**МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ**



**ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ "ОРЛОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМЕНИ И.С.ТУРГЕНЕВА" Физико-математический факультет**

**Кафедра информатики**

**Ломакин Д.Е.**

09.03.03 – 17 – **о** – 4

**Алгебра**

**Рабочая программа дисциплины**

**Тип образовательной программы**: **бакалавриат**

**Форма обучения**: **очная**

**Направление подготовки**: **09.03.03 Прикладная информатика**

**Направленность** (**профиль**): **Интеллектуальная обработка данных**

**Орел** 2020

**Автор** \_\_\_\_\_\_\_\_\_\_

**Рецензент** \_\_\_\_\_\_\_\_\_\_



**Рабочая программа разработана в соответствии с Федеральным государственным образовательным стандартом высшего образования**,

**утвержденным приказом Министерства образования и науки Российской Федерации от № по направлению подготовки 09.03.03 Прикладная информатика**

**Рабочая программа обсуждена на заседании Кафедра информатики**

**Протокол № от** «» **июня** 2020 **г**.

**Зав**. **кафедрой ФИО**\_\_\_\_\_\_\_\_\_\_

**Рабочая программа согласована с кафедрой информационных систем**, **за**

**которой закреплено направление подготовки**

**Зав**. **кафедрой**, **ФИО**\_\_\_\_\_\_\_\_\_\_

**Рабочая программа утверждена на заседании НМС Физико-математический факультет**

«» **Протокол № от** «» **июня** 2020 **г**.

**Председатель НМС ФИО** \_\_\_\_\_\_\_\_\_\_

|  |  |  |
| --- | --- | --- |
|  |  | 3 |
|  | **Содержание** |  |
| 1 | **Цели и задачи освоения дисциплины** | 4 |
| 2 | **Место дисциплины в структуре ОП** | 5 |
| 3 | **Планируемые результаты обучения по дисциплине** | 6 |
| 4 | **Структура дисциплины и распределение её трудоёмкости** | 7 |
| 5 | **Содержание дисциплины** | 8 |
| 6 | **Методические указания для обучающихся по освоению дисциплины** | 13 |
| 7 | **Фонд оценочных средств для проведения промежуточной аттестации** |  |
| **обучающихся по дисциплине** | | 15 |
| 8 | **Перечень основной и дополнительной учебной литературы**, |  |
| **необходимой для освоения дисциплины** | | 16 |
| 8.1 **Основная литература** | | 16 |
| 8.2 **Дополнительная литература** | | 16 |

1. **Перечень ресурсов информационно**-**телекоммуникационной сети**

|  |  |
| --- | --- |
| «**Интернет**», **необходимых для освоения дисциплины** | 17 |

1. **Перечень информационных технологий**, **используемых при**

|  |  |
| --- | --- |
| **осуществлении образовательного процесса по дисциплине** | 17 |
| 11 **Материально**-**техническое обеспечение дисциплины** | 17 |

4

**1 Цели и задачи освоения дисциплины**

**{}}**

**1111111111**

**Задачами изучения дисциплины** "**Алгебра**" **являются**:

**222222222**

5

**2 Место дисциплины в структуре ОП**

**ууууууууууууууууууууууу**

6

**3 Планируемые результаты обучения по дисциплине**

**Таблица** 1 - **Планируемые результаты обучения по дисциплине**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | ***Планируемые результаты обучения по*** | |  |
|  | ***Формируемые компетенции*** |  | ***дисциплине*** |  |
|  | **Требования к формируемым знаниям**, **умениям** | |  |
|  |  |  |
|  |  |  | **и навыкам** |  |
| **УК-1** | **способность использовать основы философских знаний для формирования мировоззренческой позиции** | **Знать** | основные направления, проблемы, теории и методы философии, содержание современных философских дискуссий по проблемам общественного развития. |  |
|  |  |  |  |  |
|  |  | **Владеть** | использовать положения и категории философии для оценивания и анализа различных социальных тенденций, фактов и явлений. |  |
|  |  |  |  |  |
|  |  | **Уметь** | Владеть навыками анализа текстов, имеющих философское содержание. |  |
|  |  |  |  |  |
| **УК-3** | **способность использовать основы экономических знаний в различных сферах жизнедеятельности** | **Знать** | Знать: знать базовые экономические понятия (спрос, предложение, цена, стоимость, товар, деньги, доходы, расходы, прибыль, риск, собственность, управление, рынок, фирма, государство), объективные основы функционирования экономики и поведения экономических агентов (законы спроса и предложения, принципы ценообразования, принцип ограниченной рациональности, принцип альтернативных издержек, принцип изменения ценности денег во времени). знать основные виды финансовых институтов (банк, страховая организация, брокер, биржа, негосударственный пенсионный фонд, центральный банк, агентство по страхованию вкладов, микрофинансовая организация, кредитный потребительский. кооператив, ломбард) и финансовых инструментов (банковский вклад, кредит, договор страхования, акция, облигация, пластиковая карта, индивидуальный инвестиционный счет), основы функционирования финансовых рынков. знать сущность и составные части издержек производства,источники и способыоптимизации издержек и прибыли фирм. знать основы ценообразования на рынках товаров и услуг. знать условия функционирования национальной экономики, понятие и факторы экономического роста. знать состав, структуру и способы расчета основных показателей результатов национального производства (валовой внутренний продукт, валовой национальный продукт, национальный доход, личный доход). знать значение государственной экономической политики в повышении эффективности экономики и роста благосостояния граждан, формы ее осуществления (денежно-кредитная, бюджетно-налоговая, социальная), основные методы и инструменты ее осуществления. знать основы российской налоговой системы. знать основы управления рисками (основные виды рисков, методы идентификации, измерения и оценки рисков; методы управления рисками: уклонение от риска, превентивные меры, контроль риска и финансирование риска, разделение, страховая и не страховая передача рисков) |  |
|  |  |  |  |  |
|  |  | **Владеть** | использовать положения и категории философии для оценивания и анализа различных социальных тенденций, фактов и явлений. |  |
|  |  |  |  |  |
|  |  | **Уметь** | Владеть навыками анализа текстов, имеющих философское содержание. |  |
|  |  |  |  |  |

7

**4 Структура дисциплины и распределение её трудоёмкости**

**Таблица** 2 - **Структура дисциплины и распределение ее трудоемкости**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Вид учебной работы** | **Всего**, **кол**. | | | **За** 8 **семестр**, | | |  |
| **кол**. | | |  |
|  |  |  |  |
|  |  |  |  |  |  |  |  |
|  | **часов** |  | **занятий** | **часов** |  | **занятий** |  |
| 1 | 2 |  | 3 | 4 |  | 5 |  |
| **1 Контактная работа, всего** | **72** |  | **28** | **72** |  | **28** |  |
| **Лекции** (**лек**) | 24 |  | 8 | 24 |  | 12 |  |
| **Лабораторные занятия** (**лаб**) | 32 |  | 8 | 32 |  | 8 |  |
| **Практические занятия** (**пр**) | 16 |  | 8 | 16 |  | 8 |  |
| **2 Самостоятельная работа (всего)** | **72** |  |  | **72** |  |  |  |
| **в том числе** |  |  |  |  |  |
|  |  |  |  |  |  |  |
| **Прочие виды самостоятельной работы** | 72 |  |  | 72 |  |  |  |
| **3 Промежуточная аттестация (форма)** |  | **36** | | **Экзамен** | | **(36)** |  |
|  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  | | |  | | |  |
| **Общая трудоемкость дисциплины в часах:** | **180** | | | **180** | | |  |
|  |  |  |  |  |  |  |
|  |  |  | |  |  | |  |
| **Общая трудоемкость дисциплины в зачетных единицах:** |  | **5** | |  | **5** | |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |

8

**5 Содержание дисциплины**

**Таблица** 3 – **Технологическая карта учебной дисциплины**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **занятия** | **Тема занятия** | **Контактная** | **Самостоятельна** | **Всего, час.** |  |
| **№** | **работа, час.** | **я работа, час.** |  |
|  |  |  |
| **Вид и** |  |  |  |  |  |
|  |  |  |  |  |  |
| 1 | 2 | 3 | 4 | 5 |  |
| **Семестр №8** |  |  |  |  |  |
| **Раздел №1 «Теоретические основы компьютерной лингвистики»** | |  |  |  |  |
|  | **Лекция**: **Введение в компьютерную лингвистику**. |  |  |  |  |
|  | **Изучаемые вопросы**: |  |  |  |  |
|  | **Основные понятия и определения**. **Цель и задачи компьютерной** |  |  |  |  |
|  | **лингвистики**. **Направления компьютерной лингвистики**. |  |  |  |  |
| **лек №**1 | **Компьютерный анализ и обработка неструктурированного текста**. | 2 | 4 | 6 |  |
|  | **Инструментарий компьютерной лингвистики**. |  |  |  |  |
|  | **Вопросы для самостоятельного изучения**: |  |  |  |  |
|  | **Лингвистические технологии и ресурсы**. **Направления исследований** |  |  |  |  |
|  | **и разработок компьютерной лингвистики**. |  |  |  |  |
|  | **Лекция**: **Основы лингвистики и теории речевой коммуникации**: **от** |  |  |  |  |
|  | **слова к тексту**. |  |  |  |  |
|  | **Изучаемые вопросы**: |  |  |  |  |
| **лек №**2 | **Языки**. **Тексты**. **Лингвистика**: **цели**, **методы**, **задачи**. **Принцип** | 2 | 4 | 6 |  |
|  | **моделирования**. **Инвентарные и конструктивные единицы**. **Понятие** |  |  |  |  |
|  | `**текущего словаря**`. **Избыточность и контекстная предсказуемость**. |  |  |  |  |
|  | **Вопросы для самостоятельного изучения**: |  |  |  |  |

9

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Основы речевой коммуникации**. **Коллокации и конструкции**. **Типы** |  |  |  |  |
|  | **коллокаций и конструкций**. **Принцип шкалирования**. |  |  |  |  |
|  | **Лекция**: **Семантическая и информационная структуры при анализе** |  |  |  |  |
|  | **текстов и коллекций**. |  |  |  |  |
|  | **Изучаемые вопросы**: |  |  |  |  |
| **лек №**3 | **Анализ текста в парадигмах автоматического понимания текста**. | 2 | 2 | 4 |  |
| **Коммуникативная и информационная структуры текста**. |  |
|  |  |  |  |  |
|  | **Избыточность**, **компрессия текста**, **свертки текста**. |  |  |  |  |
|  | **Вопросы для самостоятельного изучения**: |  |  |  |  |
|  | **Анализ текста в парадигме когнитивных исследований**. |  |  |  |  |
|  | **Лекция**: **Объекты исследования современной лингвистики текста**. |  |  |  |  |
|  | **Информационный поток**. |  |  |  |  |
|  | **Изучаемые вопросы**: |  |  |  |  |
|  | **Основы анализа информационного пространства и** |  |  |  |  |
| **лек №**4 | **информационных потоков**. **Информационный поток как объект** | 2 | 4 | 6 |  |
|  | **исследования**. **Моделирование информационных потоков**. **Модель** |  |  |  |  |
|  | **диффузии информации**. |  |  |  |  |
|  | **Вопросы для самостоятельного изучения**: |  |  |  |  |
|  | **Тематические информационные потоки**. |  |  |  |  |
| **Итого по разделу:** |  | *8* | *14* | *22* |  |
| **Раздел №2 «Автоматическая обработка и анализ текстов на естественном языке»** | |  |  |  |  |
|  | **Лекция**: **Компьютерная морфология**. |  |  |  |  |
|  | **Изучаемые вопросы**: |  |  |  |  |
|  | **Словарный и предиктивный морфологический анализ и синтез**. |  |  |  |  |
| **лек №**5 | **Лексическая неоднозначность**. **Автоматизированное пополнение** | 2 | 2 | 4 |  |
| **морфологического словаря**. **Методы бессловарного** |  |
|  |  |  |  |  |
|  | **морфологического анализа**. |  |  |  |  |
|  | **Вопросы для самостоятельного изучения**: |  |  |  |  |
|  | **Инструментарий для морфологического анализа и синтеза**. |  |  |  |  |
|  | **Лекция**: **Языковая модель**. **Исправление орфографических ошибок** |  |  |  |  |
| **лек №**6 | (**опечаток**). | 2 | 2 | 4 |  |
| **Изучаемые вопросы**: |  |
|  |  |  |  |  |
|  | **Языковая модель**, **цепь Маркова**, n-**граммы**. **Задача определения** |  |  |  |  |

10

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **части речи**. **Статистические методы определения части речи**. |  |  |  |  |
|  | **Частеречевая разметка на базе скрытых Марковских цепей и** |  |  |  |  |
|  | **алгоритм Витерби**. **Исправление опечаток**. **Расстояние** |  |  |  |  |
|  | **Левенштейна**, **расстояние Левенштейна**-**Дамерау**. **Подсчет** |  |  |  |  |
|  | **расстояний Левенштейна**. |  |  |  |  |
|  | **Вопросы для самостоятельного изучения**: |  |  |  |  |
|  | **Инструменты для исправления опечаток**. |  |  |  |  |
|  | **Лекция**: **Постморфологический и предсинтаксический анализ**. |  |  |  |  |
|  | **Изучаемые вопросы**: |  |  |  |  |
| **лек №**7 | **Автоматизированное снятие омонимии**. **Постморфологический** | 2 | 2 | 4 |  |
| **анализ**. **Синтаксическая сегментация**. |  |
|  |  |  |  |  |
|  | **Вопросы для самостоятельного изучения**: |  |  |  |  |
|  | **Морфологическая классификация естественных языков**. |  |  |  |  |
|  | **Лекция**: **Синтаксический анализ в естественном языке**. |  |  |  |  |
|  | **Изучаемые вопросы**: |  |  |  |  |
|  | **Синтаксическая неоднозначность**. **Подходы к описанию синтаксиса** |  |  |  |  |
|  | **в естественном языке**. **Методы и алгоритмы синтаксического** |  |  |  |  |
| **лек №**8 | **разбора в контексте грамматики зависимостей**. **Методы и** | 2 | 4 | 6 |  |
|  | **алгоритмы синтаксического разбора в контексте КС**-**грамматик**. |  |  |  |  |
|  | **Вопросы для самостоятельного изучения**: |  |  |  |  |
|  | **Иерархия Хомского**. **Задача синтаксического разбора**. **Оценка** |  |  |  |  |
|  | **точности синтаксического анализа**. |  |  |  |  |
|  | **Лекция**: **Семантический анализ**. |  |  |  |  |
|  | **Изучаемые вопросы**: |  |  |  |  |
|  | **Формальные методы семантического анализа**. **Понятие** |  |  |  |  |
| **лек №**9 | **онтологии**. **Модели представления знаний в компьютерной** | 2 | 2 | 4 |  |
| **семантике**. **Дистрибутивная семантика**. |  |
|  |  |  |  |  |
|  | **Вопросы для самостоятельного изучения**: |  |  |  |  |
|  | **Онтологические ресурсы и компьютерные тезаурусы**. |  |  |  |  |
| **лек №**10 | **Лекция**: **Алгоритмы Классификации полнотекстовых документов**. | 2 | 2 | 4 |  |
| **Изучаемые вопросы**: |  |
|  |  |  |  |  |

