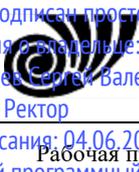


Документ подписан простой электронной подписью Информация о владельце: ФИО: Таскаев Сергей Валерьевич Должность: Ректор Дата подписания: 04.06.2025 12:28:49 Уникальный программный ключ: 04c19ed8bfb98f3b6cb77a486b9a8788b87237333	 МИНОБРАЗОВАНИЯ И НАУКИ РОССИИ Федеральное государственное бюджетное образовательное учреждение высшего образования «Челябинский государственный университет» (ФГБОУ ВО «ЧелГУ»)	Рабочая программа дисциплины "Современные технологии поиска и обработки информации" по направлению подготовки (специальности) 02.04.02 "Фундаментальная информатика и информационные технологии" направленности (профилю) Робототехника ФГБОУ ВО «ЧелГУ»	стр. 1
---	---	--	--------

**Рабочая программа дисциплины (модуля)\***  
**Современные технологии поиска и обработки информации**

Направление подготовки (специальность)

02.04.02 Фундаментальная информатика и информационные технологии

Направленность (профиль)

Робототехника

Присваиваемая квалификация (степень)

магистр

Форма обучения

очная

Год(ы) набора 2025

\*Рабочая программа дисциплины (модуля) адаптирована для инклюзивного обучения инвалидов и лиц с ограниченными возможностями здоровья

Челябинск 2025 г.



## Содержание

1. Цели освоения дисциплины
2. Место дисциплины в структуре ОПОП
3. Компетенции обучающегося, формируемые в результате освоения дисциплины (модуля)
4. Объем дисциплины (модуля)
5. Структура и содержание дисциплины (модуля)
6. Фонд оценочных средств
  - 6.1. Перечень видов оценочных средств
  - 6.2. Типовые контрольные задания и иные материалы для текущей аттестации
  - 6.3. Типовые контрольные вопросы и задания для промежуточной аттестации
  - 6.4. Критерии оценивания
7. Учебно-методическое и информационное обеспечение дисциплины (модуля)
  - 7.1. Рекомендуемая литература
  - 7.2. Перечень ресурсов информационно-телекоммуникационной сети "Интернет"
  - 7.3. Перечень информационных технологий
8. Материально-техническое обеспечение дисциплины (модуля)
9. Методические указания для обучающихся по освоению дисциплины (модуля)
10. Специальные условия освоения дисциплины обучающимися с инвалидностью и ограниченными возможностями здоровья



### 1. ЦЕЛИ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Цель изучения дисциплины - дать студентам теоретические знания и навыки их применения в области поиска информации. Разобрать модели поиска информации, выполнение эффективной индексации текста. Также рассмотрение вопроса о кластеризации и классификации документов.

Изучение дисциплины направлено на развитие следующих индикаторов

УК-1.1. Критически анализирует проблемную ситуацию с целью выработки стратегии действий, аргументированно формулирует собственные суждения и оценки.

### 2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОПОП

Цикл (раздел) ОПОП: ФТД.В.01

#### 2.1 Требования к предварительной подготовке обучающегося:

знания, умения и навыки в объеме школьного курса информатики.

Цифровая обработка сигналов

#### 2.2 Дисциплины и практики, для которых освоение данной дисциплины (модуля) необходимо как предшествующее:

дисциплины, требующие при освоении использование современных методов поиска и обработки информации.

Подготовка к процедуре защиты и защита выпускной квалификационной работы

### 3. КОМПЕТЕНЦИИ ОБУЧАЮЩЕГОСЯ, ФОРМИРУЕМЫЕ В РЕЗУЛЬТАТЕ ОСВОЕНИЯ ДИСЦИПЛИНЫ (МОДУЛЯ)

**УК-1: Способен осуществлять критический анализ проблемных ситуаций на основе системного подхода, выработать стратегию действий**

#### Знать:

Для достижения УК-1.1:

Знать основы выполнения эффективного поиска информации.

#### Уметь:

Для достижения УК-1.1:

Уметь определять критерии системного анализа для поставленных задач.

#### Владеть:

Для достижения УК-1.1:

Владеть навыками системного анализа и поиска информации.

**В результате освоения дисциплины обучающийся должен**

<b>3.1</b>	<b>Знать:</b>
3.1.1	основные существующие алгоритмы поиска информации.
<b>3.2</b>	<b>Уметь:</b>
3.2.1	пользоваться различными моделями поиска информации.
<b>3.3</b>	<b>Владеть:</b>
3.3.1	кластеризации и классификации документов.



