

Минобрнауки России

Федеральное государственное бюджетное образовательное учреждение
высшего образования

«Оренбургский государственный университет»

Кафедра геометрии и компьютерных наук

РАБОЧАЯ ПРОГРАММА

ДИСЦИПЛИНЫ

«Б.1.В.ДВ.5.2 Хранилища и аналитическая обработка данных»

Уровень высшего образования

БАКАЛАВРИАТ

Направление подготовки

02.03.02 Фундаментальная информатика и информационные технологии
(код и наименование направления подготовки)

Общий профиль

(наименование направленности (профиля) образовательной программы)

Тип образовательной программы

Программа академического бакалавриата

Квалификация

Бакалавр

Форма обучения

Очная

Год набора 2016

1060392

Рабочая программа рассмотрена и утверждена на заседании кафедры

Кафедра геометрии и компьютерных наук

наименование кафедры

протокол № 8 от "25" февраля 2016 г.

Заведующий кафедрой

Кафедра геометрии и компьютерных наук

наименование кафедры

подпись

А.Е. Шухман

расшифровка подписи



Исполнители:

Старший преподаватель кафедры ГКН

должность

подпись



И.В. Минина

расшифровка подписи

должность

подпись

расшифровка подписи

СОГЛАСОВАНО:

Председатель методической комиссии по направлению подготовки

02.03.02 Фундаментальная информатика и информационные технологии

код наименование



личная подпись

А.Е. Шухман

расшифровка подписи

Заведующий отделом комплектования научной библиотеки

личная подпись

Н.Н. Грицай

расшифровка подписи

Уполномоченный по качеству факультета

личная подпись

И.В. Крючкова

расшифровка подписи

№ регистрации _____

1 Цели и задачи освоения дисциплины

Цель (цели) освоения дисциплины: дать систематический обзор методов и принципов организации и оперирования большими объемами данных с применением современных информационных средств и технологий.

Задачи:

- изучение существующих технологий подготовки данных к анализу;
- формирование умений и навыков применения универсальных программных пакетов и аналитических платформ для построения хранилищ данных в целях их последующего анализа.

2 Место дисциплины в структуре образовательной программы

Дисциплина относится к дисциплинам (модулям) по выбору вариативной части блока 1 «Дисциплины (модули)»

Пререквизиты дисциплины: *Б.1.Б.18 Дискретная математика, Б.1.Б.23 Технологии баз данных, Б.1.В.ОД.16 Теория автоматов и формальных языков, Б.1.В.ОД.17 Интеллектуальные системы*

Постреквизиты дисциплины: *Отсутствуют*

3 Требования к результатам обучения по дисциплине

Процесс изучения дисциплины направлен на формирование следующих результатов обучения

Планируемые результаты обучения по дисциплине, характеризующие этапы формирования компетенций	Формируемые компетенции
<p>Знать: - теоретические основы проектирования информационно-аналитических систем как консолидирующего средства для создания интегрированной корпоративной информационной системы; основы построения систем поддержки принятия решений, реинжиниринга бизнес-процессов и бизнесархитектуры предприятия.</p> <p>Уметь: - определять необходимость применения технологий интеллектуального анализа данных.</p> <p>Владеть: - современными информационными технологиями для организации аналитической обработки данных</p>	ОК-7 способностью к самоорганизации самообразованию
<p>Знать: - основные принципы построения систем, ориентированных на анализ данных; - современные технологии интеллектуального анализа данных; - модели данных, используемые для построения хранилищ; - особенности построения систем на основе хранилищ данных;</p> <p>Уметь: - создавать архитектуру информационно-аналитической системы;</p> <p>Владеть: - технологиями анализа данных.</p>	ПК-2 способностью понимать, совершенствовать и применять современный математический аппарат, фундаментальные концепции и системные методологии, международные и профессиональные стандарты в области информационных технологий
<p>Знать: - концепцию построения хранилищ и витрин данных; - концепцию построения многомерных моделей данных и их обработки, а также концептуальное многомерное представление;</p>	ПК-6 способностью эффективно применять базовые математические знания и информационные

