аудиторные занятия, всего час. | 24 | 24 |
| в том числе: | | |
| лекции (Л), (час) | 10 | 10 |
| практические/семинарские занятия (ПЗ), (час) | | |
| лабораторные работы (ЛР), (час) | 14 | 14 |
| курсовой проект (работа) (КП, КР), (час) | | |
| экзамен, (час) | | |
| Самостоятельная работа, всего (час) | 48 | 48 |
| Вид промежуточной аттестации: зачет, дифф. зачет, экзамен (Зачет, Дифф. зач, Экз.**) | Зачет | Зачет |

Примечание: ** кандидатский экзамен

4. Содержание дисциплины

4.1. Распределение трудоемкости дисциплины по разделам и видам занятий.
Разделы, темы дисциплины и их трудоемкость приведены в таблице 3.

Таблица 3 – Разделы, темы дисциплины, их трудоемкость

| Разделы, темы дисциплины | Лекции (час) | ПЗ (СЗ) (час) | ЛР (час) | КП (час) | СРС (час) |
|---|-----------------|------------------|-------------|-------------|--------------|
| Семестр 5 | | | | | |
| Раздел 1. Этапы анализа текста Тема 1.1 Морфологический анализ Тема 1.2 Словарный морфологический анализ | 2 | | 4 | | 10 |
| Раздел 2. Постморфологический и предсинтаксический анализ Тема 2.1 Постморфологический анализ Тема 2.2 Синтаксическая сегментация и синтаксический анализ Тема 2.3 Автоматическое снятие омонимии | 3 | | 6 | | 12 |
| Раздел 3. Программные средства лингвистической обработки Тема 3.1 Способы представления лингвистических данных Тема 3.2 Лингвистическая разметка Тема 3.2 Лингвистические аннотации Тема 3.3 Анализ корпуса | 3 | | 4 | | 14 |
| Раздел 4 Статистические и математические методы, используемые при обработке текста | 2 | | | | 12 |
| Итого в семестре: | 10 | | 14 | | 48 |
| Итого | 10 | 0 | 14 | 0 | 48 |
| | | | | | |

Практическая подготовка заключается в непосредственном выполнении обучающимися определенных трудовых функций, связанных с будущей профессиональной деятельностью.

4.2. Содержание разделов и тем лекционных занятий.
Содержание разделов и тем лекционных занятий приведено в таблице 4.

Таблица 4 – Содержание разделов и тем лекционного цикла

| Номер раздела | Название и содержание разделов и тем лекционных занятий |
|---------------|---|
| 1 | Этапы анализа текста: этапы анализа текста системами NLP. На лексическом и морфологическом уровне: отличие морфологического анализа от токенизации, лемматизация, автоматическое определение частей речи, различие морфологии различных по структуре языков как проблема NLP, автоматическое определение частей речи, отличие контекстно обусловленного и контекстно необусловленного анализа, проблема имен собственных. |
| 2 | Постморфологический и предсинтаксический анализ. Решение проблем морфологической омонимии, автоматическое снятие проблем омонимии, синтаксическая сегментация, роль корпусов в анализе синтаксических конструкций, взаимосвязь синтаксиса и морфологии, снятие синтаксической омонимии. |

| | |
|---|---|
| 3 | Программные средства лингвистической обработки. Глоссарии, словари, корпуса, данные сети Интернет как материал для лингвистического анализа; системы, использующие определенные виды данных. Лингвистическая разметка, лингвистические аннотации, проблема анализа тональности. Корпусы как средство машинного обучения |
| 4 | Статистические и математические методы, используемые при обработке текста. Частотный анализ, ранжирование данных, контент-анализ, вероятностно-статистический анализ, в том числе закон Бедфорда. |

4.3. Практические (семинарские) занятия

Темы практических занятий и их трудоемкость приведены в таблице 5.

Таблица 5 – Практические занятия и их трудоемкость

| № п/п | Темы практических занятий | Формы практических занятий | Трудоемкость, (час) | Из них практической подготовки, (час) | № раздела дисциплины |
|---------------------------------|---------------------------|----------------------------|---------------------|---------------------------------------|----------------------|
| Учебным планом не предусмотрено | | | | | |
| | | | | | |
| Всего | | | | | |

4.4. Лабораторные занятия

Темы лабораторных занятий и их трудоемкость приведены в таблице 6.

Таблица 6 – Лабораторные занятия и их трудоемкость

| № п/п | Наименование лабораторных работ | Трудоемкость, (час) | Из них практической подготовки, (час) | № раздела дисциплины |
|-----------|--|---------------------|---------------------------------------|----------------------|
| Семестр 5 | | | | |
| 1 | Токенизация и различие морфологических систем языков как проблема NLP | 2 | 2 | 1 |
| 2 | Проблема имен собственных | 2 | 1 | 1 |
| 3 | Виды омонимии, морфологическая омонимия, способы снятия морфологической омонимии | 2 | 1 | 2 |
| 4 | Роль корпусов в анализе синтаксических конструкций | 2 | 1 | 2 |
| 5 | Снятие синтаксической омонимии | 2 | 1 | 2 |
| 6 | Способы обработки глоссариев, словарей, корпусов | 2 | 1 | 3 |
| 7 | Лингвистическая разметка | 2 | 2 | 3 |
| Всего | | 14 | 9 | |

4.5. Курсовое проектирование/ выполнение курсовой работы

Учебным планом не предусмотрено

4.6. Самостоятельная работа обучающихся

Виды самостоятельной работы и ее трудоемкость приведены в таблице 7.