#### 4. ОБЪЕМ ДИСЦИПЛИНЫ (МОДУЛЯ)

<b>Общая трудоемкость</b>	<b>1 ЗЕТ</b>
Часов по учебному плану : 36	Виды контроля в семестрах: зачеты 1
в том числе :	
аудиторные занятия : 16	
самостоятельная работа : 18,3	
контактная работа: 17,7	
ИКР: 1,7	

#### 5. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

Код занятия	Наименование разделов и тем /вид занятия/	Семестр / Курс	Часов	Литература
<b>Раздел 1. Информационный поиск</b>				
1.1	Понятие информационного поиска, его задачи и классификация. /Лек/	1	2	Л1.1 Л1.2Л2.1 Э1 Э2
1.2	Информационный поиск в вебе: его становление и развитие. Понятие информационного поиска. /Ср/	1	1	Л1.1 Л1.2Л2.1 Э1 Э2
1.3	Оценка информационного поиска. Оценка ранжированных и неранжированных результатов поиска. Оценка релевантности. Оценка информационно-поисковой системы. /Лек/	1	2	Л1.1 Л1.2Л2.1 Э1 Э2
<b>Раздел 2. Булев поиск</b>				
2.1	Пример информационного поиска. Обработка булевых запросов. Сравнение расширенной булевой модели и ранжированного поиска. Модель булева поиска. Его особенности и применение. /Лек/	1	4	Л1.1 Л1.2Л2.1 Э1
2.2	Первая попытка создать инвертированный индекс. /Ср/	1	1	Л1.1 Л1.2Л2.1 Э2
2.3	Булев поиск. Составление матрицы инцидентности и инвертированного списка. /Ср/	1	1	Л1.1 Л1.2Л2.1 Э1 Э2 Э3
<b>Раздел 3. Инвертированный индекс</b>				
3.1	Понятие и применение инвертированного индекса при поиске информации. /Лек/	1	1	Л1.1 Л1.2Л2.1 Э1
3.2	Построение индекса. Основы аппаратного обеспечения. Блочное индексирование, основанное на сортировке. /Лек/	1	1	Л1.1 Л1.2Л2.1 Э2
3.3	Однопроходное индексирование в оперативной памяти. Распределенное и динамическое индексирование, другие типы индексов. /Ср/	1	1,3	Л1.1 Л1.2Л2.1 Э1 Э2
<b>Раздел 4. Электронные библиотечные системы</b>				
4.1	Регистрация в электронных библиотечных системах "Лань" и "Университетская библиотека онлайн". Их возможности, поиск научной литературы в ЭБС. Научная электронная библиотека Elibrary: регистрация и поиск научных статей. /Ср/	1	1	Л1.1 Л1.2Л2.1 Э1 Э2 Э3
<b>Раздел 5. Поисковые системы</b>				
5.1	Особенности различных поисковых систем, принципы их работы. /Ср/	1	1	Л1.1 Л1.2Л2.1 Э1 Э2 Э3
5.2	Изучение принципов и алгоритмов, на которые опираются поисковые системы. /Ср/	1	1	Л1.1 Л1.2Л2.1 Э1 Э2 Э3
<b>Раздел 6. Основы поиска в вебе</b>				
6.1	Характеристики веба, опыт пользователей поисковых систем. /Ср/	1	1	Л1.1 Л1.2Л2.1 Э1 Э2
<b>Раздел 7. Лексикон и списки словопозиций</b>				
7.1	Схематизация документа и декодирование последовательности символов. Определение лексикона терминов. /Лек/	1	2	Л1.1 Л1.2Л2.1 Э1 Э2



7.2	Быстрое пересечение инвертированных списков с помощью указателей пропусков. /Ср/	1	1	Л1.1 Л1.2Л2.1 Э1 Э2
7.3	Поэтапное создание инвертированного индекса, координатный индекс, обработка запросов при помощи координатного индекса. /Ср/	1	1	Л1.1 Л1.2Л2.1 Э1 Э2 Э3
<b>Раздел 8. Словари и нечеткий поиск</b>				
8.1	Поисковые структуры для словарей. Запросы с джокером. Исправление опечаток. Фонетические исправления. /Лек/	1	4	Л1.1 Л1.2Л2.1 Э1 Э2
8.2	Методы обработки запросов, содержащих орфографические ошибки и другие неточности. /Ср/	1	2	Л1.1 Л1.2Л2.1 Э1 Э2
8.3	Обработка запросов, содержащих орфографические ошибки. Soundex индекс /Ср/	1	2	Л1.1 Л1.2Л2.1 Э1 Э2
<b>Раздел 9. Сжатие индекса</b>				
9.1	Сжатие инвертированного индекса. Алгоритмы сжатия словаря и списка словопозиций. /Ср/	1	2	Л1.1 Л1.2Л2.1 Э1
9.2	Алгоритмы сжатия инвертированного индекса. /Ср/	1	2	Л1.1 Л1.2Л2.1 Э2 Э3
<b>Раздел 10. Иная контактная работа</b>				
10.1	Индивидуальные консультации, текущий контроль /ИКР/	1	1,7	Л1.1 Л1.2Л2.1 Э1 Э2 Э3

## 6. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ

### 6.1. Перечень видов оценочных средств

Самостоятельная работа №1;  
Самостоятельная работа №2;  
Самостоятельная работа №3;  
Самостоятельная работа №4;  
Самостоятельная работа №5;  
Самостоятельная работа №6.  
Вопросы для зачета.

