МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ

ВЛАДИВОСТОКСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

ЭКОНОМИКИ И СЕРВИСА

КАФЕДРА информационных технологий и систем

**ХРАНИЛИЩА ДАННЫХ**

Рабочая программа дисциплины

09.03.02 Информационные системы и технологии

Владивосток 2016

Рабочая программа дисциплины «Хранилища данных» составлена в соответствии с требованиями ФГОС ВО по направлению подготовки 09.03.02 Информационные системы и технологии, бакалавр и Порядком организации и осуществления образовательной деятельности по образовательным программам высшего образования – программам бакалавриата, программам специалитета, программам магистратуры (утв. приказом Минобрнауки России от 19 декабря 2013 г. N 1367)

Составитель:

Кригер А.Б. к.ф.-м.н., доцент

Утверждена на заседании кафедры Информационных технологий и систем от 20.04.2016 г., протокол № \_9

Заведующий кафедрой (разработчика) \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Кийкова Е.В.

*подпись фамилия, инициалы*

«\_\_\_\_»\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_20\_\_г.

Заведующий кафедрой (выпускающей) \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Кийкова Е.В.

*подпись фамилия, инициалы*

«\_\_\_\_»\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_20\_\_г.

**1 Цель и задачи освоения дисциплины**

**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, соотнесенных с планируемыми результатами освоения образовательной программы**

Планируемыми результатами обучения по дисциплине, являются знания, умения, владения и/или опыт деятельности, 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 | |
| 09.03.02 Информационные системы и технологии | ОПК-1 | владением широкой общей подготовкой (базовыми знаниями) для решения практических задач в  области информационных систем и технологий | Знание: | основные модели данных |
| Умение: | создавать инфологические модели данных |
| Владение: | инструментальными средствами разработки логической и физической моделей данных |
| ОПК-3 | Способность применять основные приемы и законы создания и чтения чертежей и документации по аппаратным и программным компонентам информационных систем | Умения: | использовать стандарты. Применять современные информационные технологии разработки чертежей и документации. |
| ОПК -6 | Способность выбирать и оценивать способ реализации информационных систем и устройств | Умения: | Выбирать системы хранения данных соответствующие сущности задач обработки информации. |
| ПК-2 | Способность проводить техническое проектирование | Владения: | Методами и технологиями проектирования систем хранения данных. |

Планируемыми результатами обучения по дисциплине, являются знания, умения, владения и/или опыт деятельности, характеризующие этапы/уровни формирования компетенций и обеспечивающие достижение планируемых результатов освоения образовательной программы в целом.

**3. Место дисциплины в структуре основной образовательной программы**

профессиональный цикл, базовая часть

**4. Объем дисциплины**

Объем дисциплины 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 ООП | Форма обучения | Цикл | Семестр  курс | Трудоемкость | Объем контактной работы (час) | | | | | | СРС | Форма аттестации |
| (З.Е.) | Всего | Аудиторная | | | Внеаудитор  ная | |
| лек. | прак. | лаб. | ПА | КСР |
| 09.03.02 Информационные системы и технологии, бакалавр | ОФО | Бл1.В | 6 | 3 | 51 | 17 | - | 34 |  |  | 57 | ДЗ |
| 09.03.02 Информационные системы и технологии, бакалавр | ОЗФО | Бл1.В | 7 | 3 | 12 | 4 |  | 8 |  | 2 |  | ДЗ |

