

-

ПРОЕКТИРОВАНИЯ БД

ПОДХОДЫ В ПРОЕКТИРОВАНИИ БД

1. Классический подход к проектированию.

- Подход исходит от системы документов -на входе БД имелась одна система документов, которая при использовании БД трансформировалась в другую (выходную) систему документов (таблиц, файлов).

2. Современный подход к проектированию.

- Современный подход исходит от задач (в терминах АСУ), т.е. от приложений, под которые создается БД. Под приложением понимается программа или группа программ, предназначенных для выполнения определенных однотипных работ.

ЭТАПЫ ПРИ ПРОЕКТИРОВАНИИ БД

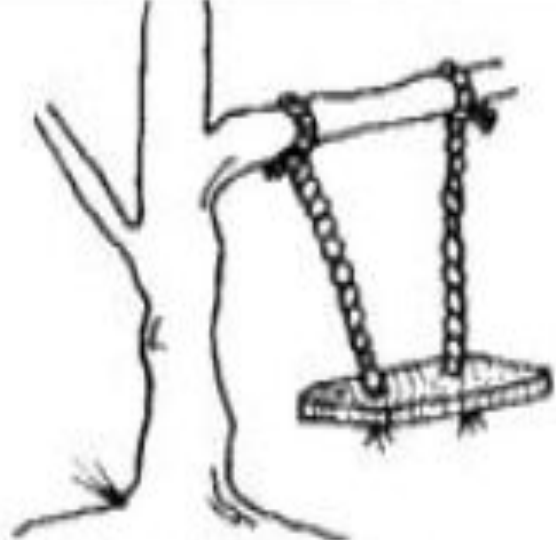
3

- **КОНЦЕПТУАЛЬНОЕ
МОДЕЛИРОВАНИЕ**(инфологическое)
- **ЛОГИЧЕСКОЕ МОДЕЛИРОВАНИЕ**
(дatalogическое)
- **ФИЗИЧЕСКОЕ МОДЕЛИРОВАНИЕ**

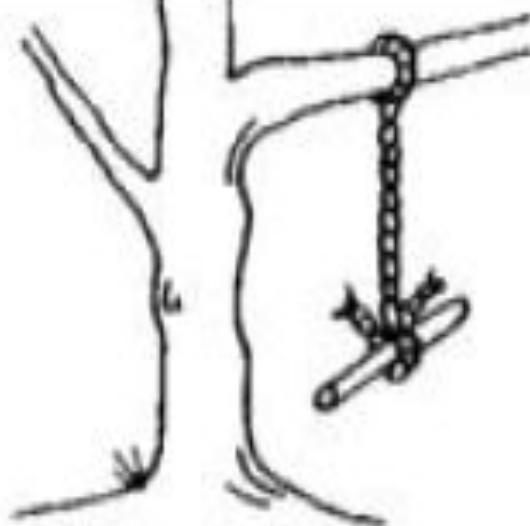
КОНЦЕПТУАЛЬНОЕ (инфологическое) ПРОЕКТИРОВАНИЕ БАЗЫ ДАННЫХ

4

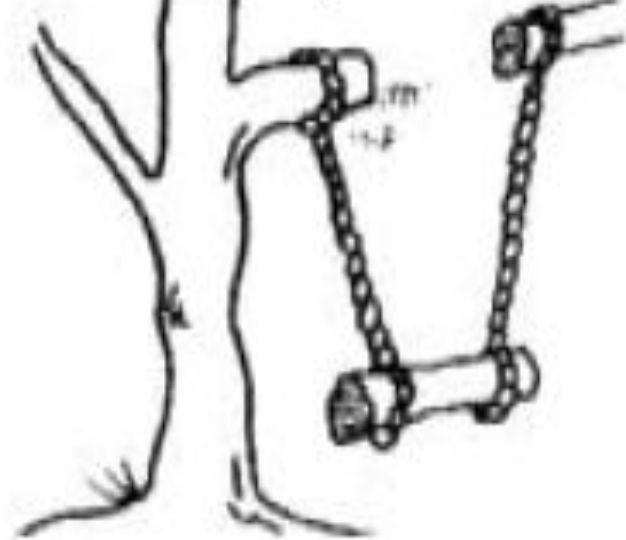
1. Создание локальной концептуальной модели данных исходя из представлений о предметной области каждого из типов пользователей.
2. Определение типов сущностей.
3. Определение типов связей.
4. Определение атрибутов, связывание их с типами сущностей, определение связей.
5. Определение доменов атрибутов.
6. Определение атрибутов, являющихся потенциальными и первичными ключами.
7. Проверка модели на отсутствие избыточности.
8. Проверка соответствия локальной концептуальной модели конкретным пользовательским транзакциям, обсуждение концептуальных моделей данных с конечными пользователями.