Таблица 7 – Виды самостоятельной работы и ее трудоемкость

| Вид самостоятельной работы | Всего, час | Семестр 5, час |
|---|---------------|-------------------|
| 1 | 2 | 3 |
| Изучение теоретического материала дисциплины (ТО) | 20 | 20 |
| Курсовое проектирование (КП, КР) | | |
| Расчетно-графические задания (РГЗ) | | |
| Выполнение реферата (Р) | | |
| Подготовка к текущему контролю успеваемости (ТКУ) | 5 | 5 |
| Домашнее задание (ДЗ) | 15 | 15 |
| Контрольные работы заочников (КРЗ) | 4 | 4 |
| Подготовка к промежуточной аттестации (ПА) | 4 | 4 |
| Всего: | 48 | 48 |

5. Перечень учебно-методического обеспечения

для самостоятельной работы обучающихся по дисциплине (модулю)

Учебно-методические материалы для самостоятельной работы обучающихся указаны в п.п. 7-11.

6. Перечень печатных и электронных учебных изданий

Перечень печатных и электронных учебных изданий приведен в таблице 8.

Таблица 8– Перечень печатных и электронных учебных изданий

| Шифр/ URL адрес | Библиографическая ссылка | Количество экземпляров в библиотеке (кроме электронных экземпляров) |
|---|--|---|
| 8 Э 34 | Лингвистический анализ текста: практикум / П. М. Эйсмонт ; С.- Петерб. гос. ун-т аэрокосм. приборостроения. - СПб. : Изд-во ГУАП, 2018. - 88 с. : ил. - Б. ц. - Текст : непосредственный | ГС (10), БМЧЗ (3) |
| 8 Л 59 | Лингвистический и переводческий анализ текста: учебно-метод. пособие/под ред. М.А.Чихановой - Спб.: ГУАП, 2019. – 125с. | ФО(5) |
| https://rusgram.narod.ru/index1 . | Русская грамматика. 1980 г. | |

| | | |
|---|------------------------------------|--|
| html | | |
| https://www.machinelearningplu.s.com/nlp/natural-language-processing-guide/ | Основы анализа естественных языков | |

7. Перечень электронных образовательных ресурсов информационно-телекоммуникационной сети «Интернет»

Перечень электронных образовательных ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины приведен в таблице 9.

Таблица 9 – Перечень электронных образовательных ресурсов информационно-телекоммуникационной сети «Интернет»

| URL адрес | Наименование |
|---|---|
| http://constructiongrammar.org/ | Ресурс и библиография по грамматике конструкций |
| https://ruscorpora.ru/new/ | Национальный корпус русского языка |
| https://www.english-corpora.org/bnc/ | Национальный корпус английского языка |
| http://constructiongrammar.org/ | Ресурс и библиография по грамматике конструкций |
| https://www.elibrary.ru/ | eLIBRARY.RU – научная электронная библиотека |
| https://nlp-training.org/free-nlp-techniques/ | Free NLP Techniques. Top 10 NLP Techniques |
| https://scholar.google.com | Академия Google |

8. Перечень информационных технологий

8.1. Перечень программного обеспечения, используемого при осуществлении образовательного процесса по дисциплине.

Перечень используемого программного обеспечения представлен в таблице 10.

Таблица 10– Перечень программного обеспечения

| № п/п | Наименование |
|-------|------------------|
| | Не предусмотрено |

8.2. Перечень информационно-справочных систем,используемых при осуществлении образовательного процесса по дисциплине

Перечень используемых информационно-справочных систем представлен в таблице 11.

Таблица 11– Перечень информационно-справочных систем

| № п/п | Наименование |
|-------|------------------|
| | Не предусмотрено |

9. Материально-техническая база

Состав материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине, представлен в таблице12.

Таблица 12 – Состав материально-технической базы

| № п/п | Наименование составной части материально-технической базы | Номер аудитории (при необходимости) |
|----------|--|--|
| 1 | Лекционная аудитория | |
| 2 | Лаборатория формальной лингвистики | Г., 34-10 |

10. Оценочные средства для проведения промежуточной аттестации

10.1. Состав оценочных средств для проведения промежуточной аттестации обучающихся по дисциплине приведен в таблице 13.

Таблица 13 – Состав оценочных средств для проведения промежуточной аттестации

| Вид промежуточной аттестации | Перечень оценочных средств |
|------------------------------|----------------------------|
| Зачет | Список вопросов; Тесты |

10.2. В качестве критериев оценки уровня сформированности (освоения) компетенций обучающимися применяется 5-балльная шкала оценки сформированности компетенций, которая приведена в таблице 14. В течение семестра может использоваться 100-балльная шкала модульно-рейтинговой системы Университета, правила использования которой, установлены соответствующим локальным нормативным актом ГУАП.

