

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
федеральное государственное бюджетное образовательное учреждение высшего
образования «Вятский государственный университет»
(ВятГУ)
г. Киров

Утверждаю
Директор/Декан Бушмелева Н. А.



Номер регистрации
РПД_3-09.04.03.03_2020_113619
Актуализировано: 02.03.2021

Рабочая программа дисциплины
Модели информационного поиска

	наименование дисциплины
Квалификация выпускника	Магистр
Направление подготовки	09.04.03 шифр
	Прикладная информатика наименование
Направленность (профиль)	3-09.04.03.03 шифр
	Машинное обучение и анализ данных наименование
Формы обучения	Очная наименование
Кафедра-разработчик	Кафедра прикладной математики и информатики (ОРУ) наименование
Выпускающая кафедра	Кафедра прикладной математики и информатики (ОРУ) наименование

Сведения о разработчиках рабочей программы дисциплины

Котельников Евгений Вячеславович

ФИО

Цели и задачи дисциплины

Цель дисциплины	ознакомление с базовыми понятиями теории информационного поиска; основными принципами организации информационно-поисковых систем; алгоритмами аналитико-синтетической переработки документов.
Задачи дисциплины	<ul style="list-style-type: none"> - определение базовых понятий теории информационного поиска, иллюстрация их тесной связи с семиотикой; - изучение основных технологий построения информационно-поисковых систем; - ознакомление с технологиями построения тезаурусов и онтологий; - изучение основных алгоритмов аналитико-синтетической переработки документов.

Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения образовательной программы

Компетенция УК-1

Способен осуществлять критический анализ проблемных ситуаций на основе системного подхода, выработать стратегию действий		
Знает	Умеет	Владеет
основы системного подхода для решения проблем информационного поиска; принципы декомпозиции задач информационного поиска; основные методологические подходы к формированию стратегии действий, направленных на решение поставленной задачи	применять методы системного подхода для критического анализа проблемных ситуаций; находить решение элементарных (атомарных) подзадач информационного поиска; формировать стратегию действий с обоснованием принимаемых решений	навыками построения стратегии действий для разрешения проблемных ситуаций; навыками получения комплексных решений путем композиции решений элементарных (атомарных) подзадач информационного поиска; навыками системного анализа стратегии действий, позволяющими выделять ошибочные элементы

Компетенция ПК-5

Способен осуществлять оценку качества формализации и алгоритмизации поставленных задач, оценку качества и эффективности программного кода, редактирование программного кода и контроль версий программного обеспечения		
Знает	Умеет	Владеет
методы и приемы формализации задач информационного поиска; методы, приемы и средства алгоритмизации поставленных задач; стандартные модели и алгоритмы	использовать методы и приемы формализации задач информационного поиска; использовать методы и приемы алгоритмизации поставленных задач; применять стандартные	навыком исследования математических и информационных моделей информационного поиска; навыком оценки качества формализации и алгоритмизации поставленных задач, оценки

информационного поиска	модели информационного поиска; использовать возможности имеющейся технической и/или программной архитектуры; применять коллективную среду разработки программного обеспечения и систему контроля версий	качества и эффективности программного кода; навыком редактирования программного кода и контроль версий программного обеспечения
------------------------	---	---

Компетенция ПК-6

Способен осуществлять анализ и согласование архитектуры программного обеспечения, оценку качества проектирования программного обеспечения, структуры базы данных, программных интерфейсов		
Знает	Умеет	Владеет
понятие и принципы работы поисковых систем, критерии качества поисковых систем	применять принципы построения архитектуры программного обеспечения и виды архитектур программного обеспечения; применять методологии и средства проектирования программного обеспечения	навыками работы с поисковыми системами