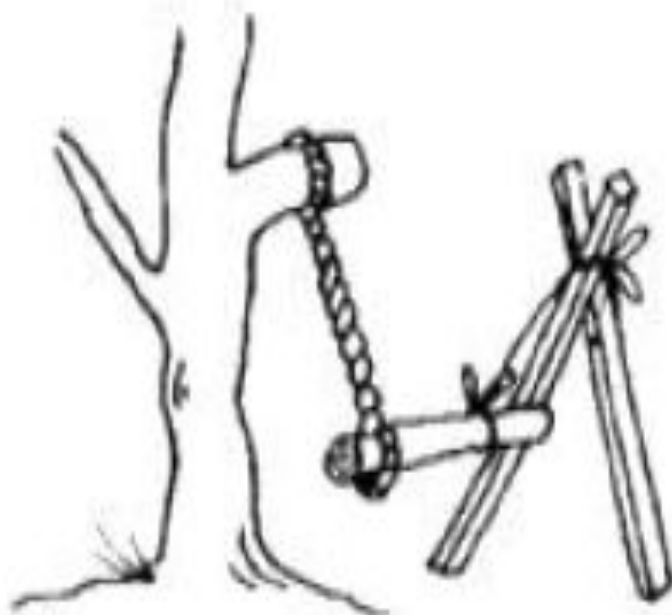
Что хотел потребитель



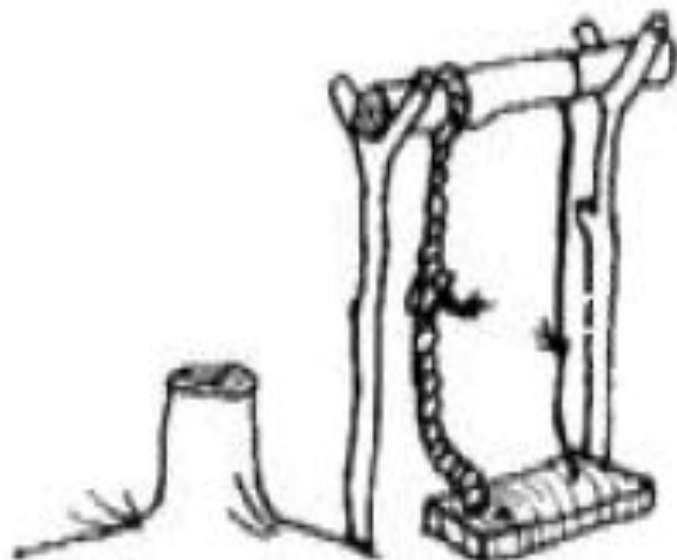
Что попросил пользователь



Как увидел аналитик



Как написал программист



Как реализовано

ПРЕДМЕТНАЯ ОБЛАСТЬ

6

Предметной областью называется часть реального мира, представляющая интерес для исследования (использования).

Описание предметной области содержит:

- цель, назначение, основные функции предприятия или организации, пользователи;
- описание входных и выходных документов, используемых при выполнении функций;
- описание всех используемых и создаваемых элементов данных;
- определение задач и запросов пользователей и их характеристик;
- направление развития.

ТЕХНИЧЕСКОЕ ЗАДАНИЕ-ТЗ

7

- В ТЗ должны быть определены основные цели приложения БД, технические требования (ТТ).
- ТТ должны содержать перечень конкретных задач, реализуемых с использованием БД.
- В разработке ТЗ участвуют инициаторы разработки проекта БД (директор или владелец предприятия).

