

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
ФИЛИАЛ ФГБОУ ВО «ВЛАДИВОСТОКСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ЭКОНОМИКИ И СЕРВИСА» В Г. НАХОДКЕ
КАФЕДРА МЕНЕДЖМЕНТА И ЭКОНОМИКИ

ХРАНИЛИЩА ДАННЫХ

Рабочая программа дисциплины

по направлению подготовки

38.03.05 Бизнес-информатика

Рабочая программа дисциплины «Хранилища данных» составлена в соответствии с требованиями ФГОС ВО по направлению подготовки 38.03.05 «Бизнес-информатика», бакалавр и Порядком организации и осуществления образовательной деятельности по образовательным программам высшего образования – программам бакалавриата, программам специалитета, программам магистратуры (утв. приказом Минобрнауки России от 5 апреля 2017 г. № 301)

Составители:


Кригер А.Б. к.ф.-м.н., доцент кафедры информационных технологий и систем,

Витязев Г.Г. преподаватель кафедры менеджмента и экономики

Рассмотрена и утверждена на заседании кафедры менеджмента и экономики от «28» апреля 2018 года, протокол № 9

Заведующий кафедрой (разработчика)

«28» апреля 2018г.



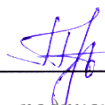
подпись

Просалова В.С.

фамилия, инициалы

Заведующий кафедрой (выпускающей)

«28» апреля 2018г.



подпись

Просалова В.С.

фамилия, инициалы

1 Цель и задачи освоения дисциплины (модуля)

Цели освоения учебной дисциплины:

- получение базовых знаний о системах хранения данных, особенностях Хранилищ данных и их назначении;
- формирование умений и навыков проектирования Хранилищ данных и систем бизнес-анализа;
- знакомство с технологиями интеллектуального анализа.

Задачи освоения дисциплины:

- Изучение принципов построения и разработки хранилищ данных;
- получение навыков настройки хранилищ данных;
- проектирование и разработка процесса наполнения Хранилища данных, реализации запросов к Хранилищам данных;

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

Планируемыми результатами обучения по дисциплине (модулю), являются знания, умения, навыки, характеризующие этапы/уровни формирования компетенций и обеспечивающие достижение планируемых результатов освоения образовательной программы в целом. Перечень компетенций, формируемых в результате изучения дисциплины, приведен в таблице 1.

Таблица 1 – Формируемые компетенции

| ОПОП | Коды компетенций | Название компетенции | Составляющие компетенции | |
|---------------------------------|------------------|--|--------------------------|--|
| | | | | |
| 38.03.05, Бизнес-информатика | ПК-2 | Проведение исследования и анализа рынка информационных систем и информационно-коммуникативных технологий | Знания: | архитектур хранилищ данных |
| | | | Умение: | выбирать систему хранения данных, соответствующую задачам профессиональной деятельности |
| | | | Навыки | настройки пользовательских инструментов промышленных хранилищ данных |
| | ПК-10 | Умения позиционировать электронное предприятие на глобальном рынке; формировать потребительскую аудиторию и осуществлять взаимодействие с потребителями. Организовывать продажи в информационно-телекоммуникационной сети «Интернет» | Знания: | технологий хранения данных (складирования) |
| | | | Умения: | применять OLAP-технологию для анализа показателей электронной коммерции |
| | | | Навыки | разработки логических моделей хранилищ данных |
| | ПК-16 | Умения разрабатывать контент и ИТ-сервисы предприятия и интернет-ресурсов | Знания: | концепции хранилищ данных |
| | | | Умения: | применять различные методы интеллектуального анализа данных для решения различных прикладных задач в |

| | | | | |
|--|-------|---|---------|---|
| | | | | профессиональной деятельности |
| | | | Навыки | разработки хранилищ данных |
| | ПК-18 | Способность использовать математический аппарат для обработки, анализа и систематизации информации по теме исследования | Знания: | теоретические основы многомерной модели данных |
| | | | Умения: | из многомерного массива данных выделять «измерения» и «факты», соответствующие методам анализа данных |
| | | | Навыки | интеллектуального анализа данных для решения различных прикладных задач в профессиональной деятельности |

3 Место дисциплины (модуля) в структуре основной образовательной программы

Освоение дисциплины формирует у обучающихся компетенции, необходимые для подготовки бакалавра в соответствии с требованиями ФГОС ВО.