Таблица 14 – Критерии оценки уровня сформированности компетенций

| Оценка компетенции 5-балльная шкала | Характеристика сформированных компетенций |
|--|---|
| «отлично» «зачтено» | <ul style="list-style-type: none"> – обучающийся глубоко и всесторонне усвоил программный материал; – уверенно, логично, последовательно и грамотно его излагает; – опираясь на знания основной и дополнительной литературы, тесно привязывает усвоенные научные положения с практической деятельностью направления; – умело обосновывает и аргументирует выдвигаемые им идеи; – делает выводы и обобщения; – свободно владеет системой специализированных понятий. |
| «хорошо» «зачтено» | <ul style="list-style-type: none"> – обучающийся твердо усвоил программный материал, грамотно и по существу излагает его, опираясь на знания основной литературы; – не допускает существенных неточностей; – увязывает усвоенные знания с практической деятельностью направления; – аргументирует научные положения; – делает выводы и обобщения; – владеет системой специализированных понятий. |
| «удовлетворительно» «зачтено» | <ul style="list-style-type: none"> – обучающийся усвоил только основной программный материал, по существу излагает его, опираясь на знания только основной литературы; – допускает несущественные ошибки и неточности; – испытывает затруднения в практическом применении знаний направления; – слабо аргументирует научные положения; – затрудняется в формулировании выводов и обобщений; – частично владеет системой специализированных понятий. |

| Оценка компетенции | Характеристика сформированных компетенций |
|---------------------------------------|---|
| 5-балльная шкала | |
| «неудовлетворительно» «не зачтено» | <ul style="list-style-type: none"> – обучающийся не усвоил значительной части программного материала; – допускает существенные ошибки и неточности при рассмотрении проблем в конкретном направлении; – испытывает трудности в практическом применении знаний; – не может аргументировать научные положения; – не формулирует выводов и обобщений. |

10.3. Типовые контрольные задания или иные материалы.

Вопросы (задачи) для экзамена представлены в таблице 15.

Таблица 15 – Вопросы (задачи) для экзамена

| № п/п | Перечень вопросов (задач) для экзамена | Код индикатора |
|-------|--|----------------|
| | Учебным планом не предусмотрено | |

Вопросы (задачи) для зачета / дифф. зачета представлены в таблице 16.

Таблица 16 – Вопросы (задачи) для зачета / дифф. зачета

| № п/п | Перечень вопросов (задач) для зачета / дифф. зачета | Код индикатора |
|-------|---|--|
| 1 | Этапы анализа текста в NLP | УК-1.3.1 ПК-3.3.1 ПК-3.У.1 ПК-3.В.1 |
| 2 | Морфологический анализ и токенизация | ПК-3.3.1 ПК-3.У.1 ПК-3.В.1 |
| 3 | Словарный морфологический анализ | ПК-3.3.1 ПК-3.У.1 ПК-3.В.1 |
| 4 | Постморфологический анализ | ПК-3.3.1 ПК-3.У.1 ПК-3.В.1 ПК-7.У.1 ПК-7.В.1 |
| 5 | Синтаксическая сегментация и синтаксический анализ | ПК-3.3.1 ПК-3.У.1 ПК-3.В.1 ПК-7.У.1 ПК-7.В.1 |
| 6 | Автоматическое снятие омонимии | ПК-3.3.1 ПК-3.У.1 ПК-3.В.1 ПК-7.У.1 ПК-7.В.1 |
| 7 | Способы представления лингвистических данных | ПК-3.3.1 ПК-3.У.1 ПК-3.В.1 ПК-7.У.1 ПК-7.В.1 |
| 8 | Лингвистическая разметка | УК-1.3.1 |

| | | |
|----|---|--|
| | | ПК-3.3.1 ПК-3.У.1 ПК-3.В.1 |
| 9 | Лингвистические аннотации | УК-1.3.1 ПК-3.3.1 ПК-3.У.1 ПК-3.В.1 |
| 10 | Анализ корпуса | УК-1.3.1 ПК-3.3.1 ПК-3.У.1 ПК-3.В.1 ПК-7.У.1 ПК-7.В.1 |
| 11 | Статистические и математические методы, используемые при обработке текста | УК-1.3.1 ПК-7.У.1 ПК-7.В.1 |

Перечень тем для курсового проектирования/выполнения курсовой работы представлены в таблице 17.

Таблица 17 – Перечень тем для курсового проектирования/выполнения курсовой работы

| № п/п | Примерный перечень тем для курсового проектирования/выполнения курсовой работы |
|-------|--|
| | Учебным планом не предусмотрено |

Вопросы для проведения промежуточной аттестации в виде тестирования представлены в таблице 18.