ТЕХНИЧЕСКОЕ ЗАДАНИЕ

- **"Каковы задачи вашей компании?"**
- **"Для чего, по вашему мнению, необходимо создать базу данных?"**
- **"Почему вы думаете, что база данных поможет решить ваши проблемы?"**

ТЕХНИЧЕСКИЕ ТРЕБОВАНИЯ

9

- "Каковы ваши должностные обязанности?"
- "Какого вида задачи вы повседневно выполняете?"
- "С данными какого рода вы обычно работаете?"
- "Какого типа отчеты вы обычно используете?"
- "Дела какого типа вам необходимо отслеживать?"
- "Какие услуги предоставляет ваша

МЕТОДИКИ СБОРА ФАКТОВ О ПРЕДМЕТНОЙ ОБЛАСТИ

10

СБОР И АНАЛИЗ ТРЕБОВАНИЙ ПОЛЬЗОВАТЕЛЕЙ

- Изучение документации;
- Проведение собеседований;
- Наблюдение за работой предприятия;
- Проведение исследований;
- Проведение анкетирования.

ИЗУЧЕНИЕ ДОКУМЕНТАЦИИ

Назначение документации	Примеры полезных источников
Описывает проблему и необходимость в базе данных	Внутренние служебные записки, электронная почта, протоколы встреч, жалобы служащих/заказчиков и документы, описывающие проблему. Обзоры и отчеты о работе
Описывает задачи предприятия, связанные с рассматриваемой проблемой	Организационный график, техническое задание и стратегический план предприятия. Технические требования для изучаемой части предприятия. Описания задач/заданий. Примеры заполненных рукописных форм и отчетов. Примеры готовых компьютеризированных форм и отчетов.
Описывает существующую систему	Различные виды блок-схем и диаграмм, словарь профессиональных терминов, словарь данных. Проект приложения базы данных. Программная документация. Руководства по обучению пользователей

СОБЕСЕДОВАНИЕ

12

Преимущества	Недостатки
Позволяет опрашиваемому лицу почувствовать себя участником проекта, позволяет опрашиваемому свободно открыто отвечать на вопросы	Успех зависит от навыков общения лица, проводящего собеседование
Позволяет лицу, проводящему собеседование изменить ход опроса в ответ на неожиданные комментарии со стороны опрашиваемого лица	Успех может зависеть от желания опрашиваемых лиц участвовать в интервью
Позволяет лицу, проводящему собеседование, переформулировать или иначе построить вопросы во время собеседования	Трудоемкий и дорогой, поэтому может быть непрактичным
Позволяет лицу, проводящему собеседование, наблюдать за поведением опрашиваемого лица	

НАБЛЮДЕНИЕ

13

Преимущества	Недостатки
Позволяет убедиться в достоверности фактов и данных	Люди, находящиеся под наблюдением, могут сознательно или бессознательно вести себя иначе
Наблюдатель может наглядно видеть, что происходит	В процессе наблюдения могут остаться незамеченными действия, выполняемые при решении задач другого уровня сложности или интенсивности, некоторые задачи могут иногда выполняться с помощью способов, отличающихся от наблюдаемых
Наблюдатель может получать данные, описывающие физические условия работы	
Относительно недорогой способ сбора фактов	

ИССЛЕДОВАНИЕ

14

Преимущества	Недостатки
Позволяет сэкономить время, если решение уже существует	Может потребовать много времени
Исследователь может узнать, как другие решают подобные проблемы, или создают системы, удовлетворяющие аналогичным требованиям	Требует доступа к соответствующим источникам информации
Позволяет исследователю быть в курсе современных достижений	Исследователь может, в конечном счете, не решить проблему, поскольку такая проблема еще нигде не описана

АНКЕТИРОВАНИЕ

15

Преимущества	Недостатки
Люди могут заполнять и возвращать анкеты в удобное для них время	Не все могут согласиться ответить на вопросы анкеты;
Относительно недорогой способ сбора данных с участием большого количества людей	Анкеты могут возвращать незаполненными (безответственное, равнодушное или скептическое отношение к процессу)
Люди склонны сообщать в ответах действительные факты, если проводится анонимное анкетирование (если нет каких-то особых причин для искажения информации)	Не предоставляют возможность пояснить или переформулировать неправильно понятые вопросы
Ответы могут быть сведены в таблицу и быстро проанализированы	Нет возможности наблюдать и анализировать реакцию респондента на отдельные вопросы. Подготовка опросных листов может потребовать много времени

