

Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования



**Пермский национальный исследовательский
политехнический университет**

УТВЕРЖДАЮ

Проректор по учебной работе


_____ Н.В.Лобов

« 20 » января 20__ г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Дисциплина: Обработка квазиструктурированных данных и распределенных систем
(наименование)

Форма обучения: очная
(очная/очно-заочная/заочная)

Уровень высшего образования: магистратура
(бакалавриат/специалитет/магистратура)

Общая трудоёмкость: 144 (4)
(часы (ЗЕ))

Направление подготовки: 09.04.04 Программная инженерия
(код и наименование направления)

Направленность: Разработка программно-информационных систем
(наименование образовательной программы)

1. Общие положения

1.1. Цели и задачи дисциплины

Формирование комплекса знаний, умений и навыков в области применения современных алгоритмических языков программирования для разработки программного обеспечения, научных вычислений, машинного обучения.

1.2. Изучаемые объекты дисциплины

Алгоритмические языки программирования;
Язык программирования Python;
СУБД MongoDB;
Средства Python для взаимодействия с MongoDB;
Модули и библиотеки для извлечения данных.

1.3. Входные требования

Не предусмотрены

2. Планируемые результаты обучения по дисциплине

Компетенция	Индекс индикатора	Планируемые результаты обучения по дисциплине (знать, уметь, владеть)	Индикатор достижения компетенции, с которым соотнесены планируемые результаты обучения	Средства оценки
ПК-2.2	ИД-1ПК-2.2	Знает правила редактирования научно-технической документации;	Знает правила редактирования научно-технической документации;	Индивидуальное задание
ПК-2.2	ИД-2ПК-2.2	Умеет применять нормативно-технические документы (стандарты и регламенты), определяющие требования к проектной и технической документации;	Умеет применять нормативно-технические документы (стандарты и регламенты), определяющие требования к проектной и технической документации;	Отчет по практике
ПК-2.2	ИД-3ПК-2.2	Владеет навыками контроля и оценки качества разработанной проектной и технической документации.	Владеет навыками контроля и оценки качества разработанной проектной и технической документации.	Защита лабораторной работы

3. Объем и виды учебной работы

Вид учебной работы	Всего часов	Распределение по семестрам в часах	
		Номер семестра	
		3	
1. Проведение учебных занятий (включая проведение текущего контроля успеваемости) в форме:	54	54	
1.1. Контактная аудиторная работа, из них:			
- лекции (Л)	18	18	
- лабораторные работы (ЛР)	18	18	
- практические занятия, семинары и (или) другие виды занятий семинарского типа (ПЗ)	16	16	
- контроль самостоятельной работы (КСР)	2	2	
- контрольная работа			
1.2. Самостоятельная работа студентов (СРС)	90	90	
2. Промежуточная аттестация			
Экзамен			
Дифференцированный зачет	9	9	
Зачет			
Курсовой проект (КП)			
Курсовая работа (КР)			
Общая трудоемкость дисциплины	144	144	

4. Содержание дисциплины

Наименование разделов дисциплины с кратким содержанием	Объем аудиторных занятий по видам в часах			Объем внеаудиторных занятий по видам в часах
	Л	ЛР	ПЗ	СРС
3-й семестр				
Понятие квазиструктурированных данных.	2	2	2	0
Понятие квазиструктурированных данных. Отличия неструктурированных, структурированных и квазиструктурированных данных.				
Обзор средств и методов извлечения данных из неструктурированных источников	4	4	2	30
Обзор средств и методов извлечения данных из неструктурированных источников Библиотеки BeautifulSoup, rutmextract, Natasha				
Описание предметной области хэш-таблицами на примере словарей Python.	2	2	2	30
Описание предметной области хэш-таблицами на примере словарей Python. Класс dict, расширения dict. Пользовательские классы.				

Наименование разделов дисциплины с кратким содержанием	Объем аудиторных занятий по видам в часах			Объем внеаудиторных занятий по видам в часах
	Л	ЛР	ПЗ	СРС
Подходы к написанию парсеров. Грамматики. Машина состояний.	2	2	4	0
Подходы к написанию парсеров. Грамматики. Машина состояний. Разбор примеров.				
Способы хранения квазиструктурированных данных.	4	4	2	0
Способы хранения квазиструктурированных данных. Отличия реляционной и документоориентированной моделей. MongoDB. Установка и настройка MongoDB. Язык запросов. Интеграция с Python.				
Приложения, управляемые данными. Распределенная обработка запросов.	4	4	4	30
Приложения, управляемые данными. Распределенная обработка запросов. Применение распределенной обработки квазиструктурированных данных в современных мобильных приложениях.				
ИТОГО по 3-му семестру	18	18	16	90
ИТОГО по дисциплине	18	18	16	90

Тематика примерных практических занятий

№ п.п.	Наименование темы практического (семинарского) занятия
1	Разработка алгоритма программы извлечения данных из lib.ru.
2	Методы и средства извлечения именованных сущностей из текста.

Тематика примерных лабораторных работ

№ п.п.	Наименование темы лабораторной работы
1	Извлечение именованных сущностей из текста.
2	Разработка программы извлечения данных из lib.ru.
3	Обработка результатов поиска в нескольких системах и интеграция данных.

5. Организационно-педагогические условия

5.1. Образовательные технологии, используемые для формирования компетенций

Проведение лекционных занятий по дисциплине основывается на активном методе обучения, при которой учащиеся не пассивные слушатели, а активные участники занятия, отвечающие на вопросы преподавателя. Вопросы преподавателя нацелены на активизацию процессов усвоения материала, а также на развитие логического мышления. Преподаватель заранее намечает список вопросов, стимулирующих ассоциативное мышление и установления связей с ранее освоенным материалом.

Практические занятия проводятся на основе реализации метода обучения действием: определяются проблемные области, формируются группы. При проведении практических занятий преследуются следующие цели: применение знаний отдельных дисциплин и креативных методов для решения проблем и принятия решений; отработка у обучающихся навыков командной работы, межличностных коммуникаций и развитие лидерских качеств; закрепление основ теоретических знаний.

Проведение лабораторных занятий основывается на интерактивном методе обучения, при котором обучающиеся взаимодействуют не только с преподавателем, но и друг с другом. При этом доминирует активность учащихся в процессе обучения. Место преподавателя в интерактивных занятиях сводится к направлению деятельности обучающихся на достижение целей занятия.

При проведении учебных занятий используются интерактивные лекции, групповые дискуссии, ролевые игры, тренинги и анализ ситуаций и имитационных моделей.

5.2. Методические указания для обучающихся по изучению дисциплины

При изучении дисциплины обучающимся целесообразно выполнять следующие рекомендации:

1. Изучение учебной дисциплины должно вестись систематически.
2. После изучения какого-либо раздела по учебнику или конспектным материалам рекомендуется по памяти воспроизвести основные термины, определения, понятия раздела.
3. Особое внимание следует уделить выполнению отчетов по практическим занятиям, лабораторным работам и индивидуальным комплексным заданиям на самостоятельную работу.
4. Вся тематика вопросов, изучаемых самостоятельно, задается на лекциях преподавателем. Им же даются источники (в первую очередь вновь изданные в периодической научной литературе) для более детального понимания вопросов, озвученных на лекции.

6. Перечень учебно-методического и информационного обеспечения для самостоятельной работы обучающихся по дисциплине

6.1. Печатная учебно-методическая литература

№ п/п	Библиографическое описание (автор, заглавие, вид издания, место, издательство, год издания, количество страниц)	Количество экземпляров в библиотеке
1. Основная литература		
1	Прохоренок Н. А. Python 3 и PyQt. Разработка приложений / Н. А. Прохоренок. - Санкт-Петербург: БХВ-Петербург, 2014.	5
2	Сузи Р. А. Язык программирования PYTHON : учебное пособие / Р. А. Сузи. - Москва: ИНТУИТ, БИНОМ. Лаб. знаний, 2006.	3
2. Дополнительная литература		
2.1. Учебные и научные издания		

1	АJAX и PHP : разработка динамических веб-приложений : пер. / К. Дари, Б. Бринзаре, Ф. Черчез-Тоза, М. Бусика. - СПб. М.: Символ, 2007.	2
2	Кузьмин И.А. Распределенная обработка информации в научных исследованиях / И.А.Кузьмин, В.А.Путилов, В.В.Фильчаков. - Л.: Наука, 1991.	2
3	Лутц М. Программирование на Python : пер. с англ. / М. Лутц. - Санкт-Петербург: Символ, 2002.	3
4	Свистунов А. Н. Построение распределенных программных систем на Java : учебное пособие / А. Н. Свистунов. - Москва: ИНТУИТ, БИНОМ. Лаб. знаний, 2011.	2
5	Сузи Р. А. Язык программирования PYTHON : учебное пособие / Р. А. Сузи. - Москва: ИНТУИТ, БИНОМ. Лаб. знаний, 2007.	3
2.2. Периодические издания		
	Не используется	
2.3. Нормативно-технические издания		
	Не используется	
3. Методические указания для студентов по освоению дисциплины		
1	Арсеньев Б. П. Интеграция распределенных баз данных / Б. П. Арсеньев, С. А. Яковлев. - Санкт-Петербург: Лань, 2001.	13
2	Свистунов А. Н. Построение распределенных программных систем на Java : учебное пособие / А. Н. Свистунов. - Москва: ИНТУИТ, БИНОМ. Лаб. знаний, 2011.	2
4. Учебно-методическое обеспечение самостоятельной работы студента		
	Не используется	

6.2. Электронная учебно-методическая литература

Вид литературы	Наименование разработки	Ссылка на информационный ресурс	Доступность (сеть Интернет / локальная сеть; авторизованный / свободный доступ)
Дополнительная литература	beautifulsoup4	https://pypi.org/project/beautifulsoup4/	сеть Интернет; свободный доступ
Учебно-методическое обеспечение самостоятельной работы студентов	Wiki Портала-Python программистов: Документации/BeautifulSoup	https://tinyurl.com/beautifulsoup-bs4	сеть Интернет; свободный доступ

6.3. Лицензионное и свободно распространяемое программное обеспечение, используемое при осуществлении образовательного процесса по дисциплине

Вид ПО	Наименование ПО
Операционные системы	Debian (GNU GPL)
Среды разработки, тестирования и отладки	NetBeans (SUN PUBLIC LICENSE)

Вид ПО	Наименование ПО
Среды разработки, тестирования и отладки	PIP (The Python Package Installer) Free

6.4. Современные профессиональные базы данных и информационные справочные системы, используемые при осуществлении образовательного процесса по дисциплине

Наименование	Ссылка на информационный ресурс
Научная библиотека Пермского национального исследовательского политехнического университета	http://lib.pstu.ru/
Электронно-библиотечная система Лань	https://e.lanbook.com/
Электронно-библиотечная система Лань	https://e.lanbook.com/
Электронно-библиотечная система IPRbooks	http://www.iprbookshop.ru/
Информационные ресурсы Сети КонсультантПлюс	http://www.consultant.ru/

7. Материально-техническое обеспечение образовательного процесса по дисциплине

Вид занятий	Наименование необходимого основного оборудования и технических средств обучения	Количество единиц
Лабораторная работа	Компьютер с подключением к сети Интернет	15
Лекция	Проектор	1
Практическое занятие	Ноутбук	1

8. Фонд оценочных средств дисциплины

Описан в отдельном документе