– Примерный перечень вопросов для тестов

| № п/п | Примерный перечень вопросов для тестов | Код индикатора |
|-------|--|----------------|
| 1 | Выберите определение поиска информации. 1) процесс сбора данных из различных источников; 2) процесс принятия решений на основе информации. | УК-1.3.1 |
| 2 | Выберите правильный ответ К методам поиска информации относятся 1) библиографический поиск, интернет-поиск, опрос; 2) анализ, синтез, сравнение; 3) классификация, систематизация, обобщение. | УК-1.3.1 |
| 3 | Выберите правильный ответ Под критическим анализом информации подразумевается 1) процесс выбора наилучшего варианта из нескольких; 2) процесс проверки и оценки достоверности информации. | УК-1.3.1 |
| 4 | Выберите правильный ответ К методам критического анализа информации относятся 1) проверка на достоверность, оценка значимости, сравнение; 2) анализ, синтез, сравнение; 3) классификация, систематизация, обобщение. | УК-1.3.1 |
| 5 | Выберите правильный ответ | УК-1.3.1 |

| | | |
|----|---|----------|
| | <p>В процессе синтеза информации производится:</p> <ol style="list-style-type: none"> 1) проверка достоверности информации; 2) выбор наилучшего варианта из нескольких; 3) объединение разрозненных данных в единое целое. | |
| 6 | <p>Выберите правильный ответ</p> <p>В процессе синтеза информации применяются методы</p> <ol style="list-style-type: none"> 1) анализ, синтез, сравнение; 2) объединение, структурирование, интерпретация; 3) классификация, систематизация, обобщение. | УК-1.3.1 |
| 7 | <p>Выберите правильный ответ</p> <p>Системный подход основан на</p> <ol style="list-style-type: none"> 1) разработке системы для анализа объекта; 2) рассмотрении объекта как системы. | УК-1.3.1 |
| 8 | <p>Выберите правильный ответ</p> <p>Принцип целостности в системном подходе основан на:</p> <ol style="list-style-type: none"> 1) выделении объекта изучения в качестве целостного образования; 2) объединении нескольких отдельных объектов в систему. | УК-1.3.1 |
| 9 | <p>Выберите правильный ответ</p> <p>Преимущество системного подхода состоит в том, что он</p> <ol style="list-style-type: none"> 1) позволяет увидеть всю картину целиком; 2) позволяет избежать ошибок при решении задач; 3) позволяет быстро найти решение задачи. | УК-1.3.1 |
| 10 | <p>Выберите правильный ответ</p> <p>Недостаток системного подхода состоит в том, что он</p> <ol style="list-style-type: none"> 1) требует много времени для решения задачи; 2) может привести к ошибкам при решении задачи; 3) может привести к неправильному решению задачи. | УК-1.3.1 |
| 11 | <p>Выберите правильный ответ</p> <p>Системный подход основан на методах</p> <ol style="list-style-type: none"> 1) анализа, синтеза, моделирования; 2) наблюдения, интервью, эксперимента. | УК-1.3.1 |
| 12 | <p>Выберите правильный ответ</p> <p>Библиографический поиск предполагает</p> <ol style="list-style-type: none"> 1) использование поисковых систем; 2) использование каталогов библиотек, баз данных и других источников для нахождения необходимой литературы; 3) работу со списком источников в монографии. | УК-1.3.1 |
| 13 | <p>Выберите правильный ответ</p> <p>Интервью представляет собой</p> <ol style="list-style-type: none"> 1) беседу с преподавателем по теме исследования; 2) метод получения информации от экспертов или людей, обладающих знаниями в определенной области. | УК-1.3.1 |
| 14 | <p>Выберите правильный ответ</p> <p>Для сохранения информации используются форматы файлов</p> <ol style="list-style-type: none"> 1) .txt, .docx, .pdf; | УК-1.3.1 |

| | | |
|----|--|----------|
| | 1) .exe, .dll, .bat. | |
| 15 | Определите последовательность этапов в системном подходе 1) постановка задачи, 2) сбор информации, 3) синтез, 4) анализ. | УК-1.3.1 |
| 16 | Охарактеризуйте способы и средства хранения информации в доцифровую эпоху | УК-1.3.1 |
| 17 | Перечислите формы, в которых информация сохраняется в цифровой форме | УК-1.3.1 |
| 18 | Назовите несколько средств хранения информации в цифровой форме | УК-1.3.1 |
| 19 | Опишите процедуру поиска с использованием поисковой системы | УК-1.3.1 |
| 20 | Охарактеризуйте основные особенности чат-бота как средства получения информации | УК-1.3.1 |
| 21 | Выберите правильный ответ Лингвистическая экспертиза связана с 1) исследованием языковой деятельности человека; 2) исследованием письменных текстов; 3) исследованием визуальных материалов. | ПК-3.3.1 |
| 22 | Выберите правильный ответ Лингвистическая экспертиза определяет 1) авторство текста; 2) определяет правила написания текста; 3) определяет значение слов в тексте. | ПК-3.3.1 |
| 23 | Выберите правильный ответ Объектом лингвистической экспертизы являются 1) книги, статьи, научные работы; 2) рекламные материалы, объявления; 3) все вышеперечисленные. | ПК-3.3.1 |
| 24 | Выберите правильный ответ Лингвистическая экспертиза осуществляется на 1) стилистическом уровне; 2) лексическом уровне; 3) морфологическом уровне; 4) всех вышеперечисленных. | ПК-3.3.1 |
| 25 | Выберите правильный ответ Контекстуальный метод лингвистической экспертизы предполагает анализ 1) речевого контекста; 2) синтаксических явлений в тексте; 3) фонетических явлений в речи. | ПК-3.3.1 |
| 26 | Выберите правильный ответ Синтаксический метод лингвистической экспертизы предполагает анализ 1) речевого контекста; 2) синтаксических явлений в тексте; 3) фонетических явлений в речи. | ПК-3.3.1 |
| 27 | Выберите определение | ПК-3.У.1 |