СБОР ИНФОРМАЦИИ О ПОЛЬЗОВАТЕЛЬСКИХ ПРЕДСТАВЛЕНИЯХ

16

Участник собеседования	Содержание собеседования
Разработчик БД	Данные какого типа должны храниться о персонале?
Менеджер	Для каждого сотрудника это — имя и фамилия, должность, пол, дата рождения и заработная плата
Разработчик БД	Что вы собираетесь делать с данными о персонале?
Менеджер	<ul style="list-style-type: none">✓ вводить данные о новых сотрудниках и удалять их при увольнении сотрудников✓ хранить данные о текущем персонале и печатать отчеты с именами и фамилиями, должностью и зарплатой каждого сотрудника в моем отделении✓ распределять сотрудников по инспекторам.✓ иногда для связи с другим отделением, нужны имена и телефоны менеджеров других отделений

СБОР ИНФОРМАЦИИ О ПОЛЬЗОВАТЕЛЬСКИХ ПРЕДСТАВЛЕНИЯХ

17

Данные	Тип доступа	Директор	Менеджер	Инспектор	Ассистент
Все отделения	Обработка				
	Запрос	x	x		
	Отчет	x	x		
Одно отделение	Обработка		x		
	Запрос		x		
	Отчет		x		
Весь персонал	Обработка				
	Запрос	x	x		
	Отчет	x	x		
Персонал одного отделения	Обработка		x		
	Запрос		x	x	

Create view as select

Grant select onto.....

СБОР ИНФОРМАЦИИ О СИСТЕМНЫХ ТРЕБОВАНИЯХ ДЛЯ ПРИЛОЖЕНИЯ БД

18

- ✓ "Какие транзакции в базе данных выполняются чаще?"
- ✓ "В какие периоды бывает низкая, нормальная и высокая загрузка по выполнению наиболее важных транзакций?"
- ✓ "Имеются ли конфиденциальные данные, к которым должны иметь доступ только определенные сотрудники?"
- ✓ "За какой прошедший период необходимо хранить данные?"
- ✓ "Какие требования к работе в сети и совместному доступу предъявляются к системе базы данных?"
- ✓ "Какого типа защиту от аварийных ситуаций или потерь данных необходимо обеспечить для приложения базы данных?"

СИСТЕМНАЯ СПЕЦИФИКАЦИЯ ДЛЯ ПРИЛОЖЕНИЯ БД

19

- ✓ начальный размер базы данных;
- ✓ темп роста базы данных;
- ✓ типы информационного поиска и их распределение по частоте использования;
- ✓ требования к работе в сети и совместному доступу;
- ✓ производительность;
- ✓ защита;
- ✓ резервное копирование и восстановление;
- ✓ юридические вопросы.

СИСТЕМНАЯ СПЕЦИФИКАЦИЯ «Риэлторское агентство» (пример)

20

Начальный размер базы данных

1. Примерно 2000 сотрудников работают в более чем 100 отделениях компании.

В среднем 20 и максимум 40 сотрудников имеются в каждом отделении.

2. Приблизительно 100 000 объектов недвижимости доступны для аренды во всех отделениях. В среднем 1000 и максимум 3000 объектов недвижимости имеются в каждом отделении.

3. Примерно 60 000 владельцев недвижимости, в среднем 600 и максимум 1000 владельцев недвижимости зарегистрированы в каждом отделении...

СИСТЕМНАЯ СПЕЦИФИКАЦИЯ «Риэлторское агентство» (пример)

21

Темп роста базы данных

1. Каждый месяц к базе данных добавляются примерно 500 новых объектов недвижимости и 200 новых владельцев недвижимости.
2. Как только объект недвижимости становится недоступным для сдачи в аренду, соответствующая запись удаляется из базы данных. Каждый месяц удаляются примерно 100 записей об объектах недвижимости.
3. Если владелец недвижимости не предоставляет для аренды объект недвижимости в течение 2 лет, запись о нем удаляется. Каждый месяц удаляются примерно 100 записей о владельцах недвижимости.
4. Каждый месяц в компанию поступают на работу и увольняются из нее приблизительно 20 сотрудников. Запись о сотрудниках удаляется через год после их увольнения. Каждый месяц удаляются примерно 20 записей о сотрудниках...