11

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Обзор терминов для классификации**. **Алгоритмы классификации с** |  |  |  |  |
|  | **учителем**. |  |  |  |  |
|  | **Алгоритм Роккио**. **Алгоритм** `**наивной**` **байесовской классификации**. |  |  |  |  |
|  | **Алгоритм деревьев принятия решений**. **Другие алгоритмы** |  |  |  |  |
|  | **классификации с обучением**. |  |  |  |  |
|  | **Вопросы для самостоятельного изучения**: |  |  |  |  |
|  | **Оценка результатов классификации**. |  |  |  |  |
|  | **Лекция**: **Кластеризация и контент**-**анализ**. |  |  |  |  |
|  | **Изучаемые вопросы**: |  |  |  |  |
| **лек №**11 | **Кластерный анализ**. **Алгоритмы кластеризации**. **Контент**-**анализ**. | 2 | 2 | 4 |  |
|  | **Вопросы для самостоятельного изучения**: |  |  |  |  |
|  | **Инструментальные средства кластеризации и контент**-**анализа**. |  |  |  |  |
|  | **Лекция**: **Инструментарий для разработки приложений по** |  |  |  |  |
|  | **автоматической обработке текстов на естественном языке**. |  |  |  |  |
|  | **Изучаемые вопросы**: |  |  |  |  |
|  | **Представление лингвистических данных**. **Архитектура** |  |  |  |  |
| **лек №**12 | **инструментальных систем для автоматической обработке текстов на** | 2 | 2 | 4 |  |
|  | **естественном языке**. **Системы автоматической обработки текстов на** |  |  |  |  |
|  | **естественном языке**. |  |  |  |  |
|  | **Вопросы для самостоятельного изучения**: |  |  |  |  |
|  | **Программные средства лингвистической обработки**. |  |  |  |  |
| **лаб №**1 | **Построение языковой модели** | 4 | 3 | 7 |  |
| **лаб №**2 | **Морфологический анализ неструктурированного текста** | 4 | 3 | 7 |  |
| **лаб №**3 | **Синтаксический анализ неструктурированного текста** | 4 | 3 | 7 |  |
| **лаб №**4 | **Семантический анализ неструктурированного текста** | 4 | 3 | 7 |  |
| **лаб №**5 | **Классификация полнотекстовых документов** | 4 | 3 | 7 |  |
| **лаб №**6 | **Кластеризация полнотекстовых документов** | 4 | 3 | 7 |  |
| **лаб №**7 | **Автоматизированное построение словаря**-**тезауруса**. **Разработка** | 4 | 3 | 7 |  |
| **лингвистической базы данных**. |  |
|  |  |  |  |  |
| **лаб №**8 | **Автоматизированное построение словаря**-**тезауруса**. **Создание** | 4 | 3 | 7 |  |
| **семантической сети и карт понятий**. |  |
|  |  |  |  |  |
| **пр №**1 | **Автоматическое аннотирование и реферирование текста**. | 2 | 2 | 4 |  |

12

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **пр №**2 | **Автоматический анализ текста** | 2 | 2 | 4 |
| **пр №**3 | **Автоматический синтез диалогов** | 2 | 2 | 4 |
| **пр №**4 | **Использование корпусов текстов** | 2 | 2 | 4 |
| **пр №**5 | **Работа с электронными словарями** | 2 | 2 | 4 |
| **пр №**6 | **Использование инструментария машинного перевода** | 2 | 2 | 4 |
| **пр №**7 | **Информационный поиск** | 2 | 2 | 4 |
| **пр №**8 | **Автоматическое распознавание текста** | 2 | 2 | 4 |
| **Итого по разделу:** |  | *64* | *58* | *122* |
| **Промежуточная аттестация**: **экз** | |  | *36* | *36* |
| **Итого по семестру:** | | *72* | *108* | *180* |
| **Итого по дисциплине:** | | *72* | *108* | *180* |
| **Примечания** |  |  |  |  |

13

**6 Методические указания для обучающихся по освоению дисциплины**

**Лекционный материал по дисциплине** "**Автоматическая обработка текстов**" **содержится в учебном пособии из перечня основной литературы**;

**при проведении аудиторных занятий излагается с применением демонстрационных слайдов**. **В ходе лекционных занятий студентам рекомендуется внимательно слушать преподавателя**, **уделять особое внимание определениям**, **формулам**, **таблицам**, **схемам**, **рисункам**. **Позднее в ходе самостоятельной работы студенты закрепляют прослушанный теоретический материал с помощью указанного учебного пособия**, **а также других рекомендованных литературных источников**.

***Методические рекомендации по подготовке к лабораторным работам и практическим занятиям***

**Лабораторные работы и практические занятия ориентированы на изучение и приобретение студентами практических навыков использования лингвистических ресурсов и технологий для автоматической обработки и анализа текстов**. **Для выполнения практических заданий студенты должны уметь работать с основными офисными приложениями**, **в частности**, **с текстовым процессором**, **использовать технологи сети Интернет для доступа к лингвистическим ресурсам**, **программы для распознавания текстов**.

**При подготовке к лабораторной работе и практическому занятию студенты должны повторить лекционный и дополнительный материал**, **относящийся к изучаемой теме**, **ознакомиться с теоретическими сведениями**.

**Подготовка к лабораторной работе и практическому занятию является самостоятельной работой студентов**. **Возникшие в ходе изучения теоретического материала вопросы при подготовке к лабораторной работе решаются совместно с преподавателем в начале аудиторного занятия**.

**Рекомендации по выполнению и объемё лабораторных и практических работ содержатся в учебно**-**методической литературе**, **представленной в конце раздела**.

