

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РФ
филиал федерального государственного бюджетного образовательного учреждения
высшего образования «Мурманский арктический государственный университет»
в г. Апатиты

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)

Б1.О.25 Технологии обработки информации

(название дисциплины (модуля) в соответствии с учебным планом)

**образовательной программы
по направлению подготовки бакалавриата**

**09.03.02 Информационные системы и технологии
направленность (профиль): «Программно-аппаратные комплексы»**

(код и наименование направления подготовки
с указанием направленности (профиля) (наименования магистерской программы))

высшее образование – бакалавриат

уровень профессионального образования: высшее образование – бакалавриат / высшее образование –
специалитет, магистратура / высшее образование – подготовка кадров высшей квалификации

бакалавр

квалификация

заочная

форма обучения

2019

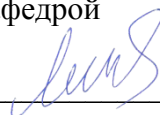
год набора

Составитель:

Тоичкин Н.А., к.т.н.,
доцент кафедры
информатики и вычислительной
техники

Утверждено на заседании кафедры
информатики и вычислительной техники
(протокол № 13 от «06» июня 2019 г.)

Зав. кафедрой



Яковлев С.Ю.

Структура рабочей программы дисциплины

1. ЦЕЛЬ ОСВОЕНИЯ ДИСЦИПЛИНЫ (МОДУЛЯ)

Целью изучения дисциплины «Технологии обработки информации» является: формирование целостной системы знаний в области создания, накопления и обработки информации, а также освоение обучающимися эффективными технологиями и методами ее обработки.

знать:

- основные виды и процедуры обработки информации;
- модели и методы решения задач обработки информации;
- современные средства хранения данных;
- принципы создания хранилищ и витрин данных;
- принципы очистки данных;
- понятие больших данных и их свойства;

уметь:

- осуществлять математическую и информационную постановку задач по обработке информации;
- использовать алгоритмы обработки информации для различных приложений;
- использовать методы оперативной аналитической обработки информации;
- выполнять обработку данных в современных ETL системах;
- применять на практике алгоритмы обработки данных.

владеть:

- методами и средствами для обработки информации;
- современными инструментальными средствами обработки информации;
- информационными технологиями поиска данных и способами их использования;
- методами навыками и фильтрации данных;
- базовыми методами обработки данных в языке программирования Python.

2. ПЕРЕЧЕНЬ ПЛАНИРУЕМЫХ РЕЗУЛЬТАТОВ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ (МОДУЛЮ), СООТНЕСЕННЫХ С ПЛАНИРУЕМЫМИ РЕЗУЛЬТАТАМИ ОСВОЕНИЯ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ

В результате освоения содержания дисциплины «Информатика» обучающийся должен обладать следующими компетенциями:

- способен использовать современные информационные технологии и программные средства, в том числе отечественного производства, при решении задач профессиональной деятельности (ОПК-2);
- способен применять математические модели, методы и средства проектирования информационных и автоматизированных систем (ОПК-8).

3. УКАЗАНИЕ МЕСТА ДИСЦИПЛИНЫ (МОДУЛЯ) В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ

Данная дисциплина относится к дисциплинам обязательной части блока Б1.О в структуре образовательной программы бакалавриата по направлению подготовки 09.03.02 Информационные системы и технологии направленность (профиль) Программно-аппаратные комплексы.

Для освоения данной дисциплины обучающиеся используют знания, умения, навыки, способы деятельности и установки, которые они получили в процессе изучения дисциплин: «Информатика», «Программирование»

В свою очередь, «Технологии обработки информации» представляет собой методологическую базу для дисциплин: «Интеллектуальные системы и технологии»,

«Анализ данных и машинное обучение».

4. ОБЪЕМ ДИСЦИПЛИНЫ (МОДУЛЯ) В ЗАЧЕТНЫХ ЕДИНИЦАХ С УКАЗАНИЕМ КОЛИЧЕСТВА АКАДЕМИЧЕСКИХ ИЛИ АСТРОНОМИЧЕСКИХ ЧАСОВ, ВЫДЕЛЕННЫХ НА КОНТАКТНУЮ РАБОТУ ОБУЧАЮЩИХСЯ С ПРЕПОДАВАТЕЛЕМ (ПО ВИДАМ УЧЕБНЫХ ЗАНЯТИЙ) И НА САМОСТОЯТЕЛЬНУЮ РАБОТУ ОБУЧАЮЩИХСЯ

Общая трудоемкость дисциплины составляет 5 зачетных единиц или 180 часов.
(из расчета 1 ЗЕТ= 36 часов).

Курс	Семестр	Трудоемкость в ЗЕТ	Общая трудоемкость (час.)	Контактная работа			Всего контактных часов	Из них в интер-активной форме	Кол-во часов на СРС	Курсовые работы	Кол-во часов на контроль	Форма контроля
				ЛК	ПР	ЛБ						
3	5	1	36	6	-	8	14	6	22	-	-	-
3	6	3	108	6	-	8	14	6	85	-	9	экзамен
4	7	1	36	-	-	-	-	-	27	-	9	экзамен
Итого:		5	180	12	-	16	28	12	134	-	18	экзамен

В интерактивной форме часы используются в виде: групповой дискуссии.

5. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ, СТРУКТУРИРОВАННОЕ ПО ТЕМАМ (РАЗДЕЛАМ) С УКАЗАНИЕМ ОТВЕДЕННОГО НА НИХ КОЛИЧЕСТВА АКАДЕМИЧЕСКИХ ИЛИ АСТРОНОМИЧЕСКИХ ЧАСОВ И ВИДОВ УЧЕБНЫХ ЗАНЯТИЙ

№ п/п	Наименование раздела, темы	Контактная работа			Всего контактных часов	Из них в интерактивной форме	Кол-во часов на СРС	Кол-во часов на контроль
		ЛК	ПР	ЛБ				
1	Введение в предмет: «Технологии обработки информации».	1	-	-	1	-	4	-
2	Базовые технологии обработки информации.	1	-	-	1	-	4	-
3	Прикладные инструменты для анализа информации.	1	-	-	1	1	8	
4	Принцип организации обработки данных в системе Loginom,	1	-	2	3			

	основанный на сценариях.							
5	Стандартизация и очистка информации.	1	-	-	1	-	8	
6	Регулярные выражения.	1	-	2	3	2	8	-
7	Основы хранилищ и витрин данных. Многомерная модель данных.	1	-	-	4	-	8	-
8	Многомерные и реляционные модели данных.	1	-	4	5	4	6	-
9	Витрины данных.	1	-	-	1	-	2	-
10	Загрузка данных в хранилище.	1	-	-	1	-	2	-
11	Метод Data Vault.	1	-	2	3	2	6	-
12	Программирование на языке Python для сбора и анализа данных.	1	-	6	7	3	6	-
	Зачет 1 семестр	-	-	-	-	-	-	-
	Экзамен 2 семестр	-	-	-	-	-	-	36
	Итого:	12	-	16	28	12	62	36

Содержание дисциплины:

Тема 1. Введение в предмет: «Технологии обработки информации». Основные понятия – информация, данные, метаинформация, знания. Виды информации. Обработка данных и ее виды. Модели процессов обработки данных. Задачи обработки данных.

Тема 2. Базовые технологии обработки информации. Создание, модификация данных, контроль, безопасность и целостность данных, поиск информации, поддержка принятия решения, создание документов, сводок, отчетов, преобразование информации. Задачи обработки данных различных типов. Прикладные области обработки данных. Сжатие информации, алгоритмы сжатия.

Тема 3. Прикладные инструменты для анализа информации. Исторический обзор инструментов бизнес-аналитики для анализа информации. Аналитическая платформа Loginom и ее редакции. Использование структурного и объектно-ориентированного подхода к моделированию в Loginom.

Тема 4. Принцип организации обработки данных в системе Loginom основанный на сценариях. Основы работы с Loginom: работа с пакетами, сценарии, компоненты и узлы, модификаторы доступа, настройки портов и автосинхронизация, компонент калькулятор, переменные и параметризация узлов, компоненты условие и замена. Компонент цикл, наследование и производные компоненты, импорт из промышленных источников данных.

Тема 5. Стандартизация и очистка информации. Причины загрязнения данных и виды ошибок. Методы очистки данных: использование словарей, и таблиц замены, анализ строк, регулярные выражения, частотный анализ, контрольные числа. Объединение различных методов очистки данных. Общий алгоритм очистки.

Тема 6. Регулярные выражения. История появления регулярных выражений. Диалекты и возможности регулярных выражений. Синтаксис регулярных выражений, языка Perl: метасимволы, квантификаторы, примеры использования.

Тема 7. Основы хранилищ и витрин данных. Основные положения концепции ХД (хранилищ данных). Определение и свойства ХД. Детализированные и агрегированные данные, метаданные, виды метаданных. Краткий обзор моделей ХД. Многомерное представление данных. Базовые понятия многомерной модели данных. Структура многомерного куба. Работа с измерениями.

Тема 9. Многомерные и реляционные модели данных. MOLAP и ROLAP. Агрегаты и нерегламентированные запросы. Пример MOLAP. Реляционные ХД ROLAP, схема «звезда» и «созвездие». Схема «созвездие». Сравнение MOLAP и ROLAP. Гибридная модель HOLAP.

Тема 10. Витрины данных. Концепция витрин данных. Проектирование витрины данных. Пример проектирование витрины данных.

Тема 11. Загрузка данных в хранилище. Поддержка историчности. Медленно меняющиеся измерения. Неполная загрузка данных. Многопоточная организация процесса загрузки данных. Постзагрузочные операции.

Тема 12. Метод Data Vault. Недостатки существующих моделей ХД. Основные концепции и определение Data Vault. Эволюция от нормальных форм к Data Vault. Основные компоненты модели Data Vault: концентраторы, ссылки, спутники. Пример модели Data Vault. Обобщенная схема хранилища Data Vault. Преимущества и недостатки Data Vault.