СИСТЕМНАЯ СПЕЦИФИКАЦИЯ «Риэлторское агентство» (пример)

22

Типы информационного поиска и их распределение по частоте использования

1. Поиск сведений об отделении — приблизительно 10 раз в день.
2. Поиск сведений о сотруднике отделения — приблизительно 20 раз в день.
3. Поиск сведений о конкретном объекте недвижимости — приблизительно 5000 раз в день (с понедельника по четверг), приблизительно 10 000 раз в день (с пятницы по субботу). Пик нагрузки — с 12.00 до 14.00 и с 17.00 до 19.00 ежедневно.
4. Поиск сведений о владельце недвижимости — приблизительно 100 раз в день.
5. Поиск сведений о клиенте — приблизительно 1000 раз в день (с понедельника по четверг), приблизительно 2000 раз в день (с пятницы по субботу). Пик нагрузки — с 12.00 до 14.00 и с 17.00 до 19.00 ежедневно...

СИСТЕМНАЯ СПЕЦИФИКАЦИЯ «Риэлторское агентство» (пример)

23

Требования к работе в сети и совместному доступу

1. Все отделения должны быть объединены в сеть с централизованной БД, находящейся в головном офисе компании в Глазго, с соблюдением мер защиты.
2. Система должна предоставлять возможность одновременного доступа к ней хотя бы 2 или 3 сотрудникам из каждого отделения.
3. Необходимо предусмотреть приобретение определенного количества пользовательских лицензий для обеспечения одновременного доступа к СУБД такому числу пользователей.

СИСТЕМНАЯ СПЕЦИФИКАЦИЯ «Риэлторское агентство» (пример)

24

Производительность

1. В утренние часы, но не в часы максимальной нагрузки, время ожидания ответа на поиск одной записи — менее 1 секунды. В часы максимальной загрузки время ожидания ответа на один поиск — менее 5 секунд.
2. В утренние часы, но не в часы максимальной загрузки, время ожидания ответа на поиск множества записей — менее 5 секунд. В часы максимальной загрузки время ожидания ответа на один поиск нескольких записей — менее 10 секунд.
3. В утренние часы, но не в часы максимальной загрузки, время выполнения операции обновления/сохранения — менее 1 секунды. В часы максимальной загрузки время выполнения операции обновления/сохранения — менее 5 секунд.

СИСТЕМНАЯ СПЕЦИФИКАЦИЯ «Риэлторское агентство» (пример)

25

Защита

1. База данных должна быть защищена паролем.
2. Каждому сотруднику должны быть присвоены привилегии (полномочия) доступа к базе данных согласно его пользовательскому представлению, а именно: **директора, менеджера, инспектора и ассистента.**
3. Сотруднику можно видеть только данные, необходимые для его работы, и в удобном для этого виде.

Копирование и восстановление

База данных должна копироваться **ежедневно в полночь.**

Юридические вопросы

1. В каждой стране имеются законы, регулирующие способ компьютеризированного хранения личных данных.
2. Так, если база данных содержит данные о персонале, клиентах и владельцах, необходимо изучить и учитывать любые правовые нормы, которым она должна удовлетворять.

ЛОГИЧЕСКОЕ ПРОЕКТИРОВАНИЕ

26

- Логическая модель данных учитывает особенности выбранной модели организации данных в целевой СУБД (например, реляционная).
- На этом этапе игнорируются остальные характеристики выбранной СУБД, например, любые особенности физической организации ее структур хранения данных и построения индексов.
- Для проверки правильности логической модели данных используется **метод нормализации**.

НОРМАЛИЗАЦИЯ

- **Нормализация** - процесс реорганизации данных путем ликвидации повторяющихся групп и иных противоречий в хранении данных с целью приведения таблиц к виду, позволяющему осуществлять непротиворечивое и корректное редактирование данных.
- **Нормальная форма** - совокупность требований, которым должно удовлетворять отношение.
- Управление данными становится очень простым, если данные организованы согласно нескольким правилам. Эти правила стали известными как правила нормализации- **правила Кодда**.

