

Туманов В.Е.

Проектирование хранилищ данных для систем бизнес-аналитики: учебное пособие / В.Е. Туманов. — М.: Интернет-Университет Информационных Технологий: БИНОМ. Лаборатория знаний, 2010. — 615 с: ил., табл. — (Основы информационных технологий).

Содержание

Предисловие

Благодарности

Лекция 1. Хранилища данных

Концепция систем складирования данных

Концепция хранилищ данных

Типы хранилищ данных

Резюме

Лекция 2. Архитектура хранилищ данных

Введение

Основные типы программно-аппаратной архитектуры хранилища данных

Подходы в организации работ по созданию хранилища данных

Характеристика решений ведущих производителей

Типовые программно-аппаратные решения (технологические решения)

Корпоративные информационные фабрики

Хранилища данных с архитектурой шины данных

Объединенное (федеративное) ХД

Резюме

Лекция 3. Модель типового проекта создания хранилища данных

Введение

Жизненный цикл разработки хранилища данных

Модель жизненного цикла хранилища данных

Бизнес-модель типового проекта создания хранилища данных

Резюме

Лекция 4. Системы деловой осведомленности (Business Intelligence Systems) и хранилища данных

Введение

Определение систем деловой осведомленности
Информационная безопасность систем
бизнес-аналитики Технологии обработки данных
Архитектура систем бизнес-уровня
Решение от компании Microsoft
Построение систем бизнес-аналитики:
проблемы и решения
Хранилища данных и системы бизнес-аналитики
Резюме

Лекция 5. Метод моделирования «сущность-связь»

Введение
Понятие предметной области и архитектура данных
Моделирование методом «сущность-связь»
Резюме

Лекция 6. Моделирование темпоральных (временных) данных в хранилищах данных

Темпоральные данные и базы данных
Введение в моделирование темпоральных данных
Основные приемы моделирования
темпоральных данных
Резюме

Лекция 7. Метод многомерного моделирования

Основные понятия метода
многомерного моделирования
Многомерная модель
Факты
Ключи в таблицах фактов
Таблицы фактов
Измерения
Основные схемы многомерной модели
Моделирование таблиц фактов
Моделирование таблиц измерений
Резюме

Лекция 8. Создание физической модели хранилища данных

Объекты физической модели данных

Моделирование объектов физической модели хранилища данных
Разработка скрипта для создания объектов

физической модели хранилища данных

Резюме

Лекция 9. Метаданные в хранилищах данных

Метаданные

Функции метаданных в хранилище данных

Состав метаданных в хранилище данных

Логическая структура метаданных хранилища данных

Стандарты метаданных

Выбор метамоделей при проектировании хранилища данных

Проектирование логической модели метаданных хранилища данных

Резюме

Лекция 10. Проектирование и разработка процесса ETL

Введение

Подходы к реализации ETL-процесса

Разработка ETL-процесса

Элементы ETL-процесса

Проектирование процессов преобразования данных с помощью CASE-инструментов

Резюме

Лекция 11. Создание модели хранилища данных на основе корпоративной модели данных

Корпоративная модель данных

Алгоритм преобразования корпоративной модели данных в модель хранилища данных

Использование CASE-инструментов для создания многомерной модели на основе корпоративной

модели данных
Секционирование таблиц многомерной
модели данных в PowerDesigner
Резюме

Лекция 12. Метод моделирования «Свод данных»

Введение
Метод моделирования «Свод данных» (Data Vault)
Пример проектирования модели «Свод данных»
Заполнение данными объектов
«Свода данных»
Резюме

Лекция 13. Физическая модель хранилища данных: учет влияния транзакций, денормализация таблиц

Определение транзакций базы данных
Денормализация
Методы разбиения таблиц
Денормализация колонок
Резюме

Лекция 14. Создание физической модели базы данных: проектирование производительности

Введение
Повышение производительности запросов: индексы
Повышение производительности запросов:
секционирование
Повышение производительности запросов:
кластеры
Резюме

Лекция 15. SQL в хранилищах данных: агрегация и суммирование

Введение
Оператор SELECT и схема «звезда»
Расширение оператора SELECT для обработки данных
Функция GROUPING
Управление иерархией в предложениях ROLLUP и CUBE

Резюме

Лекция 16. SQL в хранилищах данных: аналитическая обработка данных

SQL для анализа данных

Агрегатные и статистические функции

Медианы

Оконные функции

Функции ранжирования

Функции, генерирующие отчеты

CASE-выражения и создание гистограмм

Статистические функции в других диалектах SQL

Резюме

Лекция 17. Настройка производительности запросов к хранилищу данных

Введение

Оптимизация запросов

Оптимизатор СУБД семейства MS SQL Server

Статистики, используемые оптимизатором
запросов СУБД MS SQL Server

Статистическая коллекция MS SQL Server

Анализ запросов с целью повышения скорости
их выполнения

Оптимизация запросов для схем типа «звезда»

Резюме

Лекция 18. Проектирование кубов данных

Введение

OLAP на клиенте и на сервере

Технические аспекты многомерного
хранения данных

Основные понятия OLAP

Проектирование кубов данных с использованием
CASE-инструментов

Генерация куба данных

Резюме

Литература