Компетенция ПК-7

Способен осуществлять оценку качества разработанных процедур отладки программного кода, процедур сбора диагностических данных, измерения требуемых характеристик программного обеспечения, тестовых наборов данных в соответствии с выбранной методикой, оценку результатов проверки работоспособности программного обеспечения, рефакторинг и оптимизацию кода		
Знает	Умеет	Владеет
критерии качества и эффективности программного кода; методы и средства проверки работоспособности программного обеспечения; методы подготовки тестовых наборов данных; методы и средства оптимизации программного кода	использовать современные модели информационного поиска; использовать основные, наиболее употребительные методы математической обработки результатов исследований; производить подготовку наборов данных и проверку работоспособности программного обеспечения на их основе	навыком оценки процедур сбора диагностических данных, измерения требуемых характеристик программного обеспечения; навыком формирования тестовых наборов данных в соответствии с выбранной методикой, оценки результатов проверки работоспособности программного обеспечения, рефакторинга и оптимизации кода

Компетенция ПК-8

Способен осуществлять исследование и изучение мировых практик выполнения аналитических работ, выявление проблем и сложностей в существующих практиках,
--

описание методик проведения аналитических работ, их апробацию на выбранных проектах и доработку		
Знает	Умеет	Владеет
<p>существующие мировые практики выполнения аналитических работ их достоинства и недостатки; способы описания методик проведения аналитических работ; методы планирования проектных работ</p>	<p>выявлять проблемы и сложности в существующих практиках выполнения аналитических работ; планировать аналитические работы, планировать ресурсы, управлять проектами</p>	<p>навыком исследования и изучения мировых практик информационного поиска; навыком выявления проблем и сложностей в существующих практиках, описания методик проведения аналитических работ, их апробации на выбранных проектах и доработки</p>

Структура дисциплины
Тематический план

№ п/п	Наименование разделов дисциплины	Шифр формируемых компетенций
1	Индексация и булев поиск	ПК-5, УК-1
2	Компьютерная лингвистика в информационном поиске	ПК-6, ПК-7
3	Обработка и классификация запросов	ПК-8
4	Подготовка и прохождение промежуточной аттестации	ПК-5, ПК-6, ПК-7, ПК-8, УК-1

Формы промежуточной аттестации

Зачет	Не предусмотрен (Очная форма обучения)
Экзамен	3 семестр (Очная форма обучения)
Курсовая работа	Не предусмотрена (Очная форма обучения)
Курсовой проект	Не предусмотрена (Очная форма обучения)

Трудоемкость дисциплины

Форма обучения	Курсы	Семестры	Общий объем (трудоемкость)		Контактная работа, час	в том числе аудиторная контактная работа обучающихся с преподавателем, час				Самостоятельная работа, час	Курсовая работа (проект), семестр	Зачет, семестр	Экзамен, семестр
			Часов	ЗЕТ		Всего	Лекции	Семинарские, практические занятия	Лабораторные занятия				
Очная форма обучения	2	3	180	5	98	54	18	0	36	82			3

Содержание дисциплины

Очная форма обучения

Код занятия	Наименование тем занятий	Трудоемкость, академических часов
Раздел 1 «Индексация и булев поиск»		36.00
Лекции		
Л1.1	Булев поиск. Обработка булевых запросов. Сравнение расширенной булевой модели и ранжированного поиска	4.00
Лабораторные занятия		
Р1.1	Булев поиск	6.00
Р1.2	Инвертированный индекс	4.00
Р1.3	Обработка булевых запросов	4.00
Р1.4	Построение индекса. оценка времени создания индексов для заданной коллекции, используя заданные параметры системы	2.00
Самостоятельная работа		
С1.1	Подготовка отчета о выполнении лабораторных работ	16.00
Раздел 2 «Компьютерная лингвистика в информационном поиске»		62.50
Лекции		
Л2.1	Лексикон и списки словопозиций, определение лексикона терминов	2.00
Л2.2	Нормализация. Виды нормализации	2.00
Л2.3	Стемминг и лемматизация	4.00
Лабораторные занятия		
Р2.1	Нормализация	6.00
Р2.2	Стемминг и лемматизация	6.00
Самостоятельная работа		
С2.1	Подготовка отчета о выполнении лабораторной работы	23.00
Контактная внеаудиторная работа		
КВР2.1	Контактная внеаудиторная работа	19.50
Раздел 3 «Обработка и классификация запросов»		54.50
Лекции		
Л3.1	Словопозиции с координатами и фразовые запросы	2.00
Л3.2	Словари и нечеткий поиск. Поисковые структуры для словарей	2.00
Л3.3	Исправление опечаток	2.00
Лабораторные занятия		
Р3.1	Словари и нечеткий поиск	4.00
Р3.2	Исправление опечаток	4.00
Самостоятельная работа		
С3.1	Подготовка отчета о выполнении лабораторной работы	18.50
Контактная внеаудиторная работа		
КВР3.1	Контактная внеаудиторная работа	22.00
Раздел 4 «Подготовка и прохождение промежуточной аттестации»		27.00

Э4.1	Подготовка к сдаче экзамена	24.50
КВР4.1	Консультация перед экзаменом	2.00
КВР4.2	Сдача экзамена	0.50
ИТОГО		180.00

Содержание дисциплины данной рабочей программы используется при обучении по индивидуальному учебному плану, при ускоренном обучении, при применении дистанционных образовательных технологий и электронном обучении (при наличии).

Методические указания для обучающихся по освоению дисциплины

Успешное освоение дисциплины предполагает активное, творческое участие обучающегося на всех этапах ее освоения путем планомерной, повседневной работы. Обучающийся обязан посещать лекции, семинарские, практические и лабораторные занятия (при их наличии), получать консультации преподавателя и выполнять самостоятельную работу.

Изучение дисциплины следует начинать с проработки настоящей рабочей программы, методических указаний и разработок, указанных в программе, особое внимание уделить целям, задачам, структуре и содержанию дисциплины.

Главной задачей каждой лекции является раскрытие сущности темы и анализ ее основных положений. Тематика лекций определяется настоящей рабочей программой дисциплины.

Лекции – это систематическое устное изложение учебного материала. На них обучающийся получает основной объем информации по каждой конкретной теме. Лекции обычно носят проблемный характер и нацелены на освещение наиболее трудных и дискуссионных вопросов.

Предполагается, что обучающиеся приходят на лекции, предварительно проработав соответствующий учебный материал по источникам, рекомендованным программой. Часто обучающимся трудно разобраться с дискуссионными вопросами, дать однозначный ответ. Преподаватель, сравнивая различные точки зрения, излагает свой взгляд и нацеливает их на дальнейшие исследования и поиск научных решений. После лекции желательно вечером перечитать и закрепить полученную информацию, тогда эффективность ее усвоения значительно возрастает. При работе с конспектом лекции необходимо отметить материал, который вызывает затруднения для понимания, попытаться найти ответы на затруднительные вопросы, используя предлагаемую литературу. Если самостоятельно не удалось разобраться в материале, сформулируйте вопросы и обратитесь за помощью к преподавателю.

Целью семинарских занятий является проверка уровня понимания обучающимися вопросов, рассмотренных на лекциях и в учебной литературе.

Целью практических и лабораторных занятий является формирование у обучающихся умений и навыков применения теоретических знаний в реальной практике решения задач; восполнение пробелов в пройденной теоретической части курса.

Семинарские, практические и лабораторные занятия в равной мере направлены на совершенствование индивидуальных навыков решения теоретических и прикладных задач, выработку навыков интеллектуальной работы, а также ведения дискуссий. Для успешного участия в семинарских, практических и лабораторных занятиях обучающемуся следует тщательно подготовиться.

Основной формой подготовки обучающихся к практическим (лабораторным) занятиям является самостоятельная работа с учебно-методическими материалами, научной литературой, статистическими данными и т.п.

Изучив конкретную тему, обучающийся может определить, насколько хорошо он в ней разобрался. Если какие-то моменты остались непонятными, целесообразно составить список вопросов и на занятии задать их преподавателю. Практические (лабораторные) занятия предоставляют обучающемуся возможность творчески раскрыться, проявить инициативу и развить навыки публичного ведения дискуссий и общения.

Самостоятельная работа обучающихся включает в себя выполнение различного рода заданий (изучение учебной и научной литературы, материалов лекций, систематизацию прочитанного материала, подготовку контрольной работы, решение

задач, подготовка докладов, написание рефератов, публикация тезисов, научных статей, подготовка и защита курсовой работы / проекта и другие), которые ориентированы на глубокое усвоение материала изучаемой дисциплины.

Обучающимся рекомендуется систематически отводить время для повторения пройденного материала, проверяя свои знания, умения и навыки.

Внутренняя система оценки качества освоения дисциплины включает входной контроль уровня подготовленности обучающихся, текущий контроль успеваемости, промежуточную аттестацию, направленную на оценивание промежуточных и окончательных результатов обучения по дисциплине (в том числе результатов курсового проектирования (выполнения курсовых работ) при наличии).

При проведении промежуточной аттестации обучающегося учитываются результаты текущего контроля, проводимого в течение освоения дисциплины.

Процедура оценивания результатов освоения дисциплины осуществляется на основе действующих локальных нормативных актов ФГБОУ ВО «Вятский государственный университет», с которыми обучающиеся знакомятся на официальном сайте университета www.vyatsu.ru.

Учебно-методическое обеспечение дисциплины, в том числе учебно-методическое обеспечение самостоятельной работы обучающегося по дисциплине

Учебная литература (основная)

1) Нестеров, С. А. Интеллектуальный анализ данных средствами MS SQL Server 2008 / С.А. Нестеров. - 2-е изд., испр. - Москва : Национальный Открытый Университет «ИНТУИТ», 2016. - 338 с. : ил. - Библиогр. в кн. - Б. ц. - URL: <http://biblioclub.ru/index.php?page=book&id=429083/> (дата обращения: 24.03.2020). - Режим доступа: ЭБС Университетская библиотека ONLINE. - Текст : электронный.

Учебная литература (дополнительная)

1) Захаров, В. П. Корпусная лингвистика : учебник / В.П. Захаров, С.Ю. Богданова. - Иркутск : Иркутский государственный лингвистический университет, 2011. - 161 с. - ISBN 978-5-88267-316-0 : Б. ц. - URL: <http://biblioclub.ru/index.php?page=book&id=89753/> (дата обращения: 24.03.2020). - Режим доступа: ЭБС Университетская библиотека ONLINE. - Текст : электронный.

2) Лукашевич, Наталья Валентиновна. Тезаурусы в задачах информационного поиска : научное издание / Н. В. Лукашевич. - М. : Изд-во Моск. ун-та, 2011. - 508 с. - Библиогр.: с. 483-508. - ISBN 978-5-211-05926-9 : 300.00 р. - Текст : непосредственный.

Учебно-методические издания

1) Шкаберина, Г. Ш. Программирование. Основы языка Python : учебное пособие / Г. Ш. Шкаберина, Н. Л. Резова. - Красноярск : СибГУ им. академика М. Ф. Решетнёва, 2018. - 92 с. - Б. ц. - URL: <https://e.lanbook.com/book/147450> (дата обращения: 15.05.2020). - Режим доступа: ЭБС Лань. - Текст : электронный.

2) Балджы, А. С. Математика на Python. 1 : учебно-методическое пособие / А.С. Балджы, М.Б. Хрипунова, И.А. Александрова. - Москва : Прометей, 2018. - 76 с. : табл. - Библиогр. в кн. - ISBN 978-5-907003-86-6 : Б. ц. - URL: <http://biblioclub.ru/index.php?page=book&id=494849/> (дата обращения: 24.03.2020). - Режим доступа: ЭБС Университетская библиотека ONLINE. - Текст : электронный.

Электронные образовательные ресурсы

1) Портал дистанционного обучения ВятГУ [электронный ресурс] / - Режим доступа: <http://mooc.do-kirov.ru/>

2) Раздел официального сайта ВятГУ, содержащий описание образовательной программы [электронный ресурс] / - Режим доступа: https://www.vyatsu.ru/php/programms/eduPrograms.php?Program_ID=3-09.04.03.03

- 3) Личный кабинет студента на официальном сайте ВятГУ [электронный ресурс] / -
Режим доступа: <https://new.vyatsu.ru/account/>
- 4) Единое окно доступа к образовательным ресурсам <http://window.edu.ru/>

Электронные библиотечные системы (ЭБС)

- ЭБС «Научная электронная библиотека eLIBRARY» (<http://elibrary.ru/defaultx.asp>)
- ЭБС «Издательства Лань» (<http://e.lanbook.com/>)
- ЭБС «Университетская библиотека online» (www.biblioclub.ru)
- Внутренняя электронно-библиотечная система ВятГУ (<http://lib.vyatsu.ru/>)
- ЭБС «ЮРАЙТ» (<https://urait.ru>)

Современные профессиональные базы данных и информационные справочные системы

- ГАРАНТ
- КонсультантПлюс
- Техэксперт: Нормы, правила, стандарты
- Роспатент (<https://www1.fips.ru/elektronnye-servisy/informatsionno-poiskovaya-sistema>)
- Web of Science® (<http://webofscience.com>)

Материально-техническое обеспечение дисциплины

Демонстрационное оборудование

Перечень используемого оборудования
ПРОЕКТОР NEC V302H FULL 3D

Специализированное оборудование

Перечень используемого оборудования
ГРАФИЧЕСКАЯ РАБОЧАЯ СТАНЦИЯ DEPO Race G560S
ПЕРСОНАЛЬНЫЙ КОМПЬЮТЕР TP CORP OPTIMA
ПЕРСОНАЛЬНЫЙ КОМПЬЮТЕР В СБОРЕ

Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине, в том числе лицензионное и свободно распространяемое ПО (включая ПО отечественного производства)

№ п.п	Наименование ПО	Краткая характеристика назначения ПО
1	Программная система с модулями для обнаружения текстовых заимствований в учебных и научных работах «Антиплагиат.ВУЗ»	Программный комплекс для проверки текстов на предмет заимствования из Интернет-источников, в коллекции диссертация и авторефератов Российской государственной библиотеки (РГБ) и коллекции нормативно-правовой документации LEXPRO
2	Microsoft Office 365 ProPlusEdu ALNG SubsVL MVL AddOn toOPP	Набор веб-сервисов, предоставляющий доступ к различным программам и услугам на основе платформы Microsoft Office, электронной почте бизнес-класса, функционалу для общения и управления документами
3	Office Professional Plus 2016	Пакет приложений для работы с различными типами документов: текстами, электронными таблицами, базами данных, презентациями
4	Windows Professional	Операционная система
5	Kaspersky Endpoint Security для бизнеса	Антивирусное программное обеспечение
6	Справочная правовая система «Консультант Плюс»	Справочно-правовая система по законодательству Российской Федерации
7	Электронный периодический справочник ГАРАНТ Аналитик	Справочно-правовая система по законодательству Российской Федерации
8	Security Essentials (Защитник Windows)	Защита в режиме реального времени от шпионского программного обеспечения, вирусов.
9	МойОфис Стандартный	Набор приложений для работы с документами, почтой, календарями и контактами на компьютерах и веб браузерах
10	Python	Язык программирования
11	Anaconda	дистрибутив языков программирования Python и R с набором приложений. По умолчанию в Anaconda Navigator доступны следующие приложения: JupyterLab Jupyter Notebook QtConsole Spyder Glue Orange RStudio Visual Studio Code

Обновленный список программного обеспечения данной рабочей программы находится по адресу:
https://www.vyatsu.ru/php/list_it/index.php?op_id=113619