| | | |
|----|---|----------|
| | Под кодифицированными нормами русского языка понимаются 1) правила орфографии, орфоэпии, морфологии и синтаксиса; 2) правила создания речевого сообщения. | |
| 28 | Выберите правильный ответ Аспектами кодификации являются 1) орфография, 2) орфоэпия; 3) пунктуация, 4) все вышеперечисленные. | ПК-3.У.1 |
| 29 | Выберите правильный ответ Под орфографическими нормами русского языка понимаются правила 1) произношения слов; 2) правила написания слов; 3) расстановки знаков препинания; 4) построения предложений и текстов. | ПК-3.У.1 |
| 30 | Выберите правильный ответ Под морфологическими нормами русского языка понимаются правила 1) изменения формы слова; 2) произношения; 3) построения предложений; 4) употребления слов и выражений. | ПК-3.У.1 |
| 31 | Выберите правильный ответ Под орфоэпическими нормами русского языка понимаются правила 1) произношения слов; 2) правила написания слов; 3) расстановки знаков препинания; 4) построения предложений и текстов. | ПК-3.У.1 |
| 32 | Выберите правильный ответ Под лексическими нормами русского языка понимаются правила 1) изменения формы слова; 2) произношения; 3) построения предложений; 4) употребления слов и выражений. | ПК-3.У.1 |
| 33 | Выберите верное утверждение Грамматическое значение 1) зависит от синтаксического контекста; 2) не зависит от синтаксического контекста. | ПК-3.У.1 |
| 34 | Выберите верное утверждение Лексическое значение 1) зависит от речевого контекста; 2) не зависит от контекста и фиксируется в толковых словарях; 3) может быть как зависимым, так и независимым от контекста. | ПК-3.У.1 |
| 35 | Выберите верное утверждение Значение многозначных слов 1) зависит от речевого контекста; 2) не зависит от речевого контекста. | ПК-3.У.1 |
| 36 | Выберите и запишите единицы лексического уровня, позволяющие отнести приведенный фрагмент к просторечию | ПК-3.В.1 |

| | | |
|----|---|----------|
| | «Эй, Галка, что за фигня у тебя происходит?!» | |
| 37 | Выберите и запишите единицы лексического уровня, позволяющие отнести приведенный фрагмент к научному стилю «There is no convincing empirical evidence that handwriting is cognitively superior to keyboarding, although some make that claim.» | ПК-3.В.1 |
| 38 | Определите, для какого стиля наиболее характерно наличие многочисленных развернутых конструкций с неличными формами глагола | ПК-3.В.1 |
| 39 | На основе анализа синтаксической структуры определите стилистическую принадлежность следующего фрагмента «Освоение дисциплины предполагает - владение методами проведения лингвистических экспертиз; - владение основными математико-статистическими методами обработки лингвистической информации с учетом элементов программирования и автоматической обработки лингвистических корпусов.» | ПК-3.В.1 |
| 40 | На основе анализа синтаксической структуры определите стилистическую принадлежность следующего фрагмента “Are you sure you’re ok?” “Absolutely.” “And you still wanna go there?” | ПК-3.В.1 |
| 41 | Выберите правильный ответ Способность математической и компьютерной лингвистики эффективно описывать и представлять языковые факты основана на 1) системности языка; 2) обобщении языковых явлений. | ПК-7.У.1 |
| 42 | Выберите правильный ответ Статистические методы исследования могут применяться 1) на фонетическом уровне; 2) на морфологическом уровне; 3) на лексическом уровне; 4) на всех вышеперечисленных. | ПК-7.У.1 |
| 43 | Выберите правильный ответ Результатом статистического изучения лексики являются 1) орфографические словари; 2) орфоэпические словари; 3) частотные словари | ПК-7.У.1 |
| 44 | Выберите верное утверждение 1) В современных системах обработки письменного текста используются методы статистического анализа и вероятностного распределения 2) В современных системах обработки письменного текста используется анализ значения слова, грамматической формы и синтаксической конструкции. | ПК-7.У.1 |
| 45 | Выберите правильный ответ Статистический анализ может использоваться 1) для изучения распределения слов в тексте | ПК-7.У.1 |