Данная дисциплина относится к вариативной части учебного плана.

Входными требованиями, необходимыми для освоения дисциплины, является наличие у обучающихся компетенций, сформированных при изучении дисциплин «Технология программирования», «Компьютерный анализ данных»

Освоение дисциплины «Хранилища данных» является необходимым условием для успешного изучения курса «Интеллектуальные информационные системы».

4. Объем дисциплины (модуля)

Объем дисциплины (модуля) в зачетных единицах с указанием количества академических часов, выделенных на контактную работу с обучающимися (по видам учебных занятий) и на самостоятельную работу по всем формам обучения, приведен в таблице 2.

Таблица 2 – Общая трудоемкость дисциплины

| Название ОПОП | Форма обучения | Цикл | Курс | Трудоемкость (З.Е.) | Объем контактной работы (час) | | | | | СРС | Форма аттестации | |
|---------------|----------------|----------|------|---------------------|-------------------------------|------------|------|-----|---------------|-----|------------------|-----|
| | | | | | Всего | Аудиторная | | | Внеаудиторная | | | |
| | | | | | | лек | прак | лаб | ПА | | | КСР |
| Б-БИ | ОЗФО | Б.1.В.14 | 4 | 3 | 24 | 5 | - | 10 | 9 | 84 | ДЗ | |

5 Структура и содержание дисциплины (модуля)

5.1 Структура дисциплины (модуля)

Тематический план, отражающий содержание дисциплины (перечень разделов и тем), структурированное по видам учебных занятий с указанием их объемов в соответствии с учебным планом, приведен в таблице 3.

Таблица 3 – Структура дисциплины

| № | Название темы | Вид занятия |
|---|--|-------------|
| 1 | Тема 1.1 Архитектуры данных: история развития. | Лекция |
| 2 | Тема 2.1 Архитектуры данных: Базы данных и | Лекция |

| | | |
|----|--|---|
| | модели данных. | / Лабораторная работа: Формирование необходимых навыков работы с Deductor – мастер-класс |
| 3 | Тема 2.2 Многомерные данные. | Лекция |
| | | / Лабораторная работа: Многомерные наборы данных – использование платформы Deductor |
| 4 | Тема 2.3 Концепция хранилищ данных (ХД). OLAP как ключевой компонент ХД. | Лекция |
| | | / Лабораторная работа: Агрегированные показатели – использование платформы Deductor |
| 5 | Тема 3.1 Архитектуры хранилищ данных. | Лекция |
| 6 | Тема 3.2 Реляционные хранилища данных. | Лабораторная работа: Изучение концепции реляционного хранилища данных, используя учебный пример разработчика. |
| 7 | Тема 3.3 Реализация реляционных хранилищ данных. | Лекция |
| | | Лабораторная работа: Реализация РХД архитектуры «звезда». Тема Использование РХД: анализ многомерных данных из данных из РХД. |
| 8 | Тема 4.1 Виртуальные хранилища данных. | Лекция |
| 9 | Тема 4.2 Использование хранилищ данных. | Лекция |
| 10 | Тема 4.3. Реляционная модель данных | Лабораторная работа: Разработка экономической учетной информационной системы на основе реляционной базы данных / Разработка РХД средствами СУБД |
| 11 | Тема 4.4 Анализ данных предметной области для загрузки в ХД | Лабораторная работа: |
| 12 | Тема 4.5 Разработка модели ХД | Лабораторная работа: |

| | | |
|----|---|----------------------|
| 13 | Тема 4.6 Реализация РХД | Лабораторная работа: |
| 14 | Тема 4.7 Загрузка данных в ХД. Проверка работоспособности | Лабораторная работа: |

5.2 Содержание дисциплины (модуля)

Модуль 1 Введение в курс «Хранилища данных»

Тема 1.1 Лекция. Архитектуры данных: история развития. Эволюция задач сбора и обработки информации. Понятие архитектуры данных. Развитие систем хранения и обработки данных. Системы оперативной обработки информации – OLTP. Системы консолидации и аналитической обработки информации – ELT.