**5 Структура и содержание дисциплины**

**5.1 Структура дисциплины**

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 | Объем час | Кол-во часов в интерактивной и  электронной  форме | СРС |
| *1* | Тема 1.1 Архитектуры данных: история развития*.* | *Лекция* | 1 | - | 1 |
| 2 | Тема 2.1 Архитектуры данных: Базы данных и модели данных. | *Лекция* | 2 |  | 4 |
| */ Лабораторная работа: Формирование необходимых навыков работы с Deductor – мастер-класс* | 2 | - |
| 3 | Тема 2.2 Многомерные данные. | *Лекция* | 2 |  | 6 |
| */ Лабораторная работа:* *Многомерные наборы данных – использование платформы Deductor* | 4 |
| 4 | Тема 2.3 Концепция хранилищ данных (ХД). OLAP как ключевой компонент ХД. | *Лекция* | 4 |  | 8 |
| */ Лабораторная работа: Агрегированные показатели – использование платформы Deductor* | 4 |
| 5 | Тема 3.1 Архитектуры хранилищ данных. | *Лекция* | 3 |  | 3 |
| 6 | Тема 3.2 Реляционные хранилища данных. | *Лабораторная работа: Изучение концепции реляционного хранилища данных, используя учебный пример разработчика.* | 2 |  | 2 |
| 7 | Тема 3.3 Реализация реляционных хранилищ данных. | *Лекция* | 2 |  | 6 |
| *Лабораторная работа: Реализация РХД архитектуры «звезда».* *Тема Использование РХД: анализ многомерных данных из данных из РХД.* | 4 |
|  | Раздел 4 Проект ХД для выбранной предметной области. Работа в малых группах | | | | |
|  | Тема 4.1 Виртуальные хранилища данных. | *Лекция* | 1 |  | 2 |
|  | Тема 4.2 Использование хранилищ данных. | *Лекция* | 1 |  | 2 |
|  | Тема 4.3 Реляционная модель данных | *Лабораторная работа:* *Разработка экономической учетной информационной системы на основе реляционной базы данных /* *Разработка РХД средствами СУБД* | 6 |  | 6 |
|  | Тема 4.4 Анализ данных предметной области для загрузки в ХД | *Лабораторная работа:* | 4 |  | 4 |
|  | Тема 4.5 Разработка модели ХД | *Лабораторная работа:* | 2 |  | 2 |
|  | Тема 4.6 Реализация РХД | *Лабораторная работа:* | 2 |  | 2 |
|  | Тема 4.7 Загрузка данных в ХД. Проверка работоспособности | *Лабораторная работа:* | 4 |  | 4 |

**5.2 Содержание дисциплины**

1. **Введение в курс «Хранилища данных» -** 1 час
   1. Лекция. Архитектуры данных: история развития. – 1часа. 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.
2. **Модели данных, системы хранения данных – 4 часа**
   1. Лекция. Архитектуры данных: Базы данных и модели данных. – 2 часа. 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* (2 часа)

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

Литература

1. Туманов В.Е. Основы проектирования реляционных баз данных [Электронный ресурс]/ Интернет-Университет Информационных Технологий (ИНТУИТ), 2016. – Режим доступа: <http://www.iprbookshop.ru/52221.html>
2. Паклин Н.Б., Орешков В.И. Бизнес-аналитика: от данных к знаниям – СПб.: Питер, 2009. – 624 с.
3. Конноллн, Томас, Бегг, Карелии. Базы данных. Проектирование, реализация и сопровождение. Теория и практика. 3-е издание, : Пер. с англ. - М.: Издательский дом "Вильяме", 2003. - 1440 с.

*Формы и методы проведения занятий* по теме, применяемые образовательные технологии: метод активного обучения – «мастер-класс»

*Виды самостоятельной подготовки студентов по теме:* компьютерное моделирование

* 1. Лекция. Многомерные данные. OLAP-технология, как ключевой компонент ХД. – 2 часа. 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*. – 2 часа

* + - Загрузка многомерных массивов данных;
    - Задание структуры метаданных;
    - Формирование представления исходных данных – таблицы, диаграммы, отчеты по загрузке;
  1. Лекция. Концепция хранилищ данных (ХД). OLAP как ключевой компонент ХД. – 2 часа. Построение информационных систем на основе архитектур хранилищ данных. Операции над многомерными данными. Методы обработки агрегированных данных.

Лабораторная работа. *Агрегированные показатели. –* 2 часа.

* + - Формирование OLAP-кубов: выбор измерений, выбор агрегированных показателей.
    - Создание иерархических измерений.

Литература

1. 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>
2. Туманов В.Е. , Маклаков С.В. Проектирование реляционных хранилищ данных / М: Диалог-МИФИ, 2007 – 333 с.
3. Архипенков С., Голубев Д., Максименко О. Хранилища данных. От концепции до внедрения / М.: Диалог-МИФИ, 2002. – 528 с.

*Формы и методы проведения занятий* по теме, применяемые образовательные технологии: компьютерное моделирование; метод активного обучения – «конференция».

*Виды самостоятельной подготовки студентов по теме:* изучение теоретических вопросов, компьютерное моделирование

1. **Архитектуры хранилищ данных.** 
   1. Лекция. Архитектуры хранилищ данных. – 3 часа. Современное представление. Классификация архитектур данных. Многомерные реляционные и гибридные хранилища данных. Различие концепций и особенности построения.
   2. Лекция. Реляционные хранилища данных. – 1 час. 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.*

Литература

1. Туманов В.Е. , Маклаков С.В. Проектирование реляционных хранилищ данных / М: Диалог-МИФИ, 2007 – 333 с.
2. Паклин Н.Б., Орешков В.И. Бизнес-аналитика: от данных к знаниям – СПб.: Питер, 2009. – 624 с.

*Формы и методы проведения занятий* по теме, применяемые образовательные технологии: компьютерное моделирование,

*Виды самостоятельной подготовки студентов по теме:* изучение теоретических вопросов, компьютерное моделирование

* 1. Лекция. Реализация реляционных хранилищ данных. – 2 часа. 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***. –** 4 часа.
     + Реализация РХД архитектуры «звезда».
     + Реализация РХД архитектуры «снежинка» и загрузка данных.
     + Использование РХД: анализ многомерных данных из данных из РХД.

Литература

1. Туманов В.Е. Основы проектирования реляционных баз данных [Электронный ресурс]/ Интернет-Университет Информационных Технологий (ИНТУИТ), 2016. – Режим доступа: <http://www.iprbookshop.ru/52221.html>
2. Паклин Н.Б., Орешков В.И. Бизнес-аналитика: от данных к знаниям – СПб.: Питер, 2009. – 624 с.
3. Туманов В.Е. , Маклаков С.В. Проектирование реляционных хранилищ данных / М: Диалог-МИФИ, 2007 – 333 с.

*Формы и методы проведения занятий* по теме, применяемые образовательные технологии: компьютерное моделирование,

*Виды самостоятельной подготовки студентов по теме:* изучение теоретических вопросов, компьютерное моделирование

1. **Проектирование ХД для выбранной предметной области.**
   1. Лекция. Виртуальные хранилища данных. – 1 час
   2. Лекция. Использование хранилищ данных. Различные архитектурные решения ХД, реализация процедур ETL. – 1 час.
   3. Лабораторная работа. *Разработка экономической учетной информационной системы на основе реляционной базы данных. –* 4 часа
      * Проанализировать задачу учета для предложенного объекта автоматизации. Определить минимально необходимые атрибуты. Провести нормализацию. Предусмотреть таблицу типа «журнал регистрации». Разработать консолидирующие запросы. Сформировать отчеты по результат запросов в диаграммной форме.

Лабораторная работа. *Разработка РХД средствами СУБД* – 2 часа.

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

Литература

1. Туманов В.Е. , Маклаков С.В. Проектирование реляционных хранилищ данных / М: Диалог-МИФИ, 2007 – 333 с.
2. Инмон Б. Типы хранилищ данных. Перевод Intersoftlab, 2001, <http://www.iso.ru/journal/articles/181.html>.
3. Кузнецов С., Артемьев В. Обзор возможностей применения ведущих СУБД для построения хранилищ данных (DataWarehouse). <http://www.citforum.ru/database/kbd98/glava15.shtml>.

*Формы и методы проведения занятий* по теме, применяемые образовательные технологии: компьютерное моделирование.

*Виды самостоятельной подготовки студентов по теме:* изучение теоретических вопросов, компьютерное моделирование.

***Проектирование ХД для выбранной предметной области. Работа в малых группах***

* 1. Лабораторная работа. *Анализ данных для загрузки в ХД. –* 4 часа.
     + Выбор данных необходимых для исследования и анализа предметной области;
     + Разработка многомерной структуры, выбор измерений и фактов;
     + Выбор показателей пригодных для агрегирования, формирования агрегированных данных.
  2. Лабораторная работа. *Разработка модели ХД.* – 2часа.
     + Формирование метаданных;
     + Разработка логической модели ХД;
  3. Лабораторная работа. *Реализация РХД.* – 2часа.
     + Разработка физической модели ХД;
     + Реализация реляционного хранилища данных средствами доступной СУБД;
  4. Лабораторная работа. *Загрузка данных в ХД. Проверка работоспособности. –* 4 часа.
     + Загрузка данных в таблицы ХД;
     + Проверка целостности данных;

Литература

1. Туманов В.Е. , Маклаков С.В. Проектирование реляционных хранилищ данных / М: Диалог-МИФИ, 2007 – 333 с.
2. Паклин Н.Б., Орешков В.И. Бизнес-аналитика: от данных к знаниям – СПб.: Питер, 2009. – 624 с.
3. Инмон Б. Типы хранилищ данных. Перевод Intersoftlab, 2001, <http://www.iso.ru/journal/articles/181.html>.
4. Кузнецов С., Артемьев В. Обзор возможностей применения ведущих СУБД для построения хранилищ данных (DataWarehouse). <http://www.citforum.ru/database/kbd98/glava15.shtml>

*Формы и методы проведения занятий* по теме, применяемые образовательные технологии: компьютерное моделирование, метод активного обучения – «конференция».

*Виды самостоятельной подготовки студентов по теме:* изучение теоретических вопросов, компьютерное моделирование

**6. Методические указания для обучающихся по освоению дисциплины**

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 бля выбранной предметной области.

**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  по планируемым результатам обучения по дисциплине созданы фонды оценочных средств (Приложение 1).

**9. Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины**

**а) основная литература**

1. Туманов В.Е. Основы проектирования реляционных баз данных [Электронный ресурс]/ Интернет-Университет Информационных Технологий (ИНТУИТ), 2016. – Режим доступа: <http://www.iprbookshop.ru/52221.html>
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**

1. Туманов В.Е. , Маклаков С.В. Проектирование реляционных хранилищ данных / М: Диалог-МИФИ, 2007 – 333 с.
2. Архипенков С., Голубев Д., Максименко О. Хранилища данных. От концепции до внедрения / М.: Диалог-МИФИ, 2002. – 528 с.
3. Паклин Н.Б., Орешков В.И. Бизнес-аналитика: от данных к знаниям – СПб.: Питер, 2009. – 624 с.
4. Инмон Б. Типы хранилищ данных. Перевод Intersoftlab, 2001, <http://www.iso.ru/journal/articles/181.html>
5. Кузнецов С., Артемьев В. Обзор возможностей применения ведущих СУБД для построения хранилищ данных (DataWarehouse). <http://www.citforum.ru/database/kbd98/glava15.shtml>
6. 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.
7. 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

Электронно-библиотечная система IPRbooks. ЭБС содержит более 20 000 учебных и научных изданий по различным дисциплинам, свыше 200 наименований российских и зарубежных журналов, большая часть которых входит в перечень ВАК. <http://www.iprbookshop.ru/>

б) интернет-ресурсы

<http://www.citforum.ru/>

<http://www.basegroup.ru>

**11. Перечень информационных технологий**

программное обеспечение:

* IBM Cognos Insight – система хранения данных – хранилище данных;
* Аналитическая платформа Deductor;
* MS SQL Server / MS Access;
* MS Excel.

**12. Электронная поддержка дисциплины**

**13. Материально-техническое обеспечение дисциплины**

мультимедийное оборудование

**14. Словарь основных терминов** источник [электронный ресурс], режим доступа: http://www.osp.ru/cio/2002/03/172082/

1. **Агрегирование данных *(Data Aggregation)***

Операция над *кубом данных*, которая строит куб меньшей размерности путем замены совокупностей значений *элементов* одного или нескольких *измерений*значениями элементов более высоких уровней и формирования в качестве содержимого *ячеек* результирующего куба соответствующих суммарных значений *показателей*.

1. **Анализ данных глубинный *(Data Mining)***

Технология анализа данных в базах данных или *хранилищах данных*, основанная на статистических методах и служащая для выявления заранее неизвестных закономерностей. Широко распространена на практике для поддержки принятия стратегически важных решений. В отечественной литературе Data Mining часто переводится как «добыча данных», «исследование данных», «интеллектуальный анализ данных», «разведка данных» и т. п.

1. **Анализ данных с последовательным огрублением *(Roll-Up Analysis)***

Метод анализа данных в *хранилищах данных*, предусматривающий пошаговый переход к уровням более агрегированного представления данных с целью получения огрубленных оценок исследуемых процессов или явлений в таких условиях, когда более точные оценки не требуются или не могут быть определены.

1. **Анализ данных с последовательным уточнением *(Drill-Down Analysis)***

Метод анализа данных в хранилищах данных, предусматривающий пошаговый переход к уровням более детализированного представления данных для получения возможно более точных оценок исследуемых процессов или явлений.

1. **Атрибут элемента измерения *(Dimension Element Attribute)***

Именованный домен значений, соответствующий некоторому *элементу измерения* в *многомерной базе данных*.

1. **База данных многомерная *(Multi-Dimensional Database)***

База данных, основанная на *многомерной модели данных*. Обычно используется для целей анализа. Основным ее структурным компонентом является куб данных.

1. **Витрина данных *(Data Mart)***

См. *[Хранилище данных специализированное](http://www.osp.ru/cio/2002/03/172082/" \l "vr)*.

1. **Витрина данных независимая *(Independent Data Mart)***

*Витрина данных*, базирующаяся на собственных источниках данных, а не на едином *хранилище данных* компании. Обычно применяются в крупных организациях, где имеются независимые подразделения со своими собственными информационными службами или «продвинутые» пользователи (например, банковские аналитики), которые могут поддерживать такие витрины при минимальных внешних консультациях.

1. **Вращение измерений *(Dimension Rotation)***

Операция, обеспечивающая изменение порядка *измерений куба* данных при его визуализации. Позволяет представить его пользователю в другом ракурсе.

1. **Данные исторические *(Historical Data)***

Данные, ассоциированные со временем (обычно со временем появления соответствующего факта предметной области) и представляющие либо уже состоявшиеся *факты*, либо модельные значения фактов для анализа «что — если».

1. **Дезагрегирование данных *(Data Disaggregation)***

Операция над *кубом данных*, которая строит новый куб большей размерности путем замены значений некоторых *элементов измерений* совокупностями значений элементов более низких уровней иерархии и декомпозиции соответствующих значений исходных *показателей*. Метод декомпозиции зависит от природы представленных данных и определяется исследователем данных. Дезагрегирование данных позволяет, например, перейти от рассмотрения годовых данных по некоторым категориям продуктов к квартальным данным.

1. **Измерение *(Dimension)***

Одна из осей координат *куба данных*. Связанные с нею значения характеризуют какое-либо классификационное свойство сущностей предметной области, например временные характеристики (годы, месяцы, кварталы) или административную принадлежность (регион, город, район) и т. п. На совокупности значений, относящейся к некоторому измерению, могут быть определены иерархические отношения, которые позволяют осуществлять*агрегирование данных*, представляющих ассоциированные с *ячейками куба факты*.

1. **Куб данных *(Data Cube)***

Основная структура данных в *многомерных моделях данных*. Имеет несколько независимых *измерений* — систему координат представляемого пространства данных. Комбинации значений координат по всем измерениям определяют точки куба, называемые *ячейками*. С ячейками ассоциируются значения переменных, называемых *показателями* и имеющих, как правило, числовые типы.

1. **Куб данных плотный *(Dense Data Cube)***

* Куб данных с большой долей заполненных ячеек.

1. **Куб данных разреженный *(Sparse Data Cube)***

* Куб данных с малой долей заполненных ячеек.

1. **Модель данных многомерная *(Multi-Dimensional Data Model)***

* Модель данных, оперирующая многомерными представлениями данных в виде*кубов данных*. Такие модели данных стали широко использоваться в середине 90-х годов в связи с развитием технологий OLAP. Операционные возможности многомерных моделей данных включают, в частности, операции *агрегирования*и *дезагрегирования данных*, построения *проекций* куба, *вращения измерений*куба и др.

1. **Модель измерений *(Dimensional Model)***

См. *Модель данных многомерная*.

1. **Мультикуб данных *(Data Multicube)***

Структура *многомерной базы данных*, состоящая из нескольких *кубов данных*.

1. **Очистка данных *(Data Scrubbing, Data Cleansing)***

Процедура предварительной обработки данных, собранных из нескольких источников для загрузки в *хранилище данных*. Ее целью является фильтрация данных и их консолидация, верификация и обеспечение логической целостности, устранение несогласованности и различных ошибок, восполнение пропусков и другие действия, направленные на улучшение качества данных.

1. **Показатель *(Measure)***

Переменная, значениями которой являются *факты*, ассоциируемые с *ячейками куба данных*.

1. **Показатель аддитивный *(Additive Measure)***

Показатель, значениями которого являются *аддитивные факты*.

1. **Показатель неаддитивный *(Nonadditive Measure)***

Показатель, значениями которого являются *неаддитивные факты*.

1. **Показатель полуаддитивный *(Semiadditive Measure)***

Показатель, значениями которого являются *полуаддитивные факты*.

1. **Проекция куба данных *(Data Cube Projection)***
2. Операция над *кубом данных*, которая строит куб меньшей размерности при фиксированных значениях для каких-либо *измерений*. Этот термин обозначает также результат указанной операции.
3. **Сечение куба данных *(Data Cube Slice)***

См. *Проекция куба данных*.

1. **Сжатие куба данных *(Data Cube Compression)***

Операция над кубом данных, обеспечивающая уменьшение используемого для него объема памяти в среде хранения.

1. **Система информационная оперативная *(Operative Information System, OIS)***

Информационная система, предназначенная для поддержки деловых операций, составляющих бизнес-процессы регулярной производственной и вспомогательной деятельности предприятия. В публикациях встречаются также синонимы «операциональная ИС», «операционная ИС» (т. е. поддерживающая выполнение различных текущих операций компании, Operational IS) и «транзакционная ИС» (т. е. поддерживающая выполнение бизнес-транзакций компании, ее деловых операций или сделок в отличие от аналитической ИС).

1. **Система информационная руководителя *(Executive Information System, EIS)***

Информационная система компании, служащая для анализа и предоставления данных о ее функционировании и поддержки принятия управленческих решений. Источником данных для этой системы является база данных *оперативной информационной системы*. Информационную систему руководителя называют также «аналитической системой», если кроме средств отображения агрегированных значений регламентированных показателей система предоставляет какие-либо возможности для аналитической обработки данных.

1. **Система поддержки принятия решений *(Decision Support System, DSS)***

Система, обеспечивающая возможности изучения состояния, прогнозирования развития и оценки возможных вариантов поведения на основе анализа статистических данных, отражающих результаты деятельности компании на протяжении времени. В таких системах применяются современные технологии баз данных, OLAP, *хранилищ данных, глубинного анализа* и визуализации данных.

1. **Схема типа звезды *(Star Schema)***

Схема реляционной базы данных, служащая для поддержки многомерного представления содержащихся в ней данных. Описываемая база данных включает *таблицу фактов* и ряд связанных с нею при помощи внешних ключей*таблиц измерений*. Таблицы измерений могут иметь составные первичные ключи и являются денормализованными. Благодаря этому упрощается восприятие структуры данных пользователем и формулировка запросов, уменьшается количество операций соединения таблиц при обработке запросов. Однако в связи с избыточностью данных возрастает требуемый для их хранения объем памяти.

1. **Схема типа снежинки *(Snowflake Schema)***

Разновидность *схемы типа звезды*, предусматривающая нормализацию *таблиц измерений*. Первичные ключи в них состоят из единственного атрибута (соответствуют единственному элементу измерения). Это позволяет минимизировать избыточность данных и более эффективно выполнять запросы, связанные со структурой значений *измерений*.

1. **Таблица измерений *(Dimension Table)***

Таблица в базе данных со *схемой типа звезды* или *типа снежинки*. Строки этой таблицы содержат значения первичных ключей, представляющих значения по какому-либо *измерению*. Каждая таблица измерений с помощью внешнего ключа связана с *таблицей фактов*. В схеме типа звезды таблица измерений денормализована. Напротив, в схеме типа снежинки используются нормализованные таблицы измерений, декомпозированные по уровням иерархии *элементов измерений*. Каждая такая таблица соответствует единственному элементу измерения.

1. **Таблица фактов *(Fact Table)***

Таблица в базе данных со *схемой типа звезды* или *типа снежинки*. Каждая ее строка соответствует некоторой *ячейке куба данных*. Она содержит набор*фактов* и по одному значению внешнего ключа для каждой *таблицы измерений*.

1. **Факт *(Fact)***

Значение *показателя*, соответствующее какой-либо *ячейке куба данных*.

1. **Факт аддитивный *(Additive Fact)***

Факты, допускающие *агрегирование* относительно любого *измерения куба данных*.

1. **Факт неаддитивный *(Nonadditive Fact)***

Факты, которые не могут *агрегироваться* ни по какому *измерению куба данных*.

1. **Факт полуаддитивный *(Semiadditive Fact)***

Факты, которые допускают *агрегирование* относительно одних измерений и не допускают относительно других.

1. **Хранилище данных *(Data Warehouse)***

Информационная система, содержащая непротиворечивые консолидированные исторические данные крупной компании и предоставляющая инструментальные средства для их анализа с целью поддержки принятия стратегических решений. Информационные ресурсы хранилища данных формируются на основе фиксируемых на протяжении продолжительного периода времени моментальных снимков баз данных *оперативной информационной системы*компании и, возможно, различных внешних источников. В хранилищах данных применяются технологии баз данных, OLAP, *глубинного анализа данных*, визуализации данных.

1. **Хранилище данных двухуровневое *(Two-Level Data Warehouse)***

* Единое *хранилище данных* компании, обеспечивающее потребности всех ее подразделений, нуждающихся в средствах анализа данных. Нижний уровень его архитектуры образуют различные источники данных, в частности база данных*оперативной информационной системы*.

1. **Хранилище данных специализированное *(Data Mart)***

*Хранилище данных*, создаваемое для поддержки принятия решений в интересах какого-либо подразделения компании или для обеспечения каких-либо конкретных аспектов ее деятельности. Источником данных для специализированного хранилища данных может быть общее хранилище данных компании или оно создается и функционирует независимо. Объем данных в специализированном хранилище данных и его потребности в вычислительных ресурсах обычно существенно ограничены по сравнению с общим хранилищем данных, оно может содержать часто используемые агрегированные данные. В отечественной литературе термин Data Mart нередко переводят как «витрина данных», «киоск данных» и даже «рынок данных».

1. **Хранилище данных трехуровневое *(Three-Level Data Warehouse)***

*Хранилище данных*, архитектура которого предусматривает поддержку над единым хранилищем данных *витрин данных* для отдельных подразделений компании.

1. **Элемент измерения *(Dimension Element)***

Уровень в иерархии значений координат некоторого *измерения куба данных*. Например, для измерения времени может быть задана иерархия «год — квартал — месяц». В этом случае данному измерению соответствует три элемента измерения, каждый из которых характеризуется своим номером уровня в иерархии.

1. **Ячейка *(Cell)***

Структурный *элемент куба данных*, соответствующий набору значений по всем*измерениям*. С ячейками куба ассоциируются значения *показателей*.

1. **Схема типа звезды *(Star Schema)***

Схема реляционной базы данных, служащая для поддержки многомерного представления содержащихся в ней данных. Описываемая база данных включает *таблицу фактов* и ряд связанных с нею при помощи внешних ключей*таблиц измерений*. Таблицы измерений могут иметь составные первичные ключи и являются денормализованными. Благодаря этому упрощается восприятие структуры данных пользователем и формулировка запросов, уменьшается количество операций соединения таблиц при обработке запросов. Однако в связи с избыточностью данных возрастает требуемый для их хранения объем памяти.

1. **Схема типа снежинки *(Snowflake Schema)***

Разновидность *схемы типа звезды*, предусматривающая нормализацию *таблиц измерений*. Первичные ключи в них состоят из единственного атрибута (соответствуют единственному элементу измерения). Это позволяет минимизировать избыточность данных и более эффективно выполнять запросы, связанные со структурой значений *измерений*.