Планируемые результаты обучения по дисциплине, характеризующие этапы формирования компетенций	Формируемые компетенции
- архитектуру OLAP-систем и решаемый класс задач; – основы интеллектуального анализа данных (Data Mining). Уметь: - проектировать структуры хранилищ данных; Владеть: -инструментарием построения хранилищ и витрин данных	технологии при решении проектно-технических и прикладных задач, связанных с развитием и использованием информационных технологий
Знать: - основные определения, относящиеся к концепции управления хранилищами данных; - основные требования (и средства их обеспечения) к хранилищам данных, технологии, обеспечивающие манипулирование хранилищами данных; Уметь: - проектировать многомерные кубы данных; Владеть: - технологиями, обеспечивающими манипулирование хранилищами данных.	ПК-7 способностью разрабатывать и реализовывать процессы жизненного цикла информационных систем, программного обеспечения, сервисов систем информационных технологий, а также методы и механизмы оценки и анализа функционирования средств и систем информационных технологий
Знать: - назначение и область применения методов интеллектуального анализа данных; Уметь: - проектировать системы: загрузки данных в информационные хранилища, обработки запросов и представления результатов анализа, взаимодействия с администраторами ИАС. Владеть: - технологиями и принципами системами сбора и повышения качества исходных данных для анализа и последующего принятия решений в структурах информационных хранилищ, комплексах инструментальных средств, поддерживающих технологии анализа данных	ПК-8 способностью применять на практике международные и профессиональные стандарты информационных технологий, современные парадигмы и методологии, инструментальные и вычислительные средства

4 Структура и содержание дисциплины

4.1 Структура дисциплины

Общая трудоемкость дисциплины составляет 4 зачетных единиц (144 академических часов).

Вид работы	Трудоемкость, академических часов	
	7 семестр	всего
Общая трудоёмкость	144	144
Контактная работа:	35,25	35,25
Лекции (Л)	18	18
Лабораторные работы (ЛР)	16	16
Консультации	1	1
Промежуточная аттестация (зачет, экзамен)	0,25	0,25
Самостоятельная работа: - самоподготовка (проработка и повторение лекционного материала и материала учебников и учебных пособий); - подготовка к лабораторным занятиям; - подготовка к коллоквиумам;	108,75	108,75

Вид работы	Трудоемкость, академических часов	
	7 семестр	всего
- подготовка к рубежному контролю и т.п.)		
Вид итогового контроля (зачет, экзамен, дифференцированный зачет)	экзамен	

Разделы дисциплины, изучаемые в 7 семестре

№ раздела	Наименование разделов	Количество часов				
		всего	аудиторная работа			внеауд. работа
			Л	ПЗ	ЛР	
1	Введение в предмет	14	2		2	10
2	Общие свойства хранилищ данных	14	2		2	10
3	Архитектуры хранилищ данных	14	2		2	10
4	Многомерные кубы	32	2		10	20
5	Методология построения хранилищ данных	24	4			20
6	Выбор метода реализации хранилищ данных	12	2			10
7	Интеграция информационных ресурсов в хранилищах данных	12	2			10
8	Технология Data Mining	22	2			20
	Итого:	144	18		16	110
	Всего:	144	18		16	110

4.2 Содержание разделов дисциплины

1. Введение в предмет

Цель и задачи дисциплины, ее роль и место в общей системе подготовки бакалавров. Основные понятия. История вопроса. Понятие OLAP. Различия между транзакционными и аналитическими системами. Области применения хранилищ данных.

2. Общие свойства хранилищ данных

Введение в хранилища данных. Предпосылки создания концепции складирования данных. Основные положения концепции складирования данных. Назначение и основные компоненты хранилищ данных. Оперативный склад данных, витрина данных. Подсистемы загрузки данных, обработки запросов и представления данных. Подсистема администрирования хранилища. Классификация хранилищ данных по Инмону. Понятие жизненного цикла хранилища данных. Характеристика основных этапов жизненного цикла разработки хранилища данных

3. Архитектуры хранилищ данных

Уровни архитектуры: концептуальный, логический, физический. Типовые архитектуры хранилищ данных. Компоненты типовой архитектуры. Типовая обобщенная концептуальная схема архитектуры хранилища данных. Основные методологические подходы к реализации хранилища данных. Обзор решений основных производителей программного обеспечения для разработки хранилищ данных. Типовые технологические решения. Метаданные в хранилищах данных. Содержание термина «метаданные». Функции метаданных в хранилище данных. Логическая структура метаданных хранилища данных. Стандарты метаданных.

4. Многомерные кубы

Основные понятия кубов. Иерархии измерений. Структура ХД. Примеры кубов. Три способа хранения агрегатных данных.

5. Методология построения хранилищ данных

Подходы к стратегии построения. Модели разработки. Этапы спиральной модели применительно к разработке хранилищ данных. Компонентная архитектура. Техническая архитектура. Метод многомерного моделирования. Основные элементы многомерной модели: атрибуты, измерения, факты, гранулированность. Основные схемы данных многомерной модели. Свойства фактов. Медленно и быстро меняющиеся измерения, вырожденные измерения, иерархии измерений, множественные иерархии. Схемы «звезда» и «снежинка».

6. Выбор метода реализации хранилищ данных

Две группы аналитических платформ. Обзор рынка BI. Продукция Microsoft. Продукция Sybase. Продукция Oracle.

7. Интеграция информационных ресурсов в хранилищах данных

Проблема интеграции данных. Возможности SQL Server 2014 Integration Services. Планирование ETL проекта для хранилища данных.

8. Технология Data Mining

Общие понятия. История вопроса. Знакомство с методами интеллектуального анализа данных (Data Mining). Обзор основных алгоритмов интеллектуального анализа данных: алгоритмы кластеризации, алгоритмы классификации, алгоритм взаимосвязей, алгоритмы регрессионного анализа. Основные этапы процесса интеллектуального анализа данных. Приложения. Технология (процесс) добычи знаний. Решаемые задачи. Математические основы (РАД). Data Mining в MSSAS.

4.3 Лабораторные работы

№ ЛР	№ раздела	Наименование лабораторных работ	Кол-во часов
1	1	Развертывание служб Analysis Services	2
2	2,3	Определение представления источника данных в проекте служб Analysis Services	2
3	4	Определение и развертывание куба	2
4	4	Изменение мер, атрибутов и иерархий	2
5	4	Определение расширенных свойств атрибутов и измерений	2
6	4	Определение связей между измерениями и группами мер	2
7	4	Определение вычислений	4
		Итого:	16

5 Учебно-методическое обеспечение дисциплины

5.1 Основная литература

1. Туманов, В.Е. Проектирование хранилищ данных для систем бизнес-аналитики : учебное пособие / В.Е. Туманов. - М. : Интернет-Университет Информационных Технологий, 2010. - 616 с. : ил., табл., схем. - (Основы информационных технологий). - ISBN 978-5-9963-0353-3 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=233492> (07.04.2016).

5.2 Дополнительная литература

1. Туманов, В.Е. Проектирование реляционных хранилищ данных / В.Е. Туманов, С.В. Маклаков. - М. : Диалог-МИФИ, 2007. - 333 с. - ISBN 5864042102 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=54774> (07.04.2016).

2. Маклаков С.В. Создание информационных систем с AllFusion Modeling Suite. – М.: Издательство Диалог-МИФИ, 2007 – 400 с.

3. Полубояров В.В. Использование MS SQL Server 2008 Analysis Services для построения хранилищ данных. Интуит, 2010. – 487 с. То же [Электронный ресурс]. - URL: - <http://biblioclub.ru/index.php?page=book&id=234554>

5.3 Периодические издания

1. Информатика и системы управления : журнал. - М. : Агентство "Роспечать", 2016.
2. Информационные технологии в проектировании и производстве : журнал. - М. : Агентство "Роспечать", 2016.
3. Мир ПК : журнал. - М. : Агентство "Роспечать", 2016.
4. Программные продукты и системы : журнал. - М. : Агентство "Роспечать", 2016.

5.4 Интернет-ресурсы

1. <http://www.citforum.ru/> - портал аналитических и научных статей в области информационных технологий
2. <http://www.rsdn.ru> - сайт Российской сети разработчиков ПО, содержит статьи по современным средствам программирования.
3. <http://www.intuit.ru> – сайт Интернет-университета информационных технологий, представляет учебные курсы по разным областям ИТ.

5.5 Программное обеспечение, профессиональные базы данных и информационные справочные системы современных информационных технологий

1. Операционная система Microsoft Windows
2. Система программирования Microsoft Visual Studio 2015 Professional.
3. Система управления реляционными базами данных Microsoft SQL Server 2014 Standart

6 Материально-техническое обеспечение дисциплины

Учебные аудитории для проведения занятий лекционного типа, семинарского типа, для проведения групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации.

Аудитории оснащены комплектами ученической мебели, техническими средствами обучения, служащими для представления учебной информации большой аудитории.

Для проведения лабораторных занятий компьютерный класс, оснащенный компьютерной техникой, подключенной к сети "Интернет" с обеспечением доступа в электронную информационно-образовательную среду ОГУ.

К рабочей программе прилагаются:

- Фонд оценочных средств для проведения промежуточной аттестации обучающихся по дисциплине;
- Методические указания для обучающихся по освоению дисциплины.