Модуль 2 Модели данных, системы хранения данных

Тема 2.1 Лекция. Архитектуры данных: Базы данных и модели данных. – Иерархическая модель данных, условия целостности иерархической модели данных. Сетевая модель данных, условия целостности сетевой модели данных. Реляционная модель данных, реляционные базы данных. Хранилища данных – системы хранения данных, ориентированная на аналитическую обработку.

Лабораторная работа. *Формирование необходимых навыков работы с Deductor – мастер-класс*

- Архитектура и назначение аналитической платформы Deductor;
- Главное меню и элементы управления;
- Загрузка массивов данных;
- Создание метаданных;
- Инструменты визуализации;
- Краткий обзор видов анализа.

Тема 2.2 Лекция. Многомерные данные. OLAP-технология, как ключевой компонент ХД. Задачи OLAP-систем: представление данных, процессы обработки. Концепция многомерного представления данных – гиперкубы. Базовые понятия: измерения и факты. Формализация многомерного представления данных: метки, иерархии, ячейки, меры.

Лабораторная работа. *Многомерные наборы данных.*

- Многомерные наборы данных – использование платформы Deductor.
- Загрузка многомерных массивов данных;
- Задание структуры метаданных;
- Формирование представления исходных данных – таблицы, диаграммы, отчеты по загрузке;

Тема 2.3 Лекция. Концепция хранилищ данных (ХД). OLAP как ключевой компонент ХД. Построение информационных систем на основе архитектур хранилищ данных. Операции над многомерными данными. Методы обработки агрегированных данных. Лабораторная работа. *Агрегированные показатели.*

- Агрегированные показатели – использование платформы Deductor
- Формирование OLAP-кубов: выбор измерений, выбор агрегированных показателей.
- Создание иерархических измерений.

Модуль 3 Архитектуры хранилищ данных.

Тема 3.1 Лекция. Архитектуры хранилищ данных. Современное представление. Классификация архитектур данных. Многомерные реляционные и гибридные хранилища данных. Различия концепций и особенности построения.

Тема 3.2 Лекция. Реляционные хранилища данных. Применение реляционной модели для создания хранилищ данных (ХД). Архитектуры реляционных ХД: «звезда», «снежинка». Особенности реализации реляционных ХД.

Лабораторная работа. *Разработка реляционного хранилища данных. Изучение концепции реляционного хранилища данных на пример разработчика платформы Deductor.*

Тема 3.3 Лекция. Реализация реляционных хранилищ данных. Анализ и преобразование исходных данных. Выбор архитектуры реляционной ХД. Метаданные. Этапы реализации проекта. Реализация аналитической обработки загруженных в ХД данных.

- Лабораторная работа. *Разработка реляционного хранилища данных.*
- Реализация РХД архитектуры «звезда».
- Реализация РХД архитектуры «снежинка» и загрузка данных.
- Использование РХД: анализ многомерных данных из данных из РХД.

Модуль 4 Проектирование ХД для выбранной предметной области.

Тема 4.1 Лекция. Виртуальные хранилища данных.

Тема 4.2 Лекция. Использование хранилищ данных. Различные архитектурные решения ХД, реализация процедур ETL.

Тема 4.3 Лабораторная работа. *Разработка экономической учетной информационной системы на основе реляционной базы данных.*

- Проанализировать задачу учета для предложенного объекта автоматизации. Определить минимально необходимые атрибуты. Провести нормализацию. Предусмотреть таблицу типа «журнал регистрации». Разработать консолидирующие запросы. Сформировать отчеты по результатам запросов в диаграммной форме.

Лабораторная работа. *Разработка РХД средствами СУБД*

- Реализация РХД архитектуры «звезда» (средствами СУБД) – пробный проект. Ставится задача «трансформировать» базу данных учетной информационной системы до ХД. Сформировать срезы OLAP–кубов средствами запросов.

Тема 4.4 Лабораторная работа. *Анализ данных для загрузки в ХД.*

- Выбор данных необходимых исследования и анализа предметной области;
- Разработка многомерной структуры, выбор измерений и фактов;
- Выбор показателей пригодных для агрегирования, формирования агрегированных данных.

Тема 4.5 Лабораторная работа. *Разработка модели ХД.*

- Формирование метаданных;
- Разработка логической модели ХД;

Тема 4.6 Лабораторная работа. *Реализация РХ.*

- Разработка физической модели ХД;
- Реализация реляционного хранилища данных средствами доступной СУБД;

Тема 4.7 Лабораторная работа. *Загрузка данных в ХД. Проверка работоспособности.*

- Загрузка данных в таблицы ХД;
- Проверка целостности данных;
- Формирование срезов гиперкуба.

6. Методические указания для обучающихся по освоению дисциплины (модуля)

Современные корпоративные информационные системы содержат приложения, предназначенные для комплексного анализа данных, поддержки принятия решения, извлечения знаний.

Принятия эффективного управленческого решения невозможно, без консолидированных отчетов и глубинного анализа результатов. Для этого необходимо создание хранилищ данных (Data warehouses), специальных систем хранения, обеспечивающих загрузку предварительно извлеченных и обработанных данных и оперативный доступ к ним.

Основные требования к хранилищам данных:

- поддержка высокой скорости доступа к данным;
- поддержка внутренней непротиворечивости данных;
- возможность манипулирования данными;
- полнота и достоверность хранимых данных;
- поддержка процессов обновления данных.

В отличие от оперативных баз данных, на основе которых строятся учетные информационные системы, хранилища данных предназначены исключительно для аналитической обработки данных. Данные загружаются в хранилище из оперативных баз данных.

Перечень и тематика самостоятельных работ студентов по дисциплине

Раздел 1.

- 1) Работа с платформой Deductor Academic. Освоение загрузки данных и методов визуализации.
- 2) Методы предварительной обработки данных, реализованные в Deductor Academic. Рассмотреть назначение с методов, составить таблицу, отражающую технологию предварительной обработки и соответствующий математический метод.
- 3) Аналитическая обработка данных средствами Deductor Academic.

Раздел 2.

- 1) Изучение учебной РХД (поставляется разработчиком платформы)

Раздел 3

- 1) Разработка ER-модели для заданной предметной области.

Раздел 4.

- 1) Анализ выбранной предметной области.
- 2) Подготовка данных из открытых источников – извлечение исходных данных – для дальнейшей загрузки в РХД.
- 3) Разработка логической модели РХД для выбранной предметной области.

Контрольные вопросы для самостоятельной оценки качества освоения учебной дисциплины

1. Понятие архитектуры данных. Понятие модели данных.
2. Понятие целостности данных
3. Системы оперативной обработки информации – OLTP. Архитектура и назначение.
4. Системы консолидации и аналитической обработки информации – ELT. Архитектура и назначение.
5. Иерархическая модель данных. Условие целостности
6. Сетевая модель данных. Условие целостности
7. Реляционная модель данных.
8. Простые и агрегированные показатели.
9. Концепция многомерного представления данных – гиперкубы.
10. Измерения и факты в гиперкубах. Правил выбора измерений и фактов.
11. Формализация многомерного представления данных: метки, иерархии, ячейки, меры.
12. Операции над данными в гиперкубах: вращение, сечение (срез), свертка и детализация.
13. Агрегация в гиперкубах: виды агрегации.
14. Агрегация в гиперкубах – оценка числа агрегатов для двумерного случая.
15. Концепция хранилищ данных.
16. Понятие метаданных.
17. Классификация архитектуры хранилищ данных.
18. Многомерные хранилища данных.
19. Различие концепций и особенности построения.