***Методические рекомендации по организации*** ***самостоятельной***

***работы***

**Самостоятельная работа приводит студентов к получению нового знания**, **упорядочению и углублению имеющихся знаний**, **формированию у него профессиональных навыков и умений**.

**Самостоятельная работа выполняет ряд функций**:

1. **развивающую**;
2. **информационно**-**обучающую**;
3. **ориентирующую и стимулирующую**;
4. **воспитывающую**;
5. **исследовательскую**.

14

**Виды самостоятельной работы**, **выполняемые в ходе изучения дисциплины** "**Автоматическая обработка текстов**":

1. **Конспектирование первоисточников и другой учебной литературы**;
2. **Проработка учебного материала** (**по конспектам**, **учебной и научной литературе**);
3. **Работа с тестами и вопросами для самопроверки**;

**Студентам рекомендуется с самого начала освоения дисциплины работать с литературой и предлагаемыми заданиями в форме подготовки к очередному аудиторному занятию**. **При этом актуализируются имеющиеся знания**, **а также создается база для усвоения нового материала**, **возникают вопросы**, **ответы на которые обучающийся получает в аудитории**.

**При освоении дисциплины** "**Автоматическая обработка текстов**" **студент может пользоваться библиотекой университета**, **которая в полной мере обеспечена соответствующей литературой**, **а также свободными Интернет**-**ресурсами по данной тематике**. **Значительную помощь в подготовке к очередному занятию может оказать имеющееся в учебно**-**методическом комплексе учебное пособие по данной дисциплине**.

***Методические рекомендации по работе с литературой***

**Всю литературу можно разделить на учебники и учебные пособия**, **оригинальные научные монографические источники**, **научные публикации в периодической печати**. **Из них можно выделить литературу основную** (**рекомендуемую**), **дополнительную и литературу для углубленного изучения** **дисциплины**.

**Изучение дисциплины следует начинать с учебника**, **поскольку учебник** - **это книга**, **в которой изложены основы научных знаний по определенному предмету в соответствии с целями и задачами обучения**, **установленными программой**.

**Важной составляющей любого солидного научного издания является список литературы**, **на которую ссылается автор**. **При возникновении интереса к какой**-**то обсуждаемой в тексте проблеме всегда есть возможность обратиться к списку относящейся к ней литературы**. **В этом случае вся проблема как бы разбивается на составляющие части**, **каждая из которых может изучаться отдельно от других**. **При этом важно не терять из вида общий контекст и не погружаться чрезмерно в детали**, **потому что таким образом можно не увидеть главного**.

**Учебно**-**методическое обеспечение дисциплины** "**Автоматическая обработка текстов**":

1. **Теоретические сведения**: **Боярский К**.**К**. **Введение в компьютерную лингвистику** [**Электронный ресурс**] : **учебное пособие** / **К**.**К**. **Боярский**. — **Электрон**. **текстовые данные**. — **СПб**. : **Университет ИТМО**, 2013. — 73 c. —2227-8397. — **Режим доступа**: <http://www.iprbookshop.ru/71485.html;>

**Лабораторные работы и практические занятия**:

15

1.1. **Куликов Г**.**Г**., **Брейкин Т**.**В**., **Камалова Л**.**З** . **Интеллектуальные** **информационные системы**: **Методические указания к лабораторному практикуму по курсу** "**Интеллектуальные информационные**

**системы**"[**Электронный** **ресурс**].- **Режим** **доступа**:

<http://window.edu.ru/resource/941/23941;>

1.2. **Автоматическая обработка текстов на естественном языке и** **компьютерная лингвистика**: **учебное пособие** [**Электронный ресурс**] / **Е**.**И**. **Большакова**, **Э**.**С**. **Клышинский**, **Д**.**В**. **Ландэ**, **А**.**А**. **Носков**, **О**.**В**. **Пескова**, **Е**.**В**. **Ягунова** - **М**.: **МИЭМ**, 2011. - 272 **с**.- **Режим доступа**:<http://window.edu.ru/resource/465/78465>

**Перечисленное выше учебно**-**методическое обеспечение дисциплины** "**Автоматическая обработка текстов**" **представлено в электронных** **образовательных ресурсах на сайте библиотеки университета**.

**7 Фонд оценочных средств для проведения промежуточной аттестации обучающихся по дисциплине**

**Фонд оценочных средств представлен в Приложении к рабочей программе**.

16

**8 Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины**

**8.1 Основная литература**

1. **Боярский К**.**К**. **Введение в компьютерную лингвистику** [**Электронный ресурс**] : **учебное пособие** / **К**.**К**. **Боярский**. — **Электрон**. **текстовые данные**. — **СПб**. : **Университет ИТМО**, 2013. — 73 c. — 2227-8397. — **Режим доступа**: <http://www.iprbookshop.ru/71485.html>

**8.2 Дополнительная литература**

1. **Автоматическая обработка текстов на естественном языке и компьютерная лингвистика**: **учебное пособие** [**Электронный ресурс**] / **Е**.**И**. **Большакова**, **Э**.**С**. **Клышинский**, **Д**.**В**. **Ландэ**, **А**.**А**. **Носков**, **О**.**В**. **Пескова**, **Е**.**В**. **Ягунова** - **М**.: **МИЭМ**, 2011. - 272 **с**.- **Режим доступа**:<http://window.edu.ru/resource/465/78465>
2. **Шемякин Ю**.**И**. **Начала компьютерной лингвистики**: **Учебное пособие** [**Электронный ресурс**]. - **М**.: **Иэд**-**во МГОУ**, **А**/**О** `**Росвузнаука**`, 1992.81 **с**.- **Режим доступа**: <http://window.edu.ru/resource/098/24098>
3. **Лукашевич Н**.**В**. **Тезаурусы в задачах информационного поиска** [**Электронный ресурс**] : **монография** / **Н**.**В**. **Лукашевич**. - **Электрон**. **текстовые данные**.- **М**. : **Московский государственный университет имени М**.**В**. **Ломоносова**, 2011.-512 c.-978-5-211-05926-9.- **Режим доступа**:<http://www.iprbookshop.ru/13346.html>
4. **Методы когнитивного анализа семантики слова** [**Электронный ресурс**] : **компьютерно**-**корпусный подход** / **В**.**И**. **Заботкина** [**и др**.]. — **Электрон**. **текстовые данные**. — **М**. : **Языки славянской культуры**, 2015. —

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 342 | c. | — | 978-5-9445-7209-7. | | — | **Режим** | **доступа**: |
| <http://www.iprbookshop.ru/35673.html> | | | |  |  |  |  |

1. **Кухаренко Б**.**Г**. **Интеллектуальные системы и технологии** [**Электронный ресурс**]: **учебное пособие**/ **Кухаренко Б**.**Г**.— **Электрон**. **текстовые данные**.— **М**.: **Московская государственная академия водного**

**транспорта**, 2015.— 116 c.— **Режим доступа**:<http://www.iprbookshop.ru/47933.html>.— **ЭБС** «IPRbooks»

1. **Интеллектуальные системы** [**Электронный ресурс**]: **учебное пособие**/ **А**.**М**. **Семенов** [**и др**.].— **Электрон**. **текстовые данные**.— **Оренбург**: **Оренбургский государственный университет**, **ЭБС АСВ**, 2013.— 236 c.—

**Режим доступа**:<http://www.iprbookshop.ru/30055.html>.— **ЭБС** «IPRbooks»

1. **Сузи Р**.**А**. **Язык программирования** Python [**Электронный ресурс**]/ **Сузи Р**.**А**.— **Электрон**. **текстовые данные**.— **М**.: **Интернет**-**Университет Информационных Технологий** (**ИНТУИТ**), 2016.— 350 c.— **Режим доступа**:<http://www.iprbookshop.ru/52211.html>.— **ЭБС** «IPRbooks»

17

**9 Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины**

**Для успешного освоения дисциплины** "**Автоматическая обработка текстов**" **можно использовать материалы следующих Интернет**-**ресурсов**:

1. **Электронная библиотека образовательных ресурсов** (**ЭБОР**)[**Электронный ресурс**].- **Режим доступа**: <http://elib.oreluniver.ru>- **Систем**.