| | | |
|----|--|----------|
| | 2) выявления закономерностей в использовании определенных грамматических конструкций; 3) все вышеперечисленное. | |
| 46 | Выберите правильный ответ Математическое моделирование 1) используется для анализа языковых данных; 2) используется для создания формальных моделей языка | ПК-7.У.1 |
| 47 | Выберите правильный ответ Формальная модель языка позволяет 1) описать его структуру и функционирование; 2) описать отношения между словами в языке | ПК-7.У.1 |
| 48 | Выберите верное утверждение Различие морфологических систем языков 1) повышает сложность разработки системы автоматизированной обработки текста; 2) снижает сложность разработки системы автоматизированной обработки текста. | ПК-7.У.1 |
| 49 | Выберите верное утверждение Системы автоматизированной обработки текста 1) способны осуществлять прагматический анализ, т.е. определять намерение автора высказывания; 2) неспособны осуществлять прагматический анализ. | ПК-7.У.1 |
| 50 | Выберите верное утверждение Системы автоматизированной обработки текста 1) способны определять эмоциональную окраску текста; 2) неспособны определять эмоциональную окраску текста | ПК-7.У.1 |
| 51 | Выберите верное утверждение Системы автоматизированной обработки текста 1) способны осуществлять стилистический анализ; 2) неспособны осуществлять стилистический анализ. | ПК-7.У.1 |
| 52 | Выберите правильный ответ Первым этапом обработки текста является 1) написание текста; 2) форматирование текста; 3) распознавание текста | ПК-7.У.1 |
| 53 | Выберите правильный ответ Автоматические системы проверки орфографии 1) не относятся к системам автоматизированной обработки текста; 2) относятся к системам автоматизированной обработки текста | ПК-7.У.1 |
| 54 | Выберите правильный ответ Токенизация представляет 1) предварительную обработку текста; 2) обработку текста для публикации на сайте. | ПК-7.У.1 |
| 55 | Выберите правильный ответ В процессе токенизации в тексте могут вычленяться 1) символы; 2) слова; 3) предложения; 4) все вышеперечисленное. | ПК-7.У.1 |
| 56 | Дайте определение термину | ПК-7.В.1 |

| | | |
|---------------------------------------|---|-----------------------|
| | «лингвистическая разметка» | |
| 57 | Определите общую цель автоматической обработки корпусов языков | ПК-7.В.1 |
| 58 | Определите единицу обработки в современных системах автоматического перевода | ПК-7.В.1 |
| 59 | Определите, на основе каких данных система автоматического перевода дает перевод отдельного слова | ПК-7.В.1 |
| 60 | Приведите пример единиц, требующих специальных решений при разработке системы автоматизированной обработки текста | ПК-7.В.1 |
| Вопросы на проверку остаточных знаний | | |
| 1 | Выберите правильный ответ Современные системы автоматической обработки языка основаны на анализе 1) современной печатной продукции; 2) корпусов языков; 3) объема текстов в цифровом формате. | УК-1.3.1 |
| 2 | Язык программирования, используемый в разработке систем обработки и синтеза языка 1) Python; 2) JavaScript\$ 3) HTML | ПК-7.У.1 |
| 3 | Соотнесите единицу и уровень лингвистического анализа: 1) аллофон 2) окончание 3) несогласованное определение 4) вид глагола А) морфологический; Б) грамматический; В) синтаксический; Г) фонетический. | ПК-3.3.1 |
| 4 | Определите цель проведения сентимент-анализа | УК-1.3.1, ПК-7.У.1 |
| 5 | Расшифруйте аббревиатуру NLP | ПК-7.В.1 |

Перечень тем контрольных работ по дисциплине обучающихся заочной формы обучения, представлены в таблице 19.

Таблица 19 – Перечень контрольных работ

| № п/п | Перечень контрольных работ |
|-------|--|
| 1 | Контрольное задание: выполнение токенизации, анализ синтаксических конструкций |

10.4. Методические материалы, определяющие процедуры оценивания индикаторов, характеризующих этапы формирования компетенций, содержатся в локальных нормативных актах ГУАП, регламентирующих порядок и процедуру

проведения текущего контроля успеваемости и промежуточной аттестации обучающихся ГУАП.

11. Методические указания для обучающихся по освоению дисциплины

11.1. Методические указания для обучающихся по освоению лекционного материала *(если предусмотрено учебным планом по данной дисциплине)*.

Основное назначение лекционного материала – логически стройное, системное, глубокое и ясное изложение учебного материала. Назначение современной лекции в рамках дисциплины не в том, чтобы получить всю информацию по теме, а в освоении фундаментальных проблем дисциплины, методов научного познания, новейших достижений научной мысли. В учебном процессе лекция выполняет методологическую, организационную и информационную функции. Лекция раскрывает понятийный аппарат конкретной области знания, её проблемы, дает цельное представление о дисциплине, показывает взаимосвязь с другими дисциплинами.

Планируемые результаты при освоении обучающимися лекционного материала:

- получение современных, целостных, взаимосвязанных знаний, уровень которых определяется целевой установкой к каждой конкретной теме;
- получение опыта творческой работы совместно с преподавателем;
- развитие профессионально-деловых качеств, любви к предмету и самостоятельного творческого мышления.
- появление необходимого интереса, необходимого для самостоятельной работы;
- получение знаний о современном уровне развития науки и техники и о прогнозе их развития на ближайшие годы;
- научиться методически обрабатывать материал (выделять главные мысли и положения, приходить к конкретным выводам, повторять их в различных формулировках);
- получение точного понимания всех необходимых терминов и понятий.