Тема 13. Программирование на языке Python для сбора и анализа данных. Типы данных в Python: строки, списки, кортежи, словари, множества. Функции, их определение. Элементы функционального программирования: list comprehensions, лямбда-функции, рекурсия. Элементы объектно-ориентированного программирования.

6. ПЕРЕЧЕНЬ ОСНОВНОЙ И ДОПОЛНИТЕЛЬНОЙ УЧЕБНОЙ ЛИТЕРАТУРЫ, НЕОБХОДИМОЙ ДЛЯ ОСВОЕНИЯ ДИСЦИПЛИНЫ (МОДУЛЯ)

Основная литература:

1. Вдовин В. М., Суркова Л. Е., Шурупов А. А. Предметно-ориентированные экономические информационные системы: учебное пособие. Москва: Издательско-торговая корпорация «Дашков и К°», 2016, 386 стр. Режим доступа: https://biblioclub.ru/index.php?page=book_red&id=453951&sr=1
2. Архипенков С. Я., Голубев Д., Максименко О. Хранилища данных: от концепции до внедрения. М.: Диалог-МИФИ, 2002, 528 с. Режим доступа: https://biblioclub.ru/index.php?page=book_red&id=89285&sr=1

Дополнительная литература:

3. Васильев А.Н. Python на примерах. Практический курс по программированию. - СПб.: Наука и Техника, 2016. - 432 с.: ил.
4. Уэс Маккинли Python и анализ данных/ Пер. с англ. Слинкин А. А. - М.: ДМК Пресс, 2015. - 482 с.: ил.

7. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

В образовательном процессе используются:

- учебные аудитории для проведения занятий лекционного типа, занятий семинарского типа, курсового проектирования (выполнения курсовых работ), групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации (мебель аудиторная (столы, стулья, доска аудиторная), комплект мультимедийного оборудования, включающий мультимедиапроектор, экран, переносной ноутбук для демонстрации презентаций; учебно-наглядные пособия; обеспечивающие тематические иллюстрации);
- помещения для самостоятельной работы (оснащены компьютерными столами, стульями, доской аудиторной, компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду университета);
- помещения для хранения и профилактического обслуживания учебного оборудования (оснащены наборами инструментов, оборудованием, расходными материалами для монтажа, ремонта и обслуживания информационно-телекоммуникационной сети филиала и вычислительной техники);

- лаборатория информационных технологий (оснащена компьютерными столами, стульями, мультимедийным проектором, экраном проекционным, компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду университета).

7.1 ПЕРЕЧЕНЬ ЛИЦЕНЗИОННОГО ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ

1. Microsoft Windows XP – операционная система.
2. Loginom Studio Standard – аналитическая платформа.
3. DBeaver - платформенно-независимый клиент баз данных.
4. Vertica Community Edition - аналитическая платформа для анализа больших данных.
5. Любой web – браузер.

7.2 ЭЛЕКТРОННО-БИБЛИОТЕЧНЫЕ СИСТЕМЫ

1. ЭБС «Издательство Лань»[Электронный ресурс]: электронная библиотечная система / ООО «Издательство Лань». – Режим доступа: <https://e.lanbook.com/>;
2. ЭБС «Электронная библиотечная система ЮРАЙТ» [Электронный ресурс]: электронная библиотечная система / ООО «Электронное издательство ЮРАЙТ». – Режим доступа: <https://biblio-online.ru/>;
3. ЭБС «Университетская библиотека онлайн»[Электронный ресурс]: электронно-периодическое издание; программный комплекс для организации онлайн-доступа к лицензионным материалам / ООО «НексМедиа». – Режим доступа: <https://biblioclub.ru/>.

7.3 СОВРЕМЕННЫЕ ПРОФЕССИОНАЛЬНЫЕ БАЗЫ ДАННЫХ

1. Электронная база данных SCOPUS.
2. Электронная база данных РИНЦ.

7.4 ИНФОРМАЦИОННЫЕ СПРАВОЧНЫЕ СИСТЕМЫ

1. Справочно-правовая информационная система Консультант Плюс <http://www.consultant.ru/>
2. Электронный справочник "Информо" для высших учебных заведений <http://www.informio.ru/>

8. ИНЫЕ СВЕДЕНИЯ И МАТЕРИАЛЫ НА УСМОТРЕНИЕ ВЕДУЩЕЙ КАФЕДРЫ

Не предусмотрено.

9. ОБЕСПЕЧЕНИЕ ОБРАЗОВАНИЯ ДЛЯ ЛИЦ С ОВЗ

Для обеспечения образования инвалидов и лиц с ограниченными возможностями здоровья реализация дисциплины может осуществляться в адаптированном виде, с учетом специфики освоения и дидактических требований, исходя из индивидуальных возможностей и по личному заявлению обучающегося.