20. Реляционные хранилища данных. Применение реляционной модели для создания хранилищ данных (ХД).
21. Виртуальные хранилища данных.
22. Витрины данных. Назначение, использование в системах хранения данных.

7. Перечень учебно-методического обеспечения для самостоятельной работы

Доступ к средствам вычислительной техники и программному обеспечению, указанному в п.11.

Наличие комплектов исходных данных и методических материалов, доступных на студенческом сервере. Доступ на сервер.

Наличие комплектов индивидуальных заданий, доступных на студенческом сервере.

8. Фонд оценочных средств для проведения промежуточной аттестации

В соответствии с требованиями ФГОС ВО для аттестации обучающихся на соответствие их персональных достижений планируемым результатам обучения по дисциплине созданы фонды оценочных средств (Приложение 1).

Фонд оценочных средств по дисциплине включает в себя:

1. практические задания для промежуточной аттестации – контрольные работы;
2. перечень контрольных вопросов для итоговой аттестации;
3. темы итогового индивидуального задания – проекта ХД.

9. Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины (модуля)

а) основная литература

1. Парфенов, Ю. П. Постреляционные хранилища данных: учебное пособие для вузов / Ю. П. Парфенов ; под науч. ред. Н. В. Папуловской. — М.: Издательство Юрайт, 2017. — 121 с. — (Университеты России). — ISBN 978-5-534-03408-0. Режим доступа: <https://www.biblio-online.ru/book/628DAC6C-ECBF-45B3-BD23-F6B57148D18F>
2. Алексеева Е.В., Амириди Ю.В., Дик В.В. Информационные аналитические системы [Электронный ресурс]: учебник / Т. В. Алексеева, Ю. В. Амириди, В. В. Дик и др.; под ред. В. В. Дика. - М.: МФПУ Синергия, 2013. - 384 с. - (Университетская серия). Режим доступа: <http://znanium.com/catalog.php?bookinfo=451186>

б) дополнительная литература

3. Туманов В.Е. , Маклаков С.В. Проектирование реляционных хранилищ данных / М.: Диалог-МИФИ, 2007 – 333 с.
4. Архипенков С., Голубев Д., Максименко О. Хранилища данных. От концепции до внедрения / М.: Диалог-МИФИ, 2002. – 528 с.
5. Паклин Н.Б., Орешков В.И. Бизнес-аналитика: от данных к знаниям – СПб.: Питер, 2009. – 624 с.
6. Инмон Б. Типы хранилищ данных. Перевод Intersoftlab, 2001, <http://www.iso.ru/journal/articles/181.html>.
7. Кузнецов С., Артемьев В. Обзор возможностей применения ведущих СУБД для построения хранилищ данных (DataWarehouse). <http://www.citforum.ru/database/kbd98/glava15.shtml>
8. В. П. Божко, А. В. Хорошилов, В. А. Благодатских и др., Предметно-ориентированные экономические информационные системы: учебник для студ. вузов. / М. : Финансы и статистика, 2007. - 224 с. : ил.
9. Конноллн, Томас, Бегг, Карелии. Базы данных. Проектирование, реализация и

сопровождение. Теория и практика. 3-е издание.: Пер. с англ. - М.: Издательский дом "Вильямс", 2003. - 1440 с. : ил.

10. Перечень ресурсов информационно - телекоммуникационной сети «Интернет»

а) полнотекстовые базы данных:

1. Научная электронная библиотека eLIBRARY.RU: <http://www.eLIBRARY.RU>
2. Ресурс Цифровые учебные материалы <http://abc.vvsu.ru/>
3. ЭБС «Руконт»: <http://www.rucont.ru/>
4. ЭБС «Юрайт»: <http://www.biblio-online.ru/>

11. Материально-техническое обеспечение дисциплины (модуля)

Для проведения лекционных занятий по данной дисциплине используются аудитории, оснащенные мультимедийным оборудованием.

Помещение для самостоятельной работы обучающихся должны быть оснащены компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду университета.