

Документ подписан простой электронной подписью Информация о владельце: ФИО: Таскаев Сергей Валерьевич Должность: Ректор	МИНОБРАЗОВАНИЯ И НАУКИ РОССИИ Федеральное государственное бюджетное образовательное учреждение высшего образования «Челябинский государственный университет» (ФГБОУ ВО «ЧелГУ»)	
Дата подписания: 08.04.2025 15:37:45 Уникальный программный ключ: 04c19ed8bfb981506cb77a48609a878808522525	Рабочая программа дисциплины "Современные технологии поиска и обработки информации" по направлению подготовки (специальности) 02.03.02 "Фундаментальная информатика и информационные технологии" направленности (профилю) Прикладное программирование и системы искусственного интеллекта ФГБОУ ВО «ЧелГУ»	стр. 1

Рабочая программа дисциплины (модуля)*

Современные технологии поиска и обработки информации

Направление подготовки (специальность)

02.03.02 Фундаментальная информатика и информационные технологии

Направленность (профиль)

Прикладное программирование и системы искусственного интеллекта

Присваиваемая квалификация (степень)

бакалавр

Форма обучения

очная

Год(ы) набора 2026

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

Челябинск 2026 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. Выполняет поиск информации, определяет критерии системного анализа поставленных задач

УК-1.2. Использует критический анализ, систематизацию и обобщение информации для решения поставленных задач

ОПК-6.1. Имеет представление об основных существующих информационных технологиях, используемых при решении профессиональных задач.

ОПК-6.2. Демонстрирует умения использовать существующие информационные технологии при решении задач профессиональной деятельности.

ОПК-6.3. Имеет практический опыт использования существующих информационных технологий для решения задач профессиональной деятельности.

2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОПОП

Цикл (раздел) ОПОП: К.М.01.01

2.1 Требования к предварительной подготовке обучающегося:

Дисциплина базируется на школьной программе.

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

Данная дисциплина является основой для дальнейших практик.

Алгоритмы и структуры данных

Учебная практика (практика по программированию)

Методы машинного обучения

Программирование на Python: библиотечные технологии

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

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

Знать:

Для достижения УК-1.1. студенту необходимо знать: как выполняется поиск информации, как определяются критерии системного анализа поставленных задач.

Уметь:

Для достижения УК-1.2. студенту необходимо уметь: использовать критический анализ, систематизировать и обобщать информацию для решения поставленных задач

Владеть:

Для достижения УК-1.1:
Владеть навыками системного анализа и поиска информации.

ОПК-6: Способен понимать принципы работы современных информационных технологий и использовать их для решения задач профессиональной деятельности

Знать:

Для достижения ОПК-6.1. студенту необходимо знать: основные существующие информационные технологии, используемые при решении профессиональных задач.

Уметь:

Для достижения ОПК-6.2. студенту необходимо уметь: использовать существующие информационные технологии при решении задач профессиональной деятельности.



Владеть:

Для достижения ОПК-6.3. студенту необходимо иметь практический опыт использования существующих информационных технологий для решения задач профессиональной деятельности.

В результате освоения дисциплины обучающийся должен

3.1	Знать:
3.1.1	Основные существующие алгоритмы поиска информации
3.2	Уметь:
3.2.1	Пользоваться различными моделями поиска информации
3.3	Владеть:
3.3.1	Кластеризации и классификации документов

4. ОБЪЕМ ДИСЦИПЛИНЫ (МОДУЛЯ)

Общая трудоемкость	2 ЗЕТ
Часов по учебному плану : 72	Виды контроля в семестрах: зачеты 1
в том числе :	
аудиторные занятия : 32	
самостоятельная работа : 39,8	
контактная работа: 32,2 ИКР: 0,2	

5. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

Код занятия	Наименование разделов и тем /вид занятия/	Семестр / Курс	Часов	Литература
Раздел 1. Информационный поиск				
1.1	Введение в информационный поиск и обработку естественного языка /Лек/	1	2	Л1.1 Л1.2Л2.1 Э1 Э2
1.2	Введение в информационный поиск и обработку естественного языка /Ср/	1	4,1	Л1.1 Л1.2Л2.1 Э1 Э2
1.3	Модель Булево поиска. Инвертированный индекс. /Лаб/	1	2	Л1.1 Л1.2Л2.1 Э1 Э2
Раздел 2. Введение в обработку естественно- языковых текстов				
2.1	Нормализация текста, лексикон, списки слово позиций. /Лек/	1	2	Л1.1 Л1.2Л2.1 Э1
2.2	Лексикон и список слово позиций /Ср/	1	4	Л1.1 Л1.2Л2.1 Э2
2.3	Нормализация текста, лексикон, списки слово позиций, классы эквивалентности. /Лаб/	1	3	Л1.1 Л1.2Л2.1 Э1 Э2 Э3
Раздел 3. Словари и нечеткий поиск				
3.1	Перестановочный и k граммный индексы /Лек/	1	2	Л1.1 Л1.2Л2.1 Э1
3.2	Исправление опечаток, расстояние редактирования. /Лек/	1	2	Л1.1 Л1.2Л2.1 Э2
3.3	Словари и нечеткий поиск /Ср/	1	4	Л1.1 Л1.2Л2.1 Э1 Э2
3.4	Исправление опечаток и запросы с пропусками /Лаб/	1	3	Л1.1 Л1.2Л2.1
Раздел 4. Ранжирование, взвешивание терминов и модель векторного пространства				
4.1	Ранжирование, взвешивание терминов и модель векторного пространства /Ср/	1	4	Л1.1 Л1.2Л2.1 Э1 Э2
4.2	Модель векторного пространства для ранжирования. Ранжирование в полнофункциональной поисковой системе /Лек/	1	2	Л1.1 Л1.2Л2.1



4.3	Ранжирование документов (VSM) /Лаб/	1	2	Л1.1 Л1.2Л2.1
Раздел 5. Оценка информационного поиска				
5.1	Оценка информационного поиска План лекции /Лек/	1	2	Л1.1 Л1.2Л2.1 Э1 Э2
5.2	Оценка информационного поиска План лекции /Ср/	1	5	Л1.1 Л1.2Л2.1 Э1 Э2
5.3	Оценка ранжированных результатов поиска /Лаб/	1	1	Л1.1 Л1.2Л2.1 Э1 Э2 Э3
Раздел 6. Обработка естественного языка				
6.1	Классификация в векторном пространстве. Плоская кластеризация. Иерархическая кластеризация. /Лек/	1	2	Л1.1 Л1.2Л2.1 Э1 Э2
6.2	Классификация в векторном пространстве. Плоская кластеризация. Иерархическая кластеризация. /Ср/	1	5	Л1.1 Л1.2Л2.1 Э1 Э2
6.3	Классификация в векторном пространстве. Плоская кластеризация. Иерархическая кластеризация. /Лаб/	1	3	Л1.1 Л1.2Л2.1 Э1 Э2
6.4	Методы обработки естественных языков /Лек/	1	2	Л1.1 Л1.2Л2.1
6.5	Методы обработки естественных языков /Лаб/	1	2	Л1.1 Л1.2Л2.1
Раздел 7. Иная контактная работа				
7.1	Индивидуальные консультации, текущий контроль /ИКР/	1	0,2	Л1.1 Л1.2Л2.1 Э1 Э2 Э3
Раздел 8. Подготовка к зачету				
8.1	Подготовка к зачету /Ср/	1	13,7	Л1.1 Л1.2Л2.1

6. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ

6.1. Перечень видов оценочных средств

Лабораторная работа №1,
Лабораторная работа №2,
Лабораторная работа №3,
Лабораторная работа №4,
Лабораторная работа №5.

6.2. Типовые контрольные задания и иные материалы для текущей аттестации

Пример варианта Лабораторной:

Написать программу которая выполняет следующий функционал:

1. На вход подается файл, в котором есть список терминов (нормализованные слова) - это будет словарь
2. Пользователь вводит свое слово (не обязательно орфографически верное), программа выдает список терминов из словаря у которых минимальное расстояние редактирования к запросу (и максимальный индекс Жаккара).

6.3. Типовые контрольные вопросы и задания для промежуточной аттестации

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

Каждая лабораторная работа оценивается от 0 до 13 баллов, а именно:

1. Выполнение всех технических требований - максимум 10 баллов;
2. Ответ на 2 вопроса по программе на понимание - максимум 3 балла.

Активная работа на занятиях - максимум 10 баллов.

На зачете студент может заработать максимум 30 баллов.

Если за семестр студент заработал 61 или более баллов, тогда оценка - зачтено

Если студент за семестр заработал менее 61 балла, тогда оценка - незачтено.

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

7.1. Рекомендуемая литература

7.1.1. Основная литература

	Авторы,	Заглавие	Издательство,	Ресурс
Л1.1	Громов Ю. Ю., Дидрих И. В., Иванова О. Г., Ивановский М. А., Однолько В. Г.	Информационные технологии: учебник (https://biblioclub.ru/index.php?page=book&id=444641)	Тамбов : Тамбовский государственный технический университет (ТГТУ), 2015	ЭБС
Л1.2	Ганегедара Т.	Обработка естественного языка с TensorFlow (https://e.lanbook.com/book/140584)	Москва : ДМК Пресс, 2020	ЭБС

7.1.2. Дополнительная литература

	Авторы,	Заглавие	Издательство,	Ресурс
Л2.1	Гасанов Э. Э., Кудрявцев В. Б.	Теория хранения и поиска информации (https://znanium.com/catalog/document?id=259903)	Москва : Издательская фирма "Физико- математическая литература" (Ф ИЗМАТЛИТ), 2002	ЭБС

7.2. Перечень ресурсов информационно-телекоммуникационной сети "Интернет"

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

7.3 Перечень информационных технологий

7.3.1 Программное обеспечение



Adobe Reader

LMS Moodle

LibreOffice

Python

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, занятий семинарского типа, 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 практических 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, практические 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 практических заданий, 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, видео-конференции 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 применяться 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 учебных 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.