**требования**: P IV; 64 **Мб ОЗУ**; Windows 98 **и выше**; SVGA 32768 **и более цветов**; 640x480; **мышь**; IE 4.0 **и выше**.- **Загл**. **с экрана** - **Яз**. **рус**.

1. **Электронно**-**библиотечная система издательства** "**ЛАНЬ**"[**Электронный ресурс**].- **Режим доступа**: <http://www.e.lanbook>- **Систем**.

**требования**: P IV; 64 **Мб ОЗУ**; Windows 98 **и выше**; SVGA 32768 **и более цветов**; 640x480; **мышь**; IE 4.0 **и выше**.- **Загл**. **с экрана**

1. **Электронная библиотечная система** IPRbooks [**Электронный ресурс**].- **Режим доступа**:<http://www.IPRbooks>- **Систем**. **требования**: P IV; 64

**Мб ОЗУ**; Windows 98 **и выше**; SVGA 32768 **и более цветов**; 640x480; **мышь**;IE 4.0 **и выше**.- **Загл**. **с экрана**

1. **Федеральный портал** «**Российское образование**» [**Электронный ресурс**].- **Режим доступа**:<http://www.edu.ru>- **Систем**. **требования**: P IV; 64 **Мб ОЗУ**; Windows 98 **и выше**; SVGA 32768 **и более цветов**; 640x480; **мышь**; IE4.0 **и выше**.- **Загл**. **с экрана**

**10 Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине**

**Для выполнения лабораторных работ по дисциплине** "**Автоматическая обработка текстов**" **используется среда разработки на языке** Python, **специализированные библиотеки** (**или их аналоги**).

**Проведение лекционных занятий осуществляется с использованием презентаций и программы** PowerPoint **версии** 2007 **и выше** (**или бесплатный аналог**).

**11 Материально-техническое обеспечение дисциплины**

**Для изучения дисциплины** "**Автоматическая обработка текстов**" "

**необходимо следующее материально**-**техническое обеспечение**:

**для чтения лекций необходима лекционная аудитория**, **оснащенная мультимедийным проектором**, **с присоединенным к нему персональным компьютером для запуска презентаций**;

**для проведения лабораторных и практических работ необходим компьютерный класс**, **имеющей доступ в сеть Интернет**, **с установленным на компьютерах необходимым программным обеспечением**.

**ПРИЛОЖЕНИЕ**

**К РАБОЧЕЙ ПРОГРАММЕ**

**ДИСЦИПЛИНЫ**

**ФОНД ОЦЕНОЧНЫХ СРЕДСТВ**

**по дисциплине**

**«Автоматическая обработка текстов»**

**Направление подготовки** 09.03.03 **Прикладная информатика Направленность** (**профиль**): «**Интеллектуальная обработка данных**»

2017

**1 Перечень оценочных средств и их соответствие планируемым результатам обучения по дисциплине**

|  |  |  |
| --- | --- | --- |
| **Форма аттестации** | **Оценочные средства** | **Планируемые результаты обучения (индикаторы достижения компетенций)** |

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | **Знать**: |  |
|  |  | - **языки программирования и среды разработки для создания программных прототипов**, |  |
|  |  | **выполняющих интеллектуальную обработку и анализ данных**, **в частности**, **электронных** |  |
|  |  | **словарей**, **лингвистических баз данных**; **З** (**ПК**-8) |  |
|  |  | - **особенности обработки неструктурированных текстов на естественном языке**; **модели и** |  |
|  |  | **алгоритмы интеллектуальной обработки и анализа данных**; **методы автоматической об**- |  |
|  |  | **работки текстов с использованием лингвистических технологий**. **З** (**ДПК**-1) |  |
|  |  | **Уметь**: |  |
|  |  | - **создавать программные прототипы для интеллектуальной обработки и анализа данных**, |  |
| **Экзамен** | **Комплект экзаменацион**- | **в частности**, **электронные словари**, **лингвистические базы данных**; **У** (**ПК**-8) |  |
| **ных билетов** | - **использовать лингвистические технологии и ресурсы для автоматической обработки** |  |
|  |  |
|  |  | **неструктурированных текстов**; **модели и методы сбора**, **интеллектуальной обработки и** |  |
|  |  | **анализа данных**. **У** (**ДПК**-1) |  |
|  |  | **Владеть**: |  |
|  |  | - **практическими навыками высокоуровневого программирования для создания про**- |  |
|  |  | **граммных прототипов для интеллектуальной обработки и анализа данных**, **в частности**, |  |
|  |  | **электронных словарей**, **лингвистических баз данных**; **В** (**ПК**-8) |  |
|  |  | - **навыками создания программ с использованием специальных библиотек для анализа** |  |
|  |  | **текстов на естественном языке**, **навыками работы с инструментальными средствами сбо**- |  |
|  |  | **ра данных**. **В** (**ДПК**-1) |  |

**2 Критерии и шкалы оценивания**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Вид контроля** | **Форма атте-** | **Оценочные сред-** | **Критерии оценивания для промежуточной атте-** | **Шкала оценивания** |
|  | **стации** | **ства** | **стации** |  |
| **Промежуточная** | **Экзамен** | **Комплект экзаме**- | **Студент демонстрирует полное понимание про**- | 34 – 40 – «5» **отлично** |
| **аттестация** |  | **национных биле**- | **блемы**, **то есть**: |  |
|  |  | **тов** | - **четко знает языки программирования и среды** |  |
|  |  |  | **разработки для создания программных прототипов**, |  |
|  |  |  | **выполняющих интеллектуальную обработку и ана**- |  |
|  |  |  | **лиз данных**, **в частности**, **электронных словарей**, |  |
|  |  |  | **лингвистических баз данных**; |  |
|  |  |  | - **особенности обработки неструктурированных** |  |
|  |  |  | **текстов на естественном языке**; **модели и алгоритмы** |  |
|  |  |  | **интеллектуальной обработки и анализа данных**; **ме**- |  |
|  |  |  | **тоды автоматической обработки текстов с исполь**- |  |
|  |  |  | **зованием лингвистических технологий**; |  |
|  |  |  | - **умеет полностью создавать программные прото**- |  |
|  |  |  | **типы для интеллектуальной обработки и анализа** |  |
|  |  |  | **данных**, **в частности**, **электронные словари**, **лингви**- |  |
|  |  |  | **стические базы данных**; |  |
|  |  |  | - **использовать лингвистические технологии и ре**- |  |
|  |  |  | **сурсы для автоматической обработки неструктури**- |  |
|  |  |  | **рованных текстов**; **модели и методы сбора**, **интел**- |  |
|  |  |  | **лектуальной обработки и анализа данных**; |  |
|  |  |  | - **четко владеет навыками высокоуровневого про**- |  |
|  |  |  | **граммирования для создания программных прото**- |  |
|  |  |  | **типов для интеллектуальной обработки и анализа** |  |
|  |  |  | **данных**, **в частности**, **электронных словарей**, **линг**- |  |
|  |  |  | **вистических баз данных**; |  |
|  |  |  | - **навыками создания программ с использованием** |  |
|  |  |  | **специальных библиотек для анализа текстов на** |  |
|  |  |  | **естественном языке**, **навыками работы с инструмен**- |  |
|  |  |  | **тальными средствами сбора данных**. |  |