### 6.2. Типовые контрольные задания и иные материалы для текущей аттестации

Пример самостоятельной работы "Условное форматирование в MS Excel"

Условное форматирование в MS Excel позволяет отформатировать числовые данные или текст в таблице, в соответствии заданным условиям или правилам. Благодаря ему, взглянув на нужные ячейки, Вы сразу сможете оценить значения, так как все данные будут представлены в удобном наглядном виде.

Кнопка «Условное форматирование» находится на вкладке «Главная» в группе «Стили».

Кликнув по ней, откроется меню с видами условного форматирования.

Задание 1

- Используем в качестве исходной таблицы книгу «Проверка условий и логические функции в MS Excel», скопировав ее и присвоив ей имя «Условное форматирование в MS Excel».
- Выделим все числа в столбце Цена и нажмем на кнопку Условное форматирование. Начнем с Гистограмм. Excel предлагает 12 гистограмм. Выберем один из видов:
- В результате будет видно, какие товары имеют более высокую или более низкую стоимость:
- К столбцу Количество для заказа применим условное форматирование Цветовые шкалы:
- Для столбца количество применим условное форматирование Набор значков.
- Создадим собственное правило для столбца Количество, выполнив команду Условное форматирование – Управление правилами. В окне диспетчера правил нажмем кнопку Изменить правило и установим параметры:
- Нажмем Ок и окно диспетчера примет вид:
- Закроем окно диспетчера нажатием кнопки Ок и получим результат:

Задание 2

- Рассмотрим более сложное условие по форматированию. Выделим столбец Поставщик. В инструменте Условное форматирование выделим вариант Правила выделения ячеек.
- Выберем вариант Равно и зададим условие:
- Результат примет вид:
- Для столбца Срок годности определим условное форматирование по параметру Правила отбора первых и последних значений:
- Выберем из списка условие Ниже среднего, в результате чего получим результат:



Текст самостоятельной работы

Подготовка презентации.

Требования к презентации

1. Выполнить задание на сайте [csu.ac.ru/~les](http://csu.ac.ru/~les) Создание презентации.
2. Создать презентацию по выбранной теме (список тем смотри выше).
3. Содержание лекций в презентации не должно повторяться – только новая информация!
4. Презентация должна содержать не менее 25 слайдов.
5. Размер шрифта – 20 – 24 пт.
6. Использовать графики, рисунки, диаграммы, схемы, музыкальное оформление и т.д.
7. Смена слайдов – по щелчку. Анимационные эффекты не применять.
8. Слайды пронумеровать, в колонтитуле указать автора.
9. На последнем слайде – список использованной литературы.
10. Подготовить образец выдач 4 или 6 слайдов на листе. Вид выдач смотри по ссылке «Создание презентации». Распечатать.
11. Внимание! Сохранить файл в формате Презентация (.ppt). Структура имени файла:

Тема№\_фамилия\_группа

12. Презентация оценивается следующим образом:

s Содержание 3 б.

1. Раскрытие темы (1 б.) – представленная информация должна полностью раскрывать заданную тему
2. Оптимальность (1 б.) – структурирование, выбор и выделение наиболее важной информации
3. Новизна (1 б.) – новое в данной теме, перспективы развития, конкретные примеры

s Оформление 2 б.

1. Шрифт, колонтитулы, номера слайдов (1 б.) – применить стандартное оформление или разработать собственное, оптимальный размер шрифта – 24, в колонтитуле указать тему презентации и автора, установить нумерацию слайдов и дату.
2. Графики, рисунки, диаграммы, схемы, музыкальное оформление (0,5 б.) – можно вставлять рисунки в слайды с текстом, но слайдов, содержащих только рисунки, не должно быть более 20%. В каждой презентации должны присутствовать не менее трех объектов разных наименований (например, рисунок, схема и диаграмма).
3. Выдачи (0,5 б.) – создать, оформить образец выдач

s Снижается оценка в случае:

1. Отсутствия списка литературы (-1 б.)
2. Несвоевременной сдачи задания (-1 б. за каждую неделю)

Текст самостоятельной работы

Подготовка текста к публикации

Подготовить к публикации текст по выбранной теме, выполнив следующие этапы:

1. Создать собственный шаблон на основе шаблона Normal.dotx для оформления текста. Имя файла: Фамилия (Ваша собственная).
2. Создать свои стили и включить их в собственный шаблон (не меняя стили шаблона Normal.dotx):
  - o Стили заголовков трех уровней. Форматы стилей заголовков должны удовлетворять следующим условиям.
    - § Заголовок первого уровня с именем Глава\_фамилия/ (например, Глава\_Иванов/): записывается прописными буквами, шрифт – ARIAL (синий, полужирный, 16 кегль), выравнивание по центру, с новой страницы.
    - § Заголовок второго уровня с именем Параграф\_фамилия/ : записывается строчными буквами (начиная с прописной буквы), шрифт – Tahoma (зеленый, полужирный, 14 кегль), выравнивание по левому краю (без отступа).
    - § Заголовок второго уровня с именем Пункт\_фамилия/ : записывается строчными буквами (начиная с прописной буквы), шрифт – Garamond (сиреневый, полужирный, курсив, подчеркивание, 12 кегль), выравнивание по левому краю (с отступом на 1 см).
    - § Стил основного текста. Имя стиля – Текст\_фамилия/. Формат стиля должен удовлетворять следующим условиям: предложение записывается строчными буквами (начиная с прописной буквы), шрифт – Times New Roman (черный, 12), выравнивание по ширине (с отступом для первой строки на 1,5 см).
    - o Стил маркированного списка (имя стиля – Сп\_М\_фамилия/).
    - o Стил иерархического списка (имя стиля – Сп\_И\_фамилия/).
3. По созданному шаблону создать документ, в котором будет содержаться текст реферата. В текст скопировать справку MS Word по выбранной теме с разбиением его на части по трем уровням вложенности: не менее трех частей первого уровня (глава), каждая из которых содержит не менее двух частей второго уровня (параграф), каждая из которых содержит не менее двух частей третьего уровня (пункт). Каждая часть текста третьего уровня должна содержать не менее двух абзацев основного текста. Если текста в справке недостаточно, вставить его несколько раз для получения достаточного объема. Оформить заголовки трех уровней с помощью созданных стилей.
4. Структура реферата:



- Обложка: первая страница, на которой расположены (сверху вниз) название организации, название темы, автор (курс, факультет, Ф.И.О.) и проверяющий (кафедра, Ф.И.О, должность) в таблице без рамок.
- Оглавление: вторая страница (создается автоматически с использованием средства MS Word), которая отражает заголовки трех уровней, предметный указатель и список литературы.
- Текст реферата.
- Предметный указатель на отдельной странице (создается автоматически с использованием средства MS Word).
- Список литературы на отдельной странице.
- 5. В тексте создать два списка: маркированный и иерархический (многоуровневый).
- 6. Создать сноски для параграфов (заголовки второго уровня), с указанием места, откуда был взят текст параграфа.
- 7. Создание колонтитулов вверху на каждой странице, кроме первой. Каждая глава должна иметь свой колонтитул, включающий в себя название главы. Предметный указатель и оглавление также должны иметь свои колонтитулы.
- 8. Создать нумерацию страниц кроме первой, внизу по центру.
- 9. Текст, содержащийся в одном из пунктов, расположить в две колонки.
- 10. Создать два рисунка используемых окон MS Word (скопировать окно с помощью клавиши PrintScrin, обработать в редакторе Paint) по выбранной теме и вставить их в текст реферата с названиями и порядковыми номерами, содержащими номер главы (нумерация должна поддерживаться автоматически).
- 11. В последний пункт включить таблицу, отражающую свойства документа (количество символов в документе, знаков и т.п.).
- 12. Реферат сохранить как файл с именем Реферат\_Ваша фамилия\_№ группы.docx.

### 6.3. Типовые контрольные вопросы и задания для промежуточной аттестации

#### Вопросы к зачету

1. Понятие информационного поиска, его цель и задачи;
2. Классификация информационного поиска;
3. Модель булева поиска;
4. Определение эффективности системы информационного поиска: точность и полнота;
5. Инвертированный индекс. Этапы его построения;
6. Оптимизация запроса;
7. Структурная единица документа. Проблема детализации индексирования;
8. Лексема и термин: определение и различия;
9. Стоп-слова: определение, способ создания списка стоп-слов, использование стоп-слов в системах информационного поиска;
10. Нормализация лексем. Классы эквивалентности;
11. Индексация ударений и диакритических знаков;
12. Индексирование заглавных букв;
13. Стемминг. Пример;
14. Лемматизация. Пример;
15. Указатели пропуска;
16. Фразовые запросы. Двухсловные индексы;
17. Фразовые запросы. Координатный индекс;
18. Комбинированная схема обработки фразовых запросов;
19. Реализация словаря. Хеширование;
20. Реализация словаря. Деревья поиска;
21. Запросы с джокером и их обработка;
22. К-граммный индекс для обработки запросов с джокером;
23. Реализация исправления опечаток;
24. Фонетические исправления. Soundex-индекс;
25. Характеристики аппаратного обеспечения, влияющие на обработку запросов;
26. Архитектура MapReduce;
27. Сжатие словаря: цель и методы;
28. Сжатие инвертированного файла;
29. Классификация текстов.

### 6.4. Критерии оценивания

Зачет проводится в присутствии преподавателя и предполагает краткий ответ на вопросы. Вопросы составляются с учётом материала, пройденного на лекционных занятиях. Итоговая оценка выставляется по балльной системе. Студенту необходимо ответить на два теоретических вопроса. Каждый ответ оценивается в 5 баллов. Для того, чтобы получить максимальное количество баллов, необходимо предоставить полный ответ на вопрос. Суммируются баллы, полученные на зачете (10 максимум), и баллы, полученные на самостоятельные работы. Для получения максимального количества баллов необходимо выполнить все задания самостоятельной работы без ошибок в установленные сроки, ответить на вопросы преподавателя во время защиты работы. Каждая самостоятельная работа



МИНОБРНАУКИ РОССИИ  
Федеральное государственное бюджетное образовательное  
учреждение высшего образования  
«Челябинский государственный университет» (ФГБОУ ВО «ЧелГУ»)

Рабочая программа дисциплины "Современные технологии поиска и обработки информации" по направлению подготовки (специальности) 02.04.02 "Фундаментальная информатика и информационные технологии" направленности (профилю) Робототехника ФГБОУ ВО «ЧелГУ»

стр. 8

оценивается от 0 до 10 баллов. Частичное выполнение заданий, допущенные ошибки при их выполнении или при ответе на вопросы преподавателя приводят к снижению количества баллов за самостоятельную работу. Полученные студентами баллы суммируются, итоговая оценка выставляется исходя из полученной суммы баллов:

От 0 до 50 баллов – «не зачтено».

От 51 до 70 баллов – «зачтено».

## 7. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

### 7.1. Рекомендуемая литература

#### 7.1.1. Основная литература

	Авторы,	Заглавие	Издательство,	Ресурс
Л1.1	Гагарина Л.Г., Слюсарь В.В., Слюсарь М.В.	Основы информационных технологий: учебное пособие ( <a href="https://znanium.com/catalog/document?id=389618">https://znanium.com/catalog/document?id=389618</a> )	Москва : ООО "Научно-издательский центр ИНФРА-М", 2022	ЭБС
Л1.2	Артамонов В. Н.	Современные технологии поиска и обработки информации: учебное пособие ( <a href="http://library.csu.ru/rbooks2/view2?code=local/007972/007972">http://library.csu.ru/rbooks2/view2?code=local/007972/007972</a> )	Челябинск : Издательство Челябинского государственног о университета, 2022	ЭБС

#### 7.1.2. Дополнительная литература

	Авторы,	Заглавие	Издательство,	Ресурс
Л2.1	Резник В. Г.	Системный анализ, управление и обработка информации. Часть II: методические указания по практическим занятиям аспирантов по направлению «09.06.01 - информатика и вычислительная техника» (профиль 05.13.01 «системный анализ, управление и обработка информации (информация и информационные системы, экономика, энергетика, промышленность, образование») ( <a href="https://e.lanbook.com/book/312965">https://e.lanbook.com/book/312965</a> )	Москва : ТУСУР, 2018	ЭБС

### 7.2. Перечень ресурсов информационно-телекоммуникационной сети "Интернет"

Э1	Лань [Электронный ресурс] : электронно-библиотечная система (ЭБС) / издательство Лань. – URL: <a href="http://e.lanbook.com/">http://e.lanbook.com/</a> .
Э2	Университетская библиотека онлайн [Электронный ресурс] : электронно-библиотечная система (ЭБС) / ООО Директмедиа Паблишинг. – URL: <a href="http://biblioclub.ru/">http://biblioclub.ru/</a> .
Э3	eLIBRARY.RU [Электронный ресурс] : электронная библиотека / Науч. электрон. б-ка. – URL: <a href="http://elibrary.ru/defaultx.asp">http://elibrary.ru/defaultx.asp</a> .

### 7.3 Перечень информационных технологий

#### 7.3.1 Программное обеспечение

Adobe Reader

LMS Moodle

OpenOffice

#### 7.3.2 Профессиональные базы данных и информационно-справочные системы

1. Электронный каталог научной библиотеки ЧелГУ [Электронный ресурс] : база данных / Челяб. гос. ун-т. – Челябинск, 1992 .

2. eLIBRARY.RU : научная электронная библиотека : сайт. – Москва, 2000 – . – URL: <https://elibrary.ru> – Режим доступа: для зарегистрир. пользователей. – Текст : электронный.

3. Mathematical Reviews (MR) : реферативная база данных / American Mathematical Society. – URL: <http://www.ams.org/mathscinet/> – Яз. рус., англ. – Режим доступа: для зарегистрир. пользователей ЧелГУ. – Текст : электронный.

## 8. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)



Для реализации дисциплины используются учебные аудитории для проведения занятий лекционного типа, лабораторных занятий, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, а также помещения для самостоятельной работы.

Учебные аудитории укомплектованы специализированной мебелью и техническими средствами обучения (компьютерная техника с подключением к сети "Интернет" для лабораторных занятий).

Для проведения занятий лекционного типа предлагаются наборы демонстрационного оборудования и учебно-наглядных пособий, такие как презентации.

Помещения для самостоятельной работы обучающихся оснащены компьютерной техникой с подключением к сети "Интернет" и обеспечением доступа в электронную информационно-образовательную среду университета

## 9. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ (МОДУЛЯ)

В результате изучения теоретических основ дисциплины и реализации в процессе обучения требований по прикладной направленности дисциплины, а также самостоятельной работы студент должен знать, уметь и владеть составляющими компетенций, определенных в программе.

Основными видами учебных занятий являются: лекции, лабораторные занятия, самостоятельная работа и зачет.

Аудиторные лекции студента имеют своей целью формирование целостной системы знаний по изучаемому предмету. Студент может воспользоваться основной и дополнительной литературой.

Самостоятельная работа студента начинается с внимательного ознакомления с программой данной дисциплины.

Требуется творческое отношение к самой Программе учебного курса. Вопросы, составляющие ее содержание, обладают разной степенью важности. Есть вопросы, выполняющие функцию логической связки содержания темы и всего курса, имеются вопросы описательного или разъяснительного характера. Эти вопросы не составляют сути, понятийного, концептуального содержания темы, но необходимы для целостного восприятия изучаемых проблем. Успешно освоив теоретический материал, студент будет готов к проведению лабораторных занятий, которые рассматриваются как дальнейшее углубление и расширение знаний по предмету.

В освоении дисциплины (модуля) инвалидами и лицами с ограниченными возможностями здоровья большое значение имеет индивидуальная работа. Под индивидуальной работой подразумевается две формы взаимодействия с преподавателем: индивидуальная учебная работа (консультации), т.е. дополнительное разъяснение учебного материала и углубленное изучение материала с теми обучающимися, которые в этом заинтересованы, и индивидуальная воспитательная работа. Индивидуальные консультации по предмету является важным фактором, способствующим индивидуализации обучения и установлению воспитательного контакта между преподавателем и обучающимся инвалидом или обучающимся с ограниченными возможностями здоровья.

В случае применения при обучении дисциплины электронного обучения, дистанционных образовательных технологий общение обучающихся и преподавателя осуществляется в режиме реального времени (чаты, видеоконференции и др.) или отложенного времени (система дистанционного обучения Moodle, форумы, электронная почта).

Большую часть времени обучающиеся самостоятельно работают с учебно-методическими материалами. Студенты имеют возможность консультироваться с преподавателем по всем вопросам, возникающим в ходе самостоятельной работы посредством электронной почты, социальных сетей, Moodle.

Доступ обучающегося к учебным ресурсам в режиме отложенного времени, самостоятельной работы осуществляется через сеть Интернет в удобном для него месте, времени и темпе.

При обучении лиц с ограниченными возможностями здоровья электронное обучение, дистанционные образовательные технологии предусматривают возможность приема-передачи информации в доступных для них формах.

Реализация дисциплины с применением электронного обучения, дистанционных образовательных технологий (далее – ЭО, ДОТ) осуществляется на основании «Положения о реализации основных и дополнительных образовательных программ с применением электронного обучения и дистанционных образовательных технологий в федеральном государственном бюджетном образовательном учреждении высшего образования «Челябинский государственный университет», «Положения о порядке зачета обучающимися по основным профессиональным образовательным программам высшего образования в ФГБОУ ВО «ЧелГУ» результатов освоения в организациях, осуществляющих образовательную деятельность, учебных предметов, курсов, дисциплин (модулей), практик, дополнительных образовательных программ» посредством электронной информационно-образовательной среды ФГБОУ ВО «ЧелГУ». В исключительных случаях (форс-мажор и т.п.) при реализации образовательной деятельности с применением ЭО, ДОТ могут применять компоненты, не входящие в перечень электронной информационно-образовательной среды.

## 10. СПЕЦИАЛЬНЫЕ УСЛОВИЯ ОСВОЕНИЯ ДИСЦИПЛИНЫ ОБУЧАЮЩИМИСЯ С ИНВАЛИДНОСТЬЮ И ОГРАНИЧЕННЫМИ ВОЗМОЖНОСТЯМИ ЗДОРОВЬЯ



Освоение дисциплины инвалидами и лицами с ограниченными возможностями здоровья осуществляется с использованием специальных технических средств и информационных технологий, предоставляемых Ресурсным учебно-методическим центром по обучению инвалидов и лиц с ограниченными возможностями здоровья ЧелГУ по запросу обучающегося (мобильные специальные технические средства для лиц с нарушениями зрения и с нарушением слуха, ассистивные информационные технологии).

При необходимости для обучающихся с нарушениями зрения на рабочих местах для проведения практических или лабораторных занятий устанавливается специальное программное обеспечение (программа речевой навигации, речевые синтезаторы, экранные лупы).

В учебные аудитории обеспечивается беспрепятственный доступ для обучающихся с инвалидностью и с ограниченными возможностями здоровья. В каждой аудитории, где обучаются инвалиды и лица с ограниченными возможностями здоровья, предусматривается соответствующее количество мест для обучающихся с учетом нарушений их здоровья.

Для освоения дисциплины инвалидам и лицам с ограниченными возможностями здоровья предоставляется доступ к печатным источникам, имеющимся в научной библиотеке ЧелГУ, с помощью специальных технических средств; доступ с помощью специальных технических и программных средств к электронным источникам, представленным в форме электронного документа в фонде научной библиотеки ЧелГУ или электронно-библиотечных системах.

Учебно-методические материалы для обучающихся из числа инвалидов и лиц с ограниченными возможностями здоровья предоставляются в формах, адаптированных к ограничениям их здоровья и особенностям восприятия информации.

Для инвалидов и лиц с ограниченными возможностями здоровья освоение дисциплины может быть частично или полностью осуществлено с использованием дистанционных образовательных технологий.

При проведении промежуточной аттестации по дисциплине обучающимся с инвалидностью и с ограниченными возможностями здоровья обеспечивается по их заявлению предоставление в доступной форме в зависимости от их индивидуальных особенностей инструкции о порядке проведения промежуточной аттестации, оценочных средств и возможности ответов на задания (письменно на бумаге, набор ответов на компьютере, письменно шрифтом Брайля, с использованием услуг ассистента, устно).

При проведении процедуры оценивания результатов обучения инвалидов и лиц с ограниченными возможностями здоровья предусматривается использование предоставленных ЧелГУ или собственных технических средств, необходимых им в связи с их индивидуальными особенностями. При необходимости инвалидам и лицам с ограниченными возможностями здоровья предоставляется дополнительное время для подготовки ответа на задания, процедура оценивания результатов обучения по дисциплине может проводиться в несколько этапов.

## Самостоятельная работа №1 по теме «Булев поиск».

### *Вариант 1*

Прочитайте следующие документы:

#### Документ 1.

В Интернете с каждым днём скапливается всё больше информации, когда-либо созданной и вновь создаваемой людьми. Равнодоступность большей части информации в Интернете уравнивает возможности доступа к этой информации как обычных пользователей Интернета и журналистов локальных СМИ, так и сотрудников мировых информационных агентств. Благодаря Интернету перед каждым человеком ежедневно и даже ежесекундно открывается доступ к многомиллионной аудитории, которой он может передать свой информационный материал, полученный, например, с помощью обычного мобильного телефона с диктофоном и встроенной фотокамерой. Следовательно, уровень монополизации деятельности по распространению информации также снижается благодаря Интернету.

#### Документ 2.

До недавнего времени ограничения в прямой коммуникации между людьми, порожденные пространством и временем, во многом определяли потребность людей в услугах журналистов. По мере роста общего количества пользователей Интернета, а среди них – числа владеющих английским языком, эти ограничения всё в большей степени снимаются, что закономерно ведёт к уменьшению спроса на услуги журналистов. Одновременно с этим растёт объем «сырой» информации, доступной каждому отдельному пользователю Интернета, что актуализирует проблему её отбора и редактирования. Последнее всегда входило в перечень функций журналистики, но с ростом числа пользователей Интернета эффективный информационный поиск начинает приобретать всё большую значимость не только в журналистской деятельности, но и в других разнообразных сферах общественной деятельности. Таким образом, информационный поиск – это процесс поиска неструктурированной документальной информации.

#### Документ 3.

Поиск информации представляет собой процесс выявления в некотором множестве документов (текстов), которые посвящены заданной теме (предмету) и удовлетворяют заранее определенному условию поиска (запросу), а также содержат необходимые (соответствующие информационной потребности) факты, сведения и данные. Процесс поиска включает последовательность операций, направленных на сбор, обработку и предоставление необходимой информации заинтересованным лицам.

#### Документ 4.

Комплекс программ, предназначенных для информационного поиска, называется поисковой машиной. Обычно является частью поисковой системы – автоматизированного программно-аппаратного комплекса с веб-интерфейсом,

предоставляющего возможность поиска информации в Интернете. Самая известная поисковая система в мире – это Google, самая популярная в России – Яндекс, а одной из самых старых поисковых систем является Yahoo. Как уже было отмечено ранее, в архитектуре поисковой системы можно выделить поисковую машину – ядро системы, представленное набором программных модулей; базу данных или индекс, хранящую информацию обо всех известных поисковой системе Интернет ресурсах; и набор сайтов, являющихся собой точки входа пользователей в систему. Все это соответствует классической трехуровневой архитектуре информационных систем: есть пользовательский интерфейс, бизнес логика, которая в данном случае представлена реализацией алгоритмов поиска и база данных.

#### Документ 5.

Для того, чтобы найти в Интернете требуемую информацию, необходимо знать либо адрес её местоположения (например, адрес html-страницы или файла), либо пользователя Интернета, который может предоставить информацию. Если мы не знаем ни адреса, ни человека, который мог бы нам помочь, то следует перейти к вопросам «Как можно узнать адрес размещения информации?» или «Как найти человека, который мог бы нам помочь с поиском информации?». При этом не следует переоценивать возможности Интернета. Лучшие результаты может дать совмещение онлайн-овых и оффлайн-овых методов поиска информации.

#### Документ 6.

Сегодня существует достаточно большое количество методов информационного поиска в Интернете и через Интернет. В каждом конкретном случае успешность поиска определяется знаниями возможных методов и навыками владения ими, знанием этнических языков, на которых эта информация может быть представлена, либо нашими социальными связями.

Постройте матрицу инцидентности «термин — документ» для терминов: *Интернет, поиск, запрос, индекс, информация, система, машина*. Словоформа и регистр значения не имеют. Термины расположите в таблице в алфавитном порядке.

Обработайте запрос *поиск AND информация AND NOT интернет*, взяв векторы для терминов из матрицы инцидентности и выполнив поразрядные операции AND и NOT. В ответе укажите список документов.

Составьте инвертированный список для коллекции документов Вашего варианта используя в качестве словаря термины из пункта 2.

### Самостоятельная работа №2 по теме «Лексикон и списки словопозиций»

#### *Вариант 1*

1. Разбейте текст каждого документа на лексемы:

#### Документ 1.

Состав и интерпретация поддерживаемых метасимволов. Часто называется «диалектом» регулярного выражения.

#### Документ 2.

Особенности взаимодействия регулярных выражений с языком или программой.

#### Документ 3.

Специфика применения регулярных выражений к тексту.

Выполните предварительную лингвистическую обработку лексем с помощью нормализации лексем и игнорирования стоп-слов (союзы, предлоги).

3. Составьте инвертированный индекс для данной коллекции документов с указанием частоты и списка словопозиций для каждого термина. Термины расположите в алфавитном порядке.

4. Нормализуйте лексемы в запросе: *регулярные AND выражения* и обработайте его.

### Самостоятельная работа №3 по теме «Словопозиции с координатами и фразовые запросы»

#### *Вариант 1*

Разбейте текст каждого документа на лексемы:

#### Документ 1.

Как правило, поисковые машины поддерживают два режима: режим простого поиска и режим расширенного поиска.

#### Документ 2.

Можно просто вводить через пробел одно или несколько слов; поиск слов всевозможными окончаниями моделируется символом \* в конце слова.

#### Документ 3.

Многие системы позволяют искать словосочетания или фразу, для этого необходимо ее заключить в кавычки.

#### Документ 4.

Возможно обязательное включение или исключение определенных слов.

#### Документ 5.

Основная проблема поиска по примитивно составленному запросу заключается в том, что поисковая машина найдет все страницы, на которых указанные слова встречаются в любой части документа.

#### Документ 6.

Как правило, количество найденных страниц будет слишком велико.

. Составьте координатный индекс для данной коллекции документов в формате термин док1: <позиция1, позиция2, ...>; док2: <позиция1, позиция2, ...>; и т.д., используя следующий словарь терминов: поиск, машина, запрос, режим, слово, документ, страница. Термины расположите в алфавитном порядке.

3. Какие документы удовлетворяют фразовому запросу поисковая машина?

4. Укажите набор документов, удовлетворяющих запросу: режим /2 поиск.

#### Самостоятельная работа №4 по теме «Поиск по сходству»

##### Вариант 1

С помощью перестановочного индекса определите термины лексикона **поисковик, иск, поиск, постфикс, пост, пуск**, соответствующие запросу **по\*ск**.

Используя алгоритм динамического программирования, вычислите расстояние редактирования между строками **сильный** и **стильный**.

Составьте биграммный индекс терминов лексикона **стена, струна, сторона, оборона, рана, охрана** и найдите термины, содержащие по крайней мере две из всех биграмм в запросе **страна**.

Найдите коэффициент Жаккара для строк запроса и каждого термина, полученного в предыдущем пункте. Для каких терминов этот коэффициент наибольший?

#### Самостоятельная работа №5 по теме «Фонетические исправления»

##### Вариант 1

И  
собор, луг, забор, сбор, информатика, запор, инфракрасный, инвентарь, зубр, лук

п

о

##### Таблица кодирования русских букв

лу, е, ё, ы, а, о, э, я, и, ю, ь, ъ

ьб, п

зф, в

Уж, з, с, х

Як, г

. ц, ч, ш, щ

а д, т

л л, й

о м, н

р р

и  
Какие термины имеют одинаковые Soundex-коды?

м  
Преобразуйте термин запроса **СОБР** в Soundex-код и выполните поиск по soundex-индексу.

о

н

е

т

и

Самостоятельная работа №6 по теме «Сжатие индекса»

*Вариант 1*

1. Создайте блочное хранение ( $k = 4$ ) для следующего словаря терминов: поиска, информации, поиску, сбора, информационный, поисковой, сборный, информация, сбору, поисковик, сборного, информацию.
2. Выполните дальнейшее сжатие с помощью фронтальной упаковки. Для обозначения общего префикса используйте символ  $\diamond$  (Alt-код символа Alt+9674).
3. Вычислите коды, полученные с помощью схемы байтового кодирования для следующего инвертированного списка: <4, 10, 11, 12, 15, 62, 63, 265, 268, 270, 400>  
Запишите бинарные коды блоками по 8 бит.