Лекционный материал может сопровождаться демонстрацией слайдов и использованием раздаточного материала при проведении коротких дискуссий об особенностях применения отдельных тематик по дисциплине.

Структура предоставления лекционного материала:

- лекции;
- интернет-ресурсы с дополнительными материалами по предмету.

Конспект необходимо вести таким образом, чтобы оставались широкие поля для замечаний, вопросов или заметок. Целесообразно на одной странице разворота тетради писать текст, а другую оставлять свободной. На чистом поле можно по ходу лекции делать свои заметки, записывать сомнения, выводы. Все это поможет впоследствии вспомнить материал лекции.

Главную мысль лучше выделять красной строкой, цветным фломастером. Непременнo стоит записывать и расшифровывать новые понятия, термины. Не нужно стараться писать подробный текст, т.к. в этом случае легко упустить главное. При таком конспектировании плохо воспринимается смысл сказанного: все внимание концентрируется на записях.

Следует учиться из услышанного выделять главное, смысл, который лектор пытается донести, а конспектировать – основные положения и выводы. Запись лекции должна вестись в сжатой, лаконичной форме. Можно прибегнуть к сокращениям, использовать символы (больше, меньше и др., известные из математического школьного курса). Проанализировав, какие слова в профессиональной тематике встречаются чаще всего, составить свой словарь сокращений и использовать их при ведении конспекта.

11.2. Методические указания для обучающихся по участию в семинарах

Не предусмотрены.

11.3. Методические указания для обучающихся по прохождению практических занятий

Не предусмотрены.

11.4. Методические указания для обучающихся по выполнению лабораторных работ *(если предусмотрено учебным планом по данной дисциплине)*

В ходе выполнения лабораторных работ обучающийся должен углубить и закрепить знания, практические навыки, овладеть современной методикой и техникой эксперимента в соответствии с квалификационной характеристикой обучающегося. Выполнение лабораторных работ состоит из экспериментально-практической, расчетно-аналитической частей и контрольных мероприятий.

Выполнение лабораторных работ обучающимся является неотъемлемой частью изучения дисциплины, определяемой учебным планом, и относится к средствам, обеспечивающим решение следующих основных задач обучающегося:

- приобретение навыков исследования процессов, явлений и объектов, изучаемых в рамках данной дисциплины;
- закрепление, развитие и детализация теоретических знаний, полученных на лекциях;
- получение новой информации по изучаемой дисциплине;
- приобретение навыков самостоятельной работы с лабораторным оборудованием и приборами.

Задание и требования к проведению лабораторных работ

Лабораторная работы должна быть выполнена в полном объеме в установленные сроки, оформление отчета о лабораторной работе является обязательным. Лабораторная работа выполняется на основе полученного задания, ее выполнение предусматривает использование поисковых и экспертных лингвистических систем.

Структура и форма отчета о лабораторной работе

Теоретическая часть должна содержать минимум необходимых теоретических сведений о предметной области. Не следует копировать целиком или частично методическое пособие (описание) лабораторной работы или разделы учебника.

В разделе Программное обеспечение необходимо описать, с помощью каких инструментальных средств и каким образом были разработаны модели и получены результаты. Рисунки, блок-схемы, описание модели и её особенностей, необходимость отладки – все это должно быть представлено в указанном разделе.

Раздел Результаты включает в себя скриншоты программного приложения, полученные при выполнении лабораторной работы. Рисунки, графики и таблицы нумеруются и подписываются заголовками.

Выводы не должны быть простым перечислением того, что сделано. Здесь важно отметить, какие новые знания о предмете исследования были получены при выполнении работы, к чему привело обсуждение результатов, насколько выполнена заявленная цель работы.

Выводы по работе каждый студент делает самостоятельно. В случае необходимости в конце отчёта приводится Список литературы, использованной при подготовке к работе.

В тексте отчёта делаются краткие ссылки на литературу (учебники, справочники, иные источники...) номером в квадратных скобках, напр., [1]. Литературные источники нумеруются по мере их появления в тексте отчёта. В конце отчёта даётся их подробный

список. На все источники списка литературы должны быть ссылки в тексте отчёта, там, где это необходимо.

Требования к оформлению отчета о лабораторной работе

Объём отчёта должен быть оптимальным для понимания того, что и как сделал студент, выполняя работу. Обязательные требования к отчёту включают общую и специальную грамотность изложения, а также использование необходимых форм и аккуратность оформления. Отчет о лабораторной работе необходимо прикрепить в ЛК.

Форма отчета: <https://guap.ru/regdocs/docs/uch> (Для учебного процесса – Документация – Нормативная документация – ГУАП)

11.5. Методические указания для обучающихся по прохождению курсового проектирования/выполнения курсовой работы

Не предусмотрено

11.6. Методические указания для обучающихся по прохождению самостоятельной работы

В ходе выполнения самостоятельной работы, обучающийся выполняет работу по заданию и при методическом руководстве преподавателя, но без его непосредственного участия.