**Студент демонстрирует значительное понимание** 26 – 33 – «4» **хорошо проблемы**, **то есть**:

- **достаточно хорошо знает языки программирова**-

**ния и среды разработки для создания программных**

**прототипов**, **выполняющих интеллектуальную об**-

**работку и анализ данных**, **в частности**, **электронных**

**словарей**, **лингвистических баз данных**;

* **особенности обработки неструктурированных текстов на естественном языке**; **модели и алгоритмы интеллектуальной обработки и анализа данных**; **ме**-**тоды автоматической обработки текстов с исполь**-**зованием лингвистических технологий**;
* **достаточно хорошо умеет создавать программные прототипы для интеллектуальной обработки и ана**-**лиза данных**, **в частности**, **электронные словари**, **лингвистические базы данных**;
* **использовать лингвистические технологии и ре**-**сурсы для автоматической обработки неструктури**-**рованных текстов**; **модели и методы сбора**, **интел**-**лектуальной обработки и анализа данных**;
* **достаточно хорошо владеет практическими навыками высокоуровневого программирования для создания программных прототипов для интеллекту**-**альной обработки и анализа данных**, **в частности**, **электронных словарей**, **лингвистических баз дан**-**ных**;
* **навыками создания программ с использованием специальных библиотек для анализа текстов на естественном языке**, **навыками работы с инструмен**-**тальными средствами сбора данных**.

**Студент демонстрирует частичное понимание про**- 21 – 25 – «3» **удовл**.

**блемы**, **то есть**:

* **частично знает языки программирования и среды**

**разработки для создания программных прототипов**,

**выполняющих интеллектуальную обработку и ана**-

**лиз данных**, **в частности**, **электронных словарей**,

**лингвистических баз данных**;

* **особенности обработки неструктурированных текстов на естественном языке**; **модели и алгоритмы интеллектуальной обработки и анализа данных**; **ме**-**тоды автоматической обработки текстов с исполь**-**зованием лингвистических технологий**;
* **частично умеет создавать программные прототи**-**пы для интеллектуальной обработки и анализа дан**-**ных**, **в частности**, **электронные словари**, **лингвисти**-**ческие базы данных**;
* **использовать лингвистические технологии и ре**-**сурсы для автоматической обработки неструктури**-**рованных текстов**; **модели и методы сбора**, **интел**-**лектуальной обработки и анализа данных**;
* **частично владеет практическими навыками высо**-**коуровневого программирования для создания про**-**граммных прототипов для интеллектуальной обра**-**ботки и анализа данных**, **в частности**, **электронных словарей**, **лингвистических баз данных**; **В** (**ПК**-8)
* **навыками создания программ с использованием специальных библиотек для анализа текстов на естественном языке**, **навыками работы с инструмен**-

**тальными средствами сбора данных**

**Студент демонстрирует непонимание проблемы**, **то** 0 – 20 – «2» **неудовл**.

**есть**:

* **совершенно не знает языки программирования и среды разработки для создания программных про**-**тотипов**, **выполняющих интеллектуальную обра**-**ботку и анализ данных**, **в частности**, **электронных словарей**, **лингвистических баз данных**

* **особенности обработки неструктурированных текстов на естественном языке**; **модели и алгоритмы интеллектуальной обработки и анализа данных**; **ме**-**тоды автоматической обработки текстов с исполь**-**зованием лингвистических технологий**;
* **совершенно не умеет создавать программные прототипы для интеллектуальной обработки и ана**-**лиза данных**, **в частности**, **электронные словари**, **лингвистические базы данных**;
* **использовать лингвистические технологии и ре**-**сурсы для автоматической обработки неструктури**-**рованных текстов**; **модели и методы сбора**, **интел**-**лектуальной обработки и анализа данных**;
* **совершенно не владеет практическими навыками высокоуровневого программирования для создания программных прототипов для интеллектуальной обработки и анализа данных**, **в частности**, **элек**-**тронных словарей**, **лингвистических баз данных**;
* **навыками создания программ с использованием специальных библиотек для анализа текстов на естественном языке**, **навыками работы с инструмен**-**тальными средствами сбора данных**.

1. **Типовые оценочные средства**

**Промежуточная аттестация по дисциплине** – **экзамен в устной форме**.

**Время и место проведения экзамена устанавливается в соответствии с расписанием экзаменационной сессии**. **Продолжительность работы** – 1 **час** 30 **минут**.

**Экзаменационный билет состоит из двух частей**: **устное собеседование по двум вопросам и решение типовой задачи на компьютере**.

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **№** | **Структура экза**- | **Разделы**, **содер**- | **Проверяемые** | **Критерии оценки** | | | | | | |  | **Макс**. |
|  | **менационной ра**- | **жание дисци**- | **результаты** |  |  |  |  |  |  |  |  | **балл** |
|  | **боты** | **плины** | **обучения** |  |  |  |  |  |  |  |  |  |
| 1- | **Устное собеседова**- | **Теоретические ос**- | **З** (**ДПК**-1) | 0 **баллов** | | | **ставится**, | | | | **когда** | 11+11 |
| 2 | **ние по двум вопро**- | **новы компьютерной** | **У** (**ДПК**-1) | **студент** | |  | **демонстрирует** | | | | |  |
|  | **сам в билете** | **лингвистики**, **авто**- | **В** (**ДПК**-1) | **непонимание проблемы**, **то** | | | | | | | |  |
|  |  | **матическая обра**- |  | **есть**: **совершенно не знает** | | | | | | | |  |
|  |  | **ботка и анализ тек**- |  | **особенности** | | | |  | **обработки** | | |  |
|  |  | **стов на естествен**- |  | **неструктурированных** | | | | | | | **тек**- |  |
|  |  | **ном языке** |  | **стов на естественном язы**- | | | | | | | |  |
|  |  |  |  | **ке**; **не знает и не умеет ис**- | | | | | | | |  |
|  |  |  |  | **пользовать модели и алго**- | | | | | | | |  |
|  |  |  |  | **ритмы** | | **интеллектуальной** | | | | | |  |
|  |  |  |  | **обработки и анализа дан**- | | | | | | | |  |
|  |  |  |  | **ных**; **методы автоматиче**- | | | | | | | |  |
|  |  |  |  | **ской** | **обработки** | | | | **текстов с** | | |  |
|  |  |  |  | **использованием** | | | | |  | **лингви**- | |  |
|  |  |  |  | **стических** | | | **технологий**, **не** | | | | |  |
|  |  |  |  | **владеет** | | **навыками** | | | | **работы** | |  |
|  |  |  |  | **со специальными** | | | | | | **библио**- | |  |
|  |  |  |  | **теками** (**На** 50% **и более** | | | | | | | |  |
|  |  |  |  | **вопросов**, **связанных с ни**- | | | | | | | |  |
|  |  |  |  | **ми**, **нет ответа**); | | | | |  |  |  |  |
|  |  |  |  | 5 **баллов ставится**, **когда** | | | | | | | |  |
|  |  |  |  | **студент** | |  | **демонстрирует** | | | | |  |
|  |  |  |  | **частичное понимание про**- | | | | | | | |  |
|  |  |  |  | **блемы**, **то есть**: **частично** | | | | | | | |  |
|  |  |  |  | **знает** | **особенности** | | | | | | **обра**- |  |
|  |  |  |  | **ботки** |  | **неструктурирован**- | | | | | |  |
|  |  |  |  | **ных текстов на естествен**- | | | | | | | |  |
|  |  |  |  | **ном языке**; **частично знает** | | | | | | | |  |
|  |  |  |  | **и умеет использовать мо**- | | | | | | | |  |
|  |  |  |  | **дели и алгоритмы интел**- | | | | | | | |  |
|  |  |  |  | **лектуальной** | | | | **обработки и** | | | |  |
|  |  |  |  | **анализа** | | **данных**; | | | | **методы** | |  |
|  |  |  |  | **автоматической** | | | | | **обработки** | | |  |
|  |  |  |  | **текстов** | | **с** | **использованием** | | | | |  |
|  |  |  |  | **лингвистических** | | | | |  | **техноло**- | |  |
|  |  |  |  | **гий**, | **частично** | | | |  | **владеет** | |  |
|  |  |  |  | **навыками работы со специ**- | | | | | | | |  |
|  |  |  |  | **альными** | |  |  | **библиотеками** | | | |  |
|  |  |  |  | (**Получены положительные** | | | | | | | |  |
|  |  |  |  | **ответы на** 51 - 70 % **задан**- | | | | | | | |  |
|  |  |  |  | **ных вопросов**); | | | | |  |  |  |  |
|  |  |  |  | 8 **баллов ставится**, **когда** | | | | | | | |  |
|  |  |  |  | **студент** | |  | **демонстрирует** | | | | |  |
|  |  |  |  | **значительное** | | | | **понимание** | | | |  |
|  |  |  |  | **проблемы**, **то есть**: **доста**- | | | | | | | |  |
|  |  |  |  | **точно хорошо знает осо**- | | | | | | | |  |
|  |  |  |  | **бенности** | | | **обработки** | | | | **не**- |  |

**структурированных текстов**

**на естественном** **языке**;

**достаточно хорошо знает и**

**умеет использовать модели**

**и алгоритмы интеллекту**-

**альной обработки и анализа**

**данных**; **методы автомати**-

**ческой обработки текстов с**

**использованием лингви**-

**стических технологий**, **до**-

**статочно хорошоо владеет**

**навыками работы со специ**-

**альными** **библиотеками**

(**Получены положительные**

**ответы на** 71 - 85 % **задан**-

**ных вопросов**);

11 **баллов ставится**, **когда**

**студент демонстрирует**

**полное понимание про**-

**блемы**, **то есть**: **четко знает**

**особенности обработки**

**неструктурированных тек**-

**стов на естественном язы**-

**ке**; **четко знает и умеет ис**-

**пользовать модели и алго**-

**ритмы интеллектуальной**

**обработки и анализа дан**-

**ных**; **методы автоматиче**-

**ской обработки текстов с**

**использованием лингви**-

**стических технологий**,

**полностью владеет навы**-

**ками работы со специаль**-

**ными библиотеками** (**По**-

**лучены положительные**

**ответы на более** 85 % **за**-

**данных вопросов**).

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 3 | **Решение типовой** | **Теоретические ос**- | **З** (**ПК**-8) | 0 **баллов** | | **ставится**, | | **когда** 18 | |
|  | **задачи на компью**- | **новы компьютерной** | **У** (**ПК**-8) | **студент** | |  | **демонстрирует** | | |
|  | **тере** | **лингвистики**, **авто**- | **В** (**ПК**-8, | **непонимание проблемы**, **то** | | | | | |
|  |  | **матическая обра**- | **ДПК**-1) | **есть**: **совершенно не знает** | | | | | |
|  |  | **ботка и анализ тек**- |  | **языки программирования и** | | | | | |
|  |  | **стов на естествен**- |  | **среды разработки для со**- | | | | | |
|  |  | **ном языке** |  | **здания программных** | | | | **про**- | |
|  |  |  |  | **тотипов**, | |  | **выполняющих** | | |
|  |  |  |  | **интеллектуальную** | | | | **обра**- | |
|  |  |  |  | **ботку и анализ данных**, **не** | | | | | |
|  |  |  |  | **умеет создавать программ**- | | | | | |
|  |  |  |  | **ные прототипы для интел**- | | | | | |
|  |  |  |  | **лектуальной** | | | **обработки** | | **и** |
|  |  |  |  | **анализа данных**, **совер**- | | | | | |
|  |  |  |  | **шенно не владеет навыка**- | | | | | |
|  |  |  |  | **ми** | **создания программ** | | | | **с** |
|  |  |  |  | **использованием специаль**- | | | | | |
|  |  |  |  | **ных библиотек для анализа** | | | | | |
|  |  |  |  | **текстов** | | **на** | **естественном** | | |
|  |  |  |  | **языке**, **навыками работы с** | | | | | |
|  |  |  |  | **инструментальными** | | | | **сред**- | |
|  |  |  |  | **ствами сбора данных** (**Сде**- | | | | | |
|  |  |  |  | **лано** | 50% | | **задания** | **или** | |
|  |  |  |  | **меньше**). | |  |  |  |  |
|  |  |  |  | 9 **баллов ставится**, **когда** | | | | | |
|  |  |  |  | **студент** | |  | **демонстрирует** | | |

**частичное понимание про**-

**блемы**, **то есть**: **частично**

**знает языки программиро**-

**вания и среды разработки**

**для создания программных**

**прототипов**, **выполняющих**

**интеллектуальную обра**-

**ботку и анализ данных**,

**частично** **умеет создавать**

**программные прототипы**

**для интеллектуальной об**-

**работки и анализа данных**,

**частично владеет** **навы**-

**ками создания программ с**

**использованием специаль**-

**ных библиотек для анализа**

**текстов на** **естественном**

**языке**, **навыками работы с**

**инструментальными** **сред**-

**ствами сбора данных** (**Сде**-

**лано** 51 - 70 % **задания**).

1. **баллов ставится**, **когда студент демонстрирует значительное понимание проблемы**, **то есть**: **доста**-

**точно хорошо знает языки программирования и среды разработки для создания программных прототипов**, **выполняющих интеллекту**-

**альную обработку и анализ данных**, **хорошо умеет создавать программные прототипы для интеллекту**-

**альной обработки и анализа данных**, **хорошо владеет навыками создания про**-

**грамм с использованием специальных библиотек для анализа текстов на естественном языке**, **навы**-**ками работы с инструмен**-

**тальными средствами сбора данных** (**Сделано** 71 - 85 % **задания**).

1. **баллов ставится**, **когда студент демонстрирует полное понимание про**-**блемы**, **то есть**: **четко зна**-**ет языки программирова**-

**ния и среды разработки для создания программных прототипов**, **выполняющих интеллектуальную обра**-**ботку и анализ данных**, **полностью умеет созда**-**вать программные прото**-

**типы для интеллектуальной обработки и анализа дан**-**ных**, **в совершенстве вла**-

**деет навыками создания программ с использовани**-

**ем специальных библиотек**

**для анализа текстов на**

**естественном языке**, **навы**-

**ками работы с инструмен**-

**тальными средствами сбора**

**данных** ( (**Сделано более**

85 % **задания**).

**Теоретические вопросы для промежуточной аттестации по дисци-плине**

1. **Основные понятия и определения компьютерной лингвистики**.
2. **Цель**, **задачи**, **направления компьютерной лингвистики**.
3. **Лингвистические технологии и ресурсы**.
4. **Направления исследований и разработок компьютерной лингвисти**-

**ки**.

* 1. **Природа и сущность языка**. **Функции языка**. **Язык и речь**.
  2. **Язык как знаковая система и его особенности**.
  3. **Словосочетание** (**коллокация**). **Предложение**. **Текст**.
  4. **Инвентарные и конструктивные единицы**.
  5. **Избыточность и контекстная предсказуемость в тексте**.
  6. **Основы речевой коммуникации**. **Коллокации и кострукции**. **Прин**-**цип шкалирования**.
  7. **Анализ текста**. **Сложности и особенности автоматического анализа текста на естественном языке**.
  8. **Коммуникативная и информационная структуры текста**.
  9. **Когнитивный анализ текста**.
  10. **Избыточность**, **компрессия и свертки текста**.
  11. **Исследование информационного пространства**. **Информационный**

**поток**.

* 1. **Моделирование информационных потоков**. **Модель диффузии ин**-**формации**.
  2. **Словарный и предиктивный морфологический анализ текста**.
  3. **Словарный и бессловарный морфологический анализ текста**.
  4. **Инструментарий для морфологического анализа и синтеза текста**.
  5. **Языковые модели**. **Цепи Маркова**, n-**граммы**.
  6. **Статистические методы определения части речи**. **Частеречевая раз**-**метка на базе скрытых Марковских цепей и алгоритм Витерби**.
  7. **Исправление опечаток**. **Расстояние Левенштейна**, **расстояние Ле**-**венштейна**-**Дамерау**.
  8. **Морфологическая классификация естественных языков**.
  9. **Автоматизированное снятие омонимии**. **Постморфологический ана**-

**лиз**.

* 1. **Автоматизированное снятие омонимии**. **Синтаксическая сегмента**-

**ция**.

* 1. **Синтаксическая неоднозначность**. **Подходы к описанию синтаксиса**
* **естественном языке**.

1. **Грамматики**. **Иерархия Хомского**. **Задача синтаксического разбора**. **Оценка точности синтаксического анализа**.
2. **Методы и алгоритмы синтаксического разбора в контексте грамматики зависимостей**.
3. . **Методы и алгоритмы синтаксического разбора в контексте** **КС**-**грамматик**
4. **Формальные методы семантического анализа**.

31. **Понятие онтологии**. **Онтологические ресурсы и компьютерные те**-**заурусы**.

1. **Модели представления знаний в компьютерной семантике**.
2. **Дистрибутивная семантика**.
3. **Классификация полнотекстовых документов**. **Алгоритмы классифи**-**кации с учителем**. **Алгоритм Роккио**.
4. **Классификация полнотекстовых документов**. **Алгоритмы классифи**-**кации с учителем** . **Алгоритм** `**наивной**` **байесовской классификации**.
5. **Классификация полнотекстовых документов**. **Алгоритмы классифи**-**кации с учителем**. **Алгоритм деревьев принятия решений**.
6. **Классификация полнотекстовых документов**. **Алгоритмы классифи**-**кации с учителем**. **Оценка результатов классификации**.
7. **Кластерный анализ**. **Иерархические алгоритмы кластеризации**.
8. **Кластерный анализ** .**Неиерархические алгоритмы кластеризации**.
9. **Контент**-**анализ**. **Методы и алгоритмы контент**-**анализа**.
10. **Кластеризация и контент**-**анализ**. **Инструментальные средства кла**-**стеризации и контент**-**анализа**.
11. **Программные средства лингвистической обработки**.
12. **Системы автоматической обработки текстов на естественном языке**. **Представление лексических данных**.
13. **Архитектура инструментальных систем для автоматической обра**-**ботки и анализа текста на естественном языке**.

**Примеры типовых задач для промежуточной аттестации по дисци-плине**

1. **Вариант** 1.

**Для фрагмента текста на естественном языке выполнить его морфоло**-**гический разбор**.

1. **Вариант** 2.

**Для фрагмента текста на естественном языке выполнить его синтакси**-**ческий разбор**.

1. **Вариант** 3.

**Для фрагмента текста на естественном языке построить семантическую сеть и карту понятий**

**Макет экзаменационного билета**

**Утверждаю:**

**Зав**. **кафедрой**

**к**.**т**.**н**., **доцент**

\_\_\_\_\_\_\_\_\_\_\_**А**.**И**. **Фролов**

«\_\_\_\_» \_\_\_\_\_\_\_\_\_\_\_\_\_\_ 20\_\_**г**.

09.03.03

4 **курс**

**о**

**ФГБОУ ВО** «**Орловский государственный университет имени И**.**С**. **Тургенева** » **Институт приборостроения**, **автоматизации и информационных технологий Кафедра программной инженерии**

**Дисциплина** «**Автоматическая обработка текстов**»

**Билет № 1**

1) **Компьютерная лингвистика**. **Цель**, **задачи**, **направления компьютерной** **лингвистики**.

1. **Грамматики зависимостей**. **Методы и алгоритмы синтаксического разбора в контексте грамматики зависимостей**.
2. **Типовая задача**.

**Разработал**:

**доцент**, **к**.**т**.**н**.\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ **Конюхова О**.**В**.

**Рассмотрены и одобрены на заседании кафедры** «\_\_» \_\_\_\_\_\_\_ 20\_\_**г**.

**Протокол №** \_