НОРМАЛИЗАЦИЯ 1 НФ

Таблица 1. Состав семьи (правило 1 нарушается)

п/п	Член семьи	Степень родства	Иждивенцы
1	Борис	Глава семьи	Екатерина, Александра
2	Тамара	Мать	Нет
3	Петр	Сводный брат	Ирина
4	Михаил	Дядя	Анна, Наталья
5	Иван	Дядя	Елена

Таблица 2. Состав семьи (в первой нормальной форме)

п/п	Член семьи	Степень родства	Иждивенец1	Иждивенец2
1	Борис	Глава семьи	Екатерина	Александра
2	Тамара	Мать	Нет	Нет
3	Петр	Сводный брат	Ирина	Нет
4	Михаил	Дядя	Анна	Наталья
5	Иван	Дядя	Елена	Нет

НОРМАЛИЗАЦИЯ 2НФ

Определение 1. Пусть R - отношение. Множество атрибутов Y функционально зависит от множества атрибутов X (X функционально определяет Y) тогда и только тогда, когда во всех кортежах, имеющих одинаковые значения атрибутов X , значения атрибутов Y также совпадают в любом состоянии отношения R . Символически функциональная зависимость записывается

$$X \rightarrow Y$$

Множество атрибутов X называется детерминантом функциональной зависимости, а множество атрибутов Y называется зависимой частью.

- Таблица содержит ключ
- Все неключевые столбцы зависят от полного ключа

Составной ключ

Курс	Студент	Тип обучения	Оценка	ак. часов
10774	Е. Онегин	очно-заочное	отлично	40
10775	В. Ленский	вебинар	хорошо	40
10776	Т. Ларина	очное	отлично	40
10777	В. Ленский	очное	хорошо	40

НОРМАЛИЗАЦИЯ 3 НФ

- В фокусе внимания – неключевые столбцы
- Нет транзитивных зависимостей
 - Неключевые столбцы не зависят от других неключевых столбцов

№ заказа	Покупатель	Товар	Цена (шт.)	Кол-во	Сумма
1001	Е. Онегин	телефон	100	2	200
1002	В. Ленский	ноутбук	2000	1	2000
1003	Т. Ларина	утюг	50	1	50
1004	В. Ленский	утюг	50	2	100

НОРМАЛИЗАЦИЯ НФ БОЙСА-КОДДА (НФБК)

- Отношение R находится в НФ Б-К тогда и только тогда, когда **детерминанты всех функциональных зависимостей являются потенциальными ключами.**
- Если отношение находится в НФБК, то оно автоматически находится и в 3НФ.

ПРИВЕДЕНИЕ К 4 НФ

32

Четвёртая нормальная форма (4NF)

- Отношение находится в четвёртой нормальной форме, если она находится в нормальной форме Бойса-Кодда и все нетривиальные многозначные зависимости фактически являются функциональными зависимостями от её потенциальных ключей.
- Переменная отношения R находится в четвёртой нормальной форме тогда и только тогда, когда в случае существования таких подмножеств A и B атрибутов этой переменной отношения R , для которых выполняется нетривиальная многозначная зависимость $A \twoheadrightarrow B$, все атрибуты переменной отношения R также функционально зависят от A .



ПРИВЕДЕНИЕ К 4 НФ

33

Курсы, аудитории и преподаватели

- Проектируем базу для автоматизации учебного центра
- Основные справочники:
 - Курсы
 - Аудитории
 - Преподаватели
- Эти три сущности «сходятся» вместе во время занятий



ПРИВЕДЕНИЕ К 4 НФ

34

А вот и аномалии... ☹

№	Курс
1	Датамайнинг на платформе Microsoft
2	Анализ данных на SQL – уровень 1
3	Анализ данных на SQL – уровень 2
4	Анализ данных на языке Р
5	Язык запросов MDX
6	Углублённое изуч

№	Преподаватель
1	Михаил Гурьянов
2	Фёдор Самородов
3	Сергей Минюров

№	Аудитория
1	РА5
2	ТГ12
3	РА11
4	МС6

Как быть, если к курсу нужно подготовить аудитории, но преподаватель ещё не назначен?

Что если для курса пока ещё не подготовлена ни одна аудитория?

Если эта аудитория больше не используется для курса, то не пропадёт ли информация о преподавателе?