Для обучающихся по заочной форме обучения, самостоятельная работа может включать в себя контрольную работу.

В процессе выполнения самостоятельной работы, у обучающегося формируется целесообразное планирование рабочего времени, которое позволяет им развивать умения и навыки в усвоении и систематизации приобретаемых знаний, обеспечивает высокий уровень успеваемости в период обучения, помогает получить навыки повышения профессионального уровня.

Методическими материалами, направляющими самостоятельную работу обучающихся являются:

- учебно-методический материал по дисциплине;
- методические указания по выполнению контрольных работ (для обучающихся по заочной форме обучения).

Начинать надо с изучения рекомендованной литературы. Необходимо помнить, что на лекции обычно рассматривается не весь материал, а только его часть. Остальная его часть восполняется в процессе самостоятельной работы. В связи с этим работа с рекомендованной литературой обязательна. Особое внимание при этом необходимо обратить на содержание основных положений и выводов, объяснение явлений и фактов, уяснение практического приложения рассматриваемых теоретических вопросов. В процессе этой работы студент должен стремиться понять и запомнить основные положения рассматриваемого материала, примеры, поясняющие его, а также разобраться в иллюстративном материале.

Заканчивать подготовку следует составлением плана (конспекта) по изучаемому материалу (вопросу). Это позволяет составить концентрированное, сжатое представление по изучаемым вопросам.

Записи имеют первостепенное значение для самостоятельной работы студентов. Они помогают понять построение изучаемого материала, выделить основные положения, проследить их логику и тем самым проникнуть в творческую лабораторию автора.

Ведение записей способствует превращению чтения в активный процесс, мобилизует, наряду со зрительной, и моторную память. Следует помнить: у студента, систематически ведущего записи, создается свой индивидуальный фонд подсобных материалов для быстрого повторения прочитанного, для мобилизации накопленных

знаний. Особенно важны и полезны записи тогда, когда в них находят отражение мысли, возникшие при самостоятельной работе.

11.7. Методические указания для обучающихся по прохождению текущего контроля успеваемости.

Текущий контроль успеваемости предусматривает контроль качества знаний обучающихся, осуществляемого в течение семестра с целью оценивания хода освоения дисциплины.

Текущий контроль успеваемости представляет собой выполнение контрольной работы.

В ходе работы необходимо провести морфологический анализ заданного текста; для всех слов указать их лексико-грамматический класс, выделить омонимы, произвести токенизацию, прокомментировать отличие токенов от морфем, определить синтаксическую функцию единицы и тип синтаксической конструкции, а также составить словарь словоформ для пяти существительных и трех глаголов заданного текста.

11.8. Методические указания для обучающихся по прохождению промежуточной аттестации.

Промежуточная аттестация обучающихся предусматривает оценивание промежуточных и окончательных результатов обучения по дисциплине. Она включает в себя:

- экзамен – форма оценки знаний, полученных обучающимся в процессе изучения всей дисциплины или ее части, навыков самостоятельной работы, способности применять их для решения практических задач. Экзамен, как правило, проводится в период экзаменационной сессии и завершается аттестационной оценкой «отлично», «хорошо», «удовлетворительно», «неудовлетворительно».
- зачет – это форма оценки знаний, полученных обучающимся в ходе изучения учебной дисциплины в целом или промежуточная (по окончании семестра) оценка знаний обучающимся по отдельным разделам дисциплины с аттестационной оценкой «зачтено» или «не зачтено».
- дифференцированный зачет – это форма оценки знаний, полученных обучающимся при изучении дисциплины, при выполнении курсовых проектов, курсовых работ, научно-исследовательских работ и прохождении практик с аттестационной оценкой «отлично», «хорошо», «удовлетворительно», «неудовлетворительно».
- Лекции предназначены для изложения теоретического материала. Лабораторные работы по курсу предполагают закрепление уже полученного в ходе изучения курса студентами теоретического материала, выработку практических навыков.
- Для допуска к промежуточной аттестации необходимо выполнение студентами контрольной и лабораторных работ, в ходе выполнения которых студент демонстрирует полученные навыки и умения.
- На промежуточную аттестацию по курсу выносятся теоретический вопрос. Использование каких-либо материалов при ответе на теоретический вопрос не допускается. В ходе ответа оценка ставится исходя из знаний, продемонстрированных студентом при ответе на предложенный вопрос:
- «зачтено» - студент владеет материалом, способен ответить на 2-3 уточняющих вопроса;
- «незачтено» - студент демонстрирует крайне ограниченное или фрагментарное владение материалом, неспособен ответить на уточняющие вопросы.

Лист внесения изменений в рабочую программу дисциплины

| Дата внесения изменений и дополнений. Подпись внесшего изменения | Содержание изменений и дополнений | Дата и № протокола заседания кафедры | Подпись зав. кафедрой |
|---|-----------------------------------|--------------------------------------|-----------------------|
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |