

Министерство науки и высшего образования Российской  
Федерации Федеральное государственное бюджетное  
образовательное учреждение высшего образования  
«Кубанский государственный университет»

Факультет компьютерных технологий и прикладной  
математики Кафедра вычислительных технологий

УТВЕРЖДАЮ

Проректор по учебной работе,  
качеству образования – первый  
проректор

Хагуров Т.А.

подпись

подпись

« 30 » \_\_\_\_\_ мая \_\_\_\_\_ 2025г.

## РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ Б1.В.02 «МЕТОДЫ ИЗВЛЕЧЕНИЯ ИНФОРМАЦИИ ИЗ СЕТЕВЫХ ИСТОЧНИКОВ»

Направление подготовки/специальность 01.04.02. Прикладная математика  
и информатика

(код и наименование направления подготовки/специальности)

Направленность (профиль) /

специализация Магистерская программа 01.04.02. " Технологии  
программирования и разработки информационно-коммуникационных  
систем"

Программа подготовки академическая магистратура

Форма обучения очная

Квалификация выпускника магистр

Краснодар 2025

Рабочая программа дисциплины Б1.В.02 «МЕТОДЫ ИЗВЛЕЧЕНИЯ  
ИНФОРМАЦИИ ИЗ СЕТЕВЫХ ИСТОЧНИКОВ» составлена в соответствии  
с федеральным государственным образовательным стандартом высшего  
образования (ФГОС ВО) по направлению подготовки 02.04.02  
Фундаментальная информатика и информационные технологии

Программу составил: А.Н. Полетайкин, доц. каф. ВТ, к.т.н., доц.

\_\_\_\_\_

подпись

Рабочая программа дисциплины «Мультиагентные системы»  
утверждена на заседании кафедры Вычислительных Технологий  
протокол № 7 от «07 » мая 2025 г.

Руководитель магистерской программы Вишняков Ю.М

(фамилия, инициалы)

\_\_\_\_\_

подпись

И.о. заведующего кафедрой (разработчика) Еремин А.А.

\_\_\_\_\_

(фамилия, инициалы)

\_\_\_\_\_

подпись

Утверждена на заседании учебно-методической комиссии  
факультета Компьютерных Технологий и Прикладной  
Математики

протокол № 4 от «23» мая 2025 г

Председатель УМК факультета Коваленко А.

\_\_\_\_\_

фамилия, инициалы

\_\_\_\_\_

подпись

Рецензенты:

Гаркуша О.В., доцент кафедры  
информационных технологий ФБГОУ ВО  
«Кубанский государственный  
университет», кандидат физико-  
математических наук.

Схаляхо Ч.А., доцент КВВУ им.С.М.Штеменко, к.ф.-м.н., доцент

## 1. ЦЕЛИ И ЗАДАЧИ ОСВОЕНИЯ ДИСЦИПЛИНЫ

### 1.1 Цель освоения дисциплины

**Целью** дисциплины «Методы извлечения информации из сетевых источников» является обучение передовым методам, моделям, средствам и технологиям поиска и компьютерной обработки информации.

**Задачи** дисциплины:

Дать **знания** о:

- истории и тенденциях развития информационно-поисковых систем, крупных ученых, участвовавших в их разработке;
- основных принципах обмена данными в глобальной сети Интернет;
- основных методах функционирования информационно-поисковых систем;
- основных современных инструментальных средствах их разработки;
- основных методах программирования поиска, как на стороне сервера, так и на стороне клиента.

Привить **умение**:

- использовать в проектируемых и эксплуатируемых информационных системах и технологиях современные средства Интернет-программирования;
- применять информационно-поисковые системы при нахождении в сети Интернет требуемой информации для проектно-конструкторской, организационно-управленческой и научно-исследовательской деятельности;
- разрабатывать информационно-поисковые системы для нахождения данных на стороне сервера или клиента.

### 1.2 Задачи дисциплины

Изучить историю и тенденции развития информационно-поисковых систем, работы крупных ученых, участвовавших в их разработке. Научиться основным принципам обмена данными в глобальной сети Интернет; основным методам функционирования информационно-поисковых систем; методам программирования поиска, как на стороне сервера, так и на стороне клиента, научиться использовать современные инструментальные средства разработки поисковых систем.

### 1.3 Место дисциплины (модуля) в структуре образовательной программы

Дисциплина «Методы извлечения информации из сетевых источников» относится к вариативной части дисциплин учебного плана.

Пререквизитами данной дисциплины являются дисциплины математического и естественнонаучного цикла ООП подготовки бакалавров «Алгебра», «Иностранный язык», «Методы программирования», «WEB-программирование», а также дисциплины "Распределенные программные системы".

Знания, получаемые при изучении дисциплины, используются при работе над магистерской диссертацией.

### 1.4 Перечень планируемых результатов обучения по дисциплине (модулю), соотнесенных с планируемыми результатами освоения образовательной программы.

Изучение данной учебной дисциплины направлено на формирование у обучающихся следующих **профессиональных компетенций**:

| № п.п. | Индекс компетенции | Содержание компетенции (или её части)                                                                                                                                                   | В результате изучения учебной дисциплины обучающиеся должны                                                                                                                                  |                                                                                                                                                                                                            |                                                                                                                                                                                                                                                             |
|--------|--------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|        |                    |                                                                                                                                                                                         | Знать                                                                                                                                                                                        | Уметь                                                                                                                                                                                                      | Владеть                                                                                                                                                                                                                                                     |
| 1      | ПК-3               | Способен эффективно применять алгоритмические и программные решения в области информационно-коммуникационных технологий, а также участвовать в их проектировании и разработке           | основные принципы построения архитектур распределенных систем для поиска и анализа информации в сетевых источниках, методы научных исследований и инструменты для систематизации результатов | проводить научные исследования и получать новые научные и прикладные результаты самостоятельно и в составе научного коллектива при разработке систем для поиска и анализа информации в сетевых источниках; | методологией научных исследований и методами систематизации их результатов; устной и письменной формой изложения результатов научной деятельности, приемами разработки информационно-поисковых систем для нахождения данных на стороне сервера или клиента. |
| 2      | ПК-4               | Способен находить и извлекать актуальную научно-техническую информацию из электронных библиотек, реферативных журналов и т.п. для решения задач в области профессиональной деятельности | Принципы построения систем для сбора актуальной научно-технической информации из сетевых источников                                                                                          | находить и извлекать актуальную научно-техническую информацию из электронных библиотек, реферативных журналов, расположенных в сетевых источниках                                                          | Навыками и инструментами сбора актуальной научно-технической информации из электронных библиотек, реферативных журналов, расположенных в сетевых источниках                                                                                                 |

## 2. СОДЕРЖАНИЕ И СТРУКТУРА ДИСЦИПЛИНЫ

### 2.1 Распределение трудоёмкости дисциплины по видам работ

Общая трудоёмкость дисциплины составляет 3 зач.ед. (108 час.), их распределение по видам работ представлено в таблице (для студентов ОФО)

| Вид учебной работы                    | Всего часов | Семестры (часы) |  |  |  |
|---------------------------------------|-------------|-----------------|--|--|--|
|                                       |             | 3               |  |  |  |
| <b>Контактная работа в том числе:</b> | 28,2        | 28,2            |  |  |  |
| <b>Аудиторные занятия (всего):</b>    | 28          | 28              |  |  |  |
| В том числе:                          |             |                 |  |  |  |

|                                                                              |             |             |  |  |  |
|------------------------------------------------------------------------------|-------------|-------------|--|--|--|
| Занятия лекционного типа                                                     | 14          | 14          |  |  |  |
| Занятия семинарского типа (семинары, практ. занятия)                         | –           | –           |  |  |  |
| Лабораторные занятия                                                         | 14          | 14          |  |  |  |
| <b>Иная контрольная работа</b>                                               | –           | –           |  |  |  |
| Промежуточная аттестация (ИКР)                                               | 0,2         | 0,2         |  |  |  |
| <b>Самостоятельная работа (всего)</b>                                        | <b>79,8</b> | <b>79,8</b> |  |  |  |
| В том числе:                                                                 |             |             |  |  |  |
| Курсовая работа                                                              | –           | –           |  |  |  |
| <i>Проработка учебного (теоретического) материала</i>                        | 29,8        | 29,8        |  |  |  |
| <i>Выполнение индивидуальных заданий (подготовка сообщений, презентаций)</i> | 42          | 42          |  |  |  |
| <i>Подготовка к текущему контролю</i>                                        | 8           | 8           |  |  |  |
| <b>Контроль:</b>                                                             | зачет       | зачет       |  |  |  |
| Подготовка к экзамену:                                                       | –           | –           |  |  |  |
| Общая трудоемкость<br>в т.ч. контактная работа<br>зач. ед.                   | 108         | 108         |  |  |  |
|                                                                              | 28,2        | 28,2        |  |  |  |
|                                                                              | 3           | 3           |  |  |  |

## 2.2 Структура дисциплины:

Распределение видов учебной работы и их трудоемкости по разделам дисциплины.  
Разделы дисциплины, изучаемые в 3 семестре магистратуры (*очная форма*)

| №  | Наименование разделов                                                                      | Количество часов |                   |            |           |                      |
|----|--------------------------------------------------------------------------------------------|------------------|-------------------|------------|-----------|----------------------|
|    |                                                                                            | Всего            | Аудиторная работа |            |           | Внеаудиторная работа |
|    |                                                                                            |                  | Л                 | ИКР        | ЛР        | СРС                  |
| 1  | 2                                                                                          | 3                | 4                 | 5          | 6         | 7                    |
| 1. | <b>Раздел 1. Аналитика в сети Интернет</b>                                                 | 34               | 2                 |            | 2         | 30                   |
| 2. | <b>Раздел 2.</b> Методология сбора данных из сетевых источников                            | 42               | 8                 | 0,2        | 8         | 25,8                 |
| 3. | <b>Раздел 3.</b> Типы информационных систем. Устройство и принцип работы поисковых систем. | 32               | 4                 |            | 4         | 24                   |
| 4. | <b>Итого</b>                                                                               | <b>108</b>       | <b>14</b>         | <b>0,2</b> | <b>14</b> | <b>79,8</b>          |
| 5. | <i>Итого по дисциплине:</i>                                                                | <b>108</b>       |                   |            |           |                      |

Примечание: Л – лекции, КСР – контрольные и самостоятельные работы, ЛР – лабораторные занятия, СРС – самостоятельная работа студента

## 2.3 Содержание разделов дисциплины:

### 2.3.1 Занятия лекционного типа

| № раздела | Наименование раздела                       | Содержание раздела                                                                                                               | Форма текущего контроля |
|-----------|--------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------|-------------------------|
| 1         | 2                                          | 3                                                                                                                                | 4                       |
| 1         | <b>Раздел 1. Аналитика в сети Интернет</b> | <b>Тема 1.</b> Генезис сети Интернет.<br><b>Тема 2.</b> Структура WEB, Deep WEB.<br><b>Тема 3.</b> Системы управления контентом. | ЛР                      |

|   |                                                                                                  |                                                                                                                                                                                                                                                                                                                                                              |    |
|---|--------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| 2 | <b>Раздел 2.</b><br>Методология сбора и анализа данных из сетевых источников                     | <b>Тема 4.</b> Технологии извлечения знаний из WEB - WEB-mining.<br><b>Тема5.</b> Понятие <i>data scraping</i> или «срезание данных с поверхности». Классификация способов извлечения информации из WEB-источников.<br><b>Тема 6.</b> Модели информационного поиска.                                                                                         | ЛР |
| 3 | <b>Раздел 3.</b><br>Типы информационных систем.<br>Устройство и принцип работы поисковых систем. | <b>Тема 7.</b> Типология, структура и функция информационных систем. Системы переработки информации. Типы информационных систем. Информационные системы Интернета.<br><b>Тема 8.</b> Автоматическое индексирование. Семантический взб. Искусственный интеллект. Разработка ИПТ. Отраслевой тезаурус.<br><b>Тема 9.</b> Способы хранения больших данных в WEB | ЛР |

ЛР – лабораторные работы

### Расшифровка тем дисциплины

- Генезис сети Интернет.** История создания Сети. Развитие электрических и электронных средств связи. ARPANET. Всемирная паутина. Развитие интернет в XXI веке. Организационная структура Интернета. Схема адресации в сети Интернет. Модель BOW TIE. Понятия и различия WEB 2.0- WEB 4.0.
- Структура WEB, Deep WEB.** Невидимый WEB, его возможности и характеристики. Инструменты и технологии работы в невидимом WEB.
- Системы управления контентом.** Проблемы, возникающие при поддержании актуальности информации на сайте. Определение CMS. Краткое описание CMS. Динамический и статический сайты. Характеристика контента. Создание контента. Управление автоматизированными деловыми процессами. Распространение контента. Персонализация и глобализация контента. Критерии классификации систем управления контентом. Простая CMS. Шаблонная CMS. Профессиональная CMS. Универсальная CMS. Функциональные и технологические возможности систем управления контентом. Требования к системам управления контентом. Вопросы, решаемые при выборе системы управления контентом.
- Технологии извлечения знаний из WEB - WEB-mining.** Определение понятий WEB Mining и Data Mining? Отличия между ними. Задачи и этапы извлечения знаний из WEB. Направления WEB-mining: Извлечение Web-контента (Web Content Mining); Извлечение Web-структур (Web Structure Mining); Исследование использования Web-ресурсов (Web Usage Mining)
- Понятие *data scraping* или «срезание данных с поверхности».** Понятие бизнес-аналитического решения. Анализ журнала посещаемости сайта. Заказные статистические исследования. Определение профиля сайта. Определение перечня сайтов, посещаемых вашей аудиторией. Определение целевой аудитории сайта. Типы посетителей сайтов. Модели поведения посетителей сайта. Пользователи Интернет магазинов.
- Модели информационного поиска.** Булева модель, векторная модель, вероятностная модель, гибридная модель. Математические особенности обработки информации разными моделями. Сферы их применения.
- Типология, структура и функция информационных систем.** Системы переработки информации. Типы информационных систем. Уточнение структуры информационных систем. Информационные системы Интернета.

8. **Устройство и принцип работы поисковых систем.** Понятие поисковой системы. Принципы работы поисковых систем, которые нужно учитывать при продвижении сайта. Виды поисковых роботов. Порядок индексации сайтов. Порядок поисковой выдачи. Принципы алгоритмов выдачи поисковой системы Яндекс и Google. Выбор ключевых слов для продвижения сайта. Типы запросов по частотности. Типы запросов по степени конверсии. Понятие семантического ядра. Создание семантического ядра. Выбор ключевых страниц сайта. Распределение семантического ядра. Анализ сайтов конкурентов. Расчет сложности продвижения сайта. Выбор основной стратегии поискового продвижения сайта.
9. **Способы хранения данных в WEB.** Требования к хранилищам данных, OLTP и OLAP системы. Нереляционные базы данных.

### 2.3.3 Лабораторные занятия

| № работы | № раздела дисциплины | Наименование лабораторных работ                                                                                                                                                                                                                                                                                                       | Форма текущего контроля                |
|----------|----------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------|
| 1        | 1                    | Вводное занятие. Настройка необходимого ПО и среды разработки.                                                                                                                                                                                                                                                                        |                                        |
| 2        | 1                    | <b>Классификация интернет-сервисов.</b> Составление запросов по теме магистерской работы, выполнение поиска в открытых и закрытых сетевых источниках, сравнение эффективности поиска с помощью различных инструментов.                                                                                                                | Отчет по лабораторной работе           |
| 3        | 1                    | <b>Системы управления контентом.</b> Обсуждение преимуществ и недостатков различных CMS, особенностей разработки WEB-ресурсов с их помощью.                                                                                                                                                                                           | Отчет по лабораторной работе           |
| 4        | 2                    | <b>Технологии извлечения знаний из WEB – WEB Mining.</b> Понятие <i>data scraping</i> или «срезание данных с поверхности». Используя любой из приведенных либо найденных вами способов извлечения информации с web страниц, разработать программу по сбору информации методами Web-scraping и продемонстрировать результат ее работы. | Отчет по лабораторной работе           |
| 5        | 3                    | <b>Модели информационного поиска. Устройство и принцип работы поисковых систем.</b> Определение и анализ характеристик выбранной поисковой системы: Google, Yandex, Rambler, Yahoo, Bing, AltaVista.                                                                                                                                  | Отчет по лабораторной работе<br>Доклад |
| 6        | 3                    | <b>Способы хранения данных в WEB.</b>                                                                                                                                                                                                                                                                                                 | Устный опрос                           |

### 2.3.4 Примерная тематика курсовых работ (проектов)

Учебным планом не предусмотрены.

### 2.3.5 Расчетно-графические задания

Учебным планом не предусмотрены.

## Самостоятельное изучение разделов дисциплины

**Раздел 1.** Историческое развитие Интернет и сервисов.

**Раздел 2.** Виды, структуры и функции систем управления контентом.

**Раздел 3.** Устройство и принцип работы поисковых систем. Инструментарий для сбора интернет-статистики.

### 2.4 Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине

| № | Вид СРС                                                                                                      | Перечень учебно-методического обеспечения дисциплины по выполнению самостоятельной работы |
|---|--------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------|
| 1 | 2                                                                                                            | 3                                                                                         |
| 1 | <b>Раздел 1.</b> Историческое развитие Интернет и сервисов.                                                  | Список основной литературы                                                                |
| 2 | <b>Раздел 2.</b> Виды, структуры и функции систем управления контентом.                                      | Список основной литературы                                                                |
| 3 | <b>Раздел 3.</b> Устройство и принцип работы поисковых систем. Инструментарий для сбора интернет-статистики. | Список основной литературы                                                                |

Учебно-методические материалы для самостоятельной работы обучающихся из числа инвалидов и лиц с ограниченными возможностями здоровья (ОВЗ) предоставляются в формах, адаптированных к ограничениям их здоровья и восприятия информации:

Для лиц с нарушениями зрения:

- в печатной форме увеличенным шрифтом,
- в форме электронного документа,

Для лиц с нарушениями слуха:

- в печатной форме,
- в форме электронного документа.

Для лиц с нарушениями опорно-двигательного аппарата:

- в печатной форме,
- в форме электронного документа,

Данный перечень может быть конкретизирован в зависимости от контингента обучающихся.

## 3. ОБРАЗОВАТЕЛЬНЫЕ ТЕХНОЛОГИИ

| Семестр | Вид занятия (Л, ПР, ЛР) | Используемые интерактивные образовательные технологии                                                                                     | Количество часов |
|---------|-------------------------|-------------------------------------------------------------------------------------------------------------------------------------------|------------------|
| А       | Л                       | Компьютерные презентации и обсуждение                                                                                                     | 14               |
|         | ЛР                      | Разбор конкретных ситуаций (задач), тренинги по решению задач, компьютерные симуляции (программирование алгоритмов), коллективные проекты | 14               |
| Итого:  |                         |                                                                                                                                           | 28               |

Для лиц с ограниченными возможностями здоровья предусмотрена организация консультаций с использованием электронной почты.

#### **4. ОЦЕНОЧНЫЕ СРЕДСТВА ДЛЯ ТЕКУЩЕГО КОНТРОЛЯ УСПЕВАЕМОСТИ И ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ**

##### **4.1 Фонд оценочных средств для проведения текущего контроля**

Фонд оценочных средств дисциплины состоит из средств текущего контроля выполнения лабораторных работ, коллективных проектов, средств для итоговой аттестации (экзамена).

Оценка успеваемости осуществляется по результатам:

- выполнения лабораторных работ, коллективных проектов;
- ответа на зачете (для выявления знания и понимания теоретического материала дисциплины).

#### **Перечень вопросов, которые выносятся на зачет**

1. Поиск информации в Интернет и его документирование. Требования к поиску
2. Поиск информации в Интернет. Основные поисковые системы, их особенности и отличия
3. Поиск информации в Интернет: основные тенденции и рост
4. Модель Web-пространства Bow Tie: концепция, основные и дополнительные свойства.
5. Открытый и скрытый Web: особенности и отличия, примеры
6. Научные публикации в Интернет: основные ресурсы, особенности поиска
7. Трансграничный научный поиск в Интернет: основные проблемы и особенности
8. Библиотеки в сети Интернет: обзор и сравнительный анализа
9. Современные технологии парсинга данных
10. Понятие парсинга. Виды и методы парсинга данных
11. Библиотеки парсинга данных: особенности и сравнительный анализ
12. Сетевые источники и особенности их парсинга. Особенности источников данных в «обертке» парсинга
13. Правовые аспекты, связанные с использованием парсинга
14. Сервис парсинга данных: архитектура и этапы работы
15. Понятие Data Mining и этапы процесса Data Mining по стандарту CRISP-DM
16. Понятие Web Mining. Технология Web Mining, ее состав и задачи и этапы реализации
17. Data Mining и Web Mining: особенности и отличия
18. Основные направления анализа Web Mining
19. Web Content Mining: основные проблемы и особенности
20. Web Structure Mining: основные проблемы и особенности
21. Понятие «Сильной связности» WEB-графа, типы его узлов. Какому функциональному закону подчиняются сети «тесного мира»?
22. Закономерности и ограничения модели Bow Tie.
23. Понятие WEB 2.0.
24. Deep WEB. Какие ресурсы его составляют. Какими средствами его можно исследовать.
25. Понятия Web Mining и Web Analytics. Этапы аналитики в соответствии со стандартом CRISP-DM.
26. Задачи Data Mining. Направления Data Mining.

27. Понятие и задачи Web Content Mining.
28. Перечислите и охарактеризуйте средства WEB scraping.
29. Методы Text Mining в приложении к специфике WWW.
30. Методологии Web Graph Mining для подхода Web Structure Mining.
31. Основные задачи Web Usage Mining, средства их решения, назначение кластерного анализа в контексте Web Usage Mining.
32. Классификация способов извлечения информации из WEB-источников.
33. Задачи Web-scraping, механизм его работы. Разновидность методов Web-scraping.
34. Этапы работы поисковой системы. Компоненты поискового движка.
35. Суть векторной и вероятностной моделей ИП, их достоинства и недостатки.
36. Способы хранения словарей. Способы нечеткого поиска.
37. Технология Map-Reduce, механизмы работы, примеры использования. Как обеспечивается отказоустойчивость Map-Reduce.
38. Технология Hadoop. MapReduce в Hadoop. Структура программы в Hadoop.
39. Хранилища Больших данных. Примеры распределенных хранилищ.
40. Понятия OLAP и OLTP. Характеристики Больших данных.

## **Критерии оценивания**

### **Критерии оценивания**

#### **Критерии оценивания к зачету:**

Оценка “зачтено” - Практические задания выполнены в срок в объеме не менее 80%. Студент демонстрирует правильные, уверенные действия по применению полученных знаний на практике, грамотное и логически стройное изложение материала при аргументации ответов на вопросы при защите лабораторных.

Оценка «не зачтено» - Практические задания не выполнены либо предоставлены не в срок в объеме менее 60%, Студент демонстрирует наличие грубых ошибок в ответе, непонимание сущности излагаемого вопроса, неумение применять знания на практике, неуверенность и неточность ответов на дополнительные и наводящие вопросы.

Оценочные средства для инвалидов и лиц с ограниченными возможностями здоровья выбираются с учетом их индивидуальных психофизических особенностей.

– при необходимости инвалидам и лицам с ограниченными возможностями здоровья предоставляется дополнительное время для подготовки ответа на экзамене;

– при проведении процедуры оценивания результатов обучения инвалидов и лиц с ограниченными возможностями здоровья предусматривается использование технических средств, необходимых им в связи с их индивидуальными особенностями;

– при необходимости для обучающихся с ограниченными возможностями здоровья и инвалидов процедура оценивания результатов обучения по дисциплине может проводиться в несколько этапов.

Процедура оценивания результатов обучения инвалидов и лиц с ограниченными возможностями здоровья по дисциплине предусматривает предоставление информации в формах, адаптированных к ограничениям их здоровья и восприятия информации:

Для лиц с нарушениями зрения:

- в печатной форме увеличенным шрифтом,
- в форме электронного документа.

Для лиц с нарушениями слуха:

- в печатной форме,
- в форме электронного документа.

Для лиц с нарушениями опорно-двигательного аппарата:

- в печатной форме,

– в форме электронного документа.

Данный перечень может быть конкретизирован в зависимости от контингента обучающихся.

## 5. ПЕРЕЧЕНЬ ОСНОВНОЙ И ДОПОЛНИТЕЛЬНОЙ УЧЕБНОЙ ЛИТЕРАТУРЫ, НЕОБХОДИМОЙ ДЛЯ ОСВОЕНИЯ ДИСЦИПЛИНЫ

### 5.1 Основная литература:

1. Система формирования знаний в среде Интернет / В. И. Аверченков, А. В. Заболеева-Зотова, Ю. М. Казаков [и др.]. – 4-е изд., стер. – Москва : ФЛИНТА, 2021. – 181 с. : ил., схем., табл. – Режим доступа: по подписке. – URL: <https://biblioclub.ru/index.php?page=book&id=93354> (дата обращения: 01.06.2024). – Библиогр. в кн. – ISBN 978-5-9765-1266-5. – Текст : электронный.
2. Симанков, В. С. Методы и алгоритмы поиска информации в Интернете=Search methods and algorithms for information retrieval on the Internet / В. С. Симанков, Д. М. Толкачев. – Москва : Библио-Глобус, 2017. – 332 с. : граф., схем., ил. – Режим доступа: по подписке. – URL: <https://biblioclub.ru/index.php?page=book&id=499077> (дата обращения: 01.06.2024). – Библиогр. в кн. – ISBN 978-5-9500501-8-3. – DOI 10.18334/9785950050183. – Текст : электронный.

Для освоения дисциплины инвалидами и лицами с ограниченными возможностями здоровья имеются издания в электронном виде в электронно-библиотечных системах «*Лань*» и «*Юрайт*».

### 5.2. Дополнительная литература:

1. Ромм, Я.Е. Детерминированный информационный поиск на основе сортировки с распараллеливанием базовых операций / Я.Е. Ромм, С.С. Белоконова. - Москва : Издательство Научный мир, 2014. - 197 с. [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=468725>
2. Артемов, А.В. Мониторинг информации в интернете : учебно-методическое пособие / А.В. Артемов ; Межрегиональная Академия безопасности и выживания.- Орел : МАБИВ, 2014. - 160 с. [Электронный ресурс]. - URL:<http://biblioclub.ru/index.php?page=book&id=428606>

### 5.2. Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины

1. Web Page Scraping using Java, ресурс: URL: <http://half-wit4u.blogspot.ru/2011/01/web-scraping-using-java-api.html>
2. Web Page Scraping with Java, ресурс: URL: <https://www.packtpub.com/books/content/creating-sample-web-scraper>
3. Jaunt Java Web Scraping & JSON Querying , ресурс: [jaunt-api.com](http://jaunt-api.com)
4. Web scraping с Node.js, ресурс: URL: <http://www.webdesignmagazine.ru/internet-technology/other/web-scraping-s-node-js/>
5. Web-scraping средствами R, примеры. Ресурс: URL: [kek.ksu.ru/EOS/WM/WebScraping.docx](http://kek.ksu.ru/EOS/WM/WebScraping.docx)
6. Chang, C.-H.,Kayed, M.,Girgis, M. R.,and Shaalan, K. F. 2006. A survey of web information extraction systems. IEEE Trans. on Knowl. and Data Eng. 18,10, 1411-1428.

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

По курсу предусмотрено проведение лекционных занятий, на которых дается основной систематизированный материал, лабораторных работ, контрольной работы, зачета и экзамена.

Важнейшим этапом курса является самостоятельная работа по дисциплине с использованием указанных литературных источников и методических указаний автора курса.

Виды и формы СР, сроки выполнения, формы контроля приведены выше в данном документе.

Для лучшего освоения дисциплины при защите ЛР студент должен ответить на несколько вопросов из лекционной части курса.

В освоении дисциплины инвалидами и лицами с ограниченными возможностями здоровья большое значение имеет индивидуальная учебная работа (консультации) – дополнительное разъяснение учебного материала.

Индивидуальные консультации по предмету являются важным фактором, способствующим индивидуализации обучения и установлению воспитательного контакта между преподавателем и обучающимся инвалидом или лицом с ограниченными возможностями здоровья.

## **7. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю)**

### **7.1 Перечень информационных технологий.**

- Проверка домашних заданий и консультирование посредством электронной почты.
- Использование электронных презентаций при проведении лекций и практических занятий.

### **7.2 Перечень необходимого программного обеспечения**

1. NetBeans или Eclipse IDE, JDK.
2. Программы для демонстрации и создания презентаций («Microsoft Power Point»).

### **7.3 Перечень информационных справочных систем:**

1. Электронная библиотечная система eLIBRARY.RU (<http://www.elibrary.ru/>)

## **8. Материально-техническая база, необходимая для осуществления образовательного процесса по дисциплине**

| №  | Вид работ            | Материально-техническое обеспечение дисциплины (модуля) и оснащенность                                                                                                                                                                                                                          |
|----|----------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1. | Лекционные занятия   | Лекционная аудитория, оснащенная презентационной техникой (проектор, экран, компьютер/ноутбук) и соответствующим программным обеспечением (ПО) PowerPoint. ауд. 129, 131, А305.                                                                                                                 |
| 2. | Лабораторные занятия | Лаборатория, укомплектованная специализированными техническими средствами обучения – компьютерный класс, с возможностью подключения к сети «Интернет», программой экранного увеличения и обеспеченный доступом в электронную информационно-образовательную среду университета. (лаб. 102-106.). |
| 3. | Групповые            | Аудитория, (кабинет) – компьютерный класс                                                                                                                                                                                                                                                       |

|    |                                                  |                                                                                    |
|----|--------------------------------------------------|------------------------------------------------------------------------------------|
|    | (индивидуальные)<br>консультации                 |                                                                                    |
| 4. | Текущий контроль,<br>промежуточная<br>аттестация | Аудитория, приспособленная для письменного ответа при<br>промежуточной аттестации. |