Курс	Преподаватель	Аудитория
2	NULL	1
2	1	NULL
3	3	3

ПРИВЕДЕНИЕ К 4 НФ

35

Многозначные зависимости

Один курс
несколько

**Несколько многозначных зависимостей
в одной таблице – это проблема!**

Один курс могут читать
несколько преподавателей

▪ (Курс)

▪ (Курс) →> (Преподаватель)

Преподаватель

Одна строка определяет
несколько зависимостей

**Несколько многозначных зависимостей
в одной таблице – это проблема!**

Один курс могут читать
несколько преподавателей

Один курс может проводиться
в разных аудиториях

▪ (Курс) →> (Преподаватель)

▪ (Курс) →> (Аудитория)

Преподаватель	Курс	Аудитория
1	2	1
1	2	2
2	2	3
2	2	4

Одна строка определяет сразу
несколько зависимостей

Хотим изменить одну
зависимость – неизбежно
затронем и все остальные!

ПРИВЕДЕНИЕ К 4 НФ

36

- В таблице должно быть не больше одной многозначной зависимости
- Если таких зависимостей несколько, надо выносить их в отдельные таблицы
- Не ошибитесь с многозначными зависимостями!
 - От курса зависит набор преподавателей, способных читать этот курс
 - От курса зависит набор пригодных для его проведения аудиторий
 - От курса зависит набор доступных режимов обучения
 - От курса зависит список экзаменов, к которым готовит этот курс
 - **Преподаватели, аудитории и режимы обучения не зависят друг от друга**

ПРИВЕДЕНИЕ К 1 НФ

37

Накладная № 123				
<u>Дата</u>	<u>Покупатель</u>	<u>Адрес</u>		
10.01.2001	ТОО "Суперпулс"	г. Кукуюевск ул. Большая Трубная д.6		
<u>Отпущен товар</u>	<u>Количество</u>	<u>Ед. Изм.</u>	<u>Цена за ед.</u>	<u>Общая стоимость</u>
Банка стеклянная	100	шт.	3,45	345
Стакан граненый	34	шт.	1,34	45,56
Бутылка "чебурашка"	367	шт.	0,45	165,15
Вода минеральная	40	б	7,85	314
Водка "Столичная"	25	б	10,90	262,5
Пиво "Амур ДВ"	40	б	4,56	182,4

ОТПУСК ТОВАРОВ СО СКЛАДА
Дата
Покупатель
Город
Адрес
Товар
Ед_Измерения
Цена_за_ед
Отпущено_ед
Общая_стоимость
Номер_накладной

ПРИВЕДЕНИЕ К 2 НФ

38

Накладная № 123				
<u>Дата</u>		<u>Покупатель</u>		<u>Адрес</u>
10.01.2001		ООО "Суперпулс"		г. Куккуевск ул. Большая Трубная д.6
<u>Отпущен товар</u>	<u>Количество</u>	<u>Ед. Изм.</u>	<u>Цена за ед.</u>	<u>Общая стоимость</u>
Банка стеклянная	100	шт.	3,45	345
Стакан граненый	34	шт.	1,34	45,56
Бутылка "чебурашка"	367	шт.	0,45	165,15
Вода минеральная	40	бутылка	7,85	314
Водка "Столичная"	25	бутылка	10,50	262,5
Пиво "Амур ДВ"	40	банка	4,56	182,4

ОТПУСК ТОВАРОВ СО СКЛАДА
Дата
Покупатель
Город
Адрес
Товар
Ед_Измерения
Цена_за_ед
Отпущено_ед
Общая_стоимость
Номер_накладной



ОТПУСК ТОВАРОВ СО СКЛАДА
Номер_накладной
Товар
Дата
Покупатель
Город
Адрес
Ед_Измерения
Цена_за_ед
Отпущено_ед
Общая_стоимость

ПРИВЕДЕНИЕ К 2 НФ

39

ОТПУСК ТОВАРОВ СО СКЛАДА
Номер_накладной
Товар
Дата
Покупатель
Город
Адрес
Ед_Измерения
Цена за ед
Отпущено_ед
Общая_стоимость



Выделение таблицы "Товар"

ОТПУСК ТОВАРОВ СО СКЛАДА
Номер_накладной
Товар (FK)
Дата
Покупатель
Город
Адрес
Отпущено_ед
Общая_стоимость