1. **Таблица измерений *(Dimension Table)***

Таблица в базе данных со *схемой типа звезды* или *типа снежинки*. Строки этой таблицы содержат значения первичных ключей, представляющих значения по какому-либо *измерению*. Каждая таблица измерений с помощью внешнего ключа связана с *таблицей фактов*. В схеме типа звезды таблица измерений денормализована. Напротив, в схеме типа снежинки используются нормализованные таблицы измерений, декомпозированные по уровням иерархии *элементов измерений*. Каждая такая таблица соответствует единственному элементу измерения.

1. **Таблица фактов *(Fact Table)***

Таблица в базе данных со *схемой типа звезды* или *типа снежинки*. Каждая ее строка соответствует некоторой *ячейке куба данных*. Она содержит набор*фактов* и по одному значению внешнего ключа для каждой *таблицы измерений*.

1. **Факт *(Fact)***

Значение *показателя*, соответствующее какой-либо *ячейке куба данных*.

1. **Факт аддитивный *(Additive Fact)***

Факты, допускающие *агрегирование* относительно любого *измерения куба данных*.

1. **Факт неаддитивный *(Nonadditive Fact)***

Факты, которые не могут *агрегироваться* ни по какому *измерению куба данных*.

1. **Факт полуаддитивный *(Semiadditive Fact)***

Факты, которые допускают *агрегирование* относительно одних измерений и не допускают относительно других.

1. **Хранилище данных *(Data Warehouse)***

Информационная система, содержащая непротиворечивые консолидированные исторические данные крупной компании и предоставляющая инструментальные средства для их анализа с целью поддержки принятия стратегических решений. Информационные ресурсы хранилища данных формируются на основе фиксируемых на протяжении продолжительного периода времени моментальных снимков баз данных *оперативной информационной системы*компании и, возможно, различных внешних источников. В хранилищах данных применяются технологии баз данных, OLAP, *глубинного анализа данных*, визуализации данных. Термин Data Warehouse был введен Б. Инмоном в 1990 году. В отечественных публикациях встречается также дословный перевод — «склад данных».

1. **Хранилище данных двухуровневое *(Two-Level Data Warehouse)***

Единое *хранилище данных* компании, обеспечивающее потребности всех ее подразделений, нуждающихся в средствах анализа данных. Нижний уровень его архитектуры образуют различные источники данных, в частности база данных *оперативной информационной системы*.

1. **Хранилище данных специализированное *(Data Mart)***

*Хранилище данных*, создаваемое для поддержки принятия решений в интересах какого-либо подразделения компании или для обеспечения каких-либо конкретных аспектов ее деятельности. Источником данных для специализированного хранилища данных может быть общее хранилище данных компании или оно создается и функционирует независимо. Объем данных в специализированном хранилище данных и его потребности в вычислительных ресурсах обычно существенно ограничены по сравнению с общим хранилищем данных, оно может содержать часто используемые агрегированные данные. В отечественной литературе термин Data Mart нередко переводят как «витрина данных», «киоск данных» и даже «рынок данных».

1. **Хранилище данных трехуровневое *(Three-Level Data Warehouse)***

*Хранилище данных*, архитектура которого предусматривает поддержку над единым хранилищем данных *витрин данных* для отдельных подразделений компании.

1. **Элемент измерения *(Dimension Element)***

Уровень в иерархии значений координат некоторого *измерения куба данных*. Например, для измерения времени может быть задана иерархия «год — квартал — месяц». В этом случае данному измерению соответствует три элемента измерения, каждый из которых характеризуется своим номером уровня в иерархии.

1. **Ячейка *(Cell)***

Структурный *элемент куба данных*, соответствующий набору значений по всем *измерениям*. С ячейками куба ассоциируются значения *показателей*.

***Лист изменений и согласований***

Дополнения и изменения в учебной программе на 201 \_\_/201\_\_ учебный год.

В рабочую программу вносятся следующие изменения: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Редакция \_\_\_\_\_\_\_\_\_г. утверждена на заседании кафедры \_\_\_\_\_\_\_\_\_\_\_\_\_от \_\_.\_\_.\_\_.\_\_\_г., протокол № \_\_

Заведующий кафедрой (разработчика) \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

*подпись фамилия, инициалы*

«\_\_\_\_»\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_20\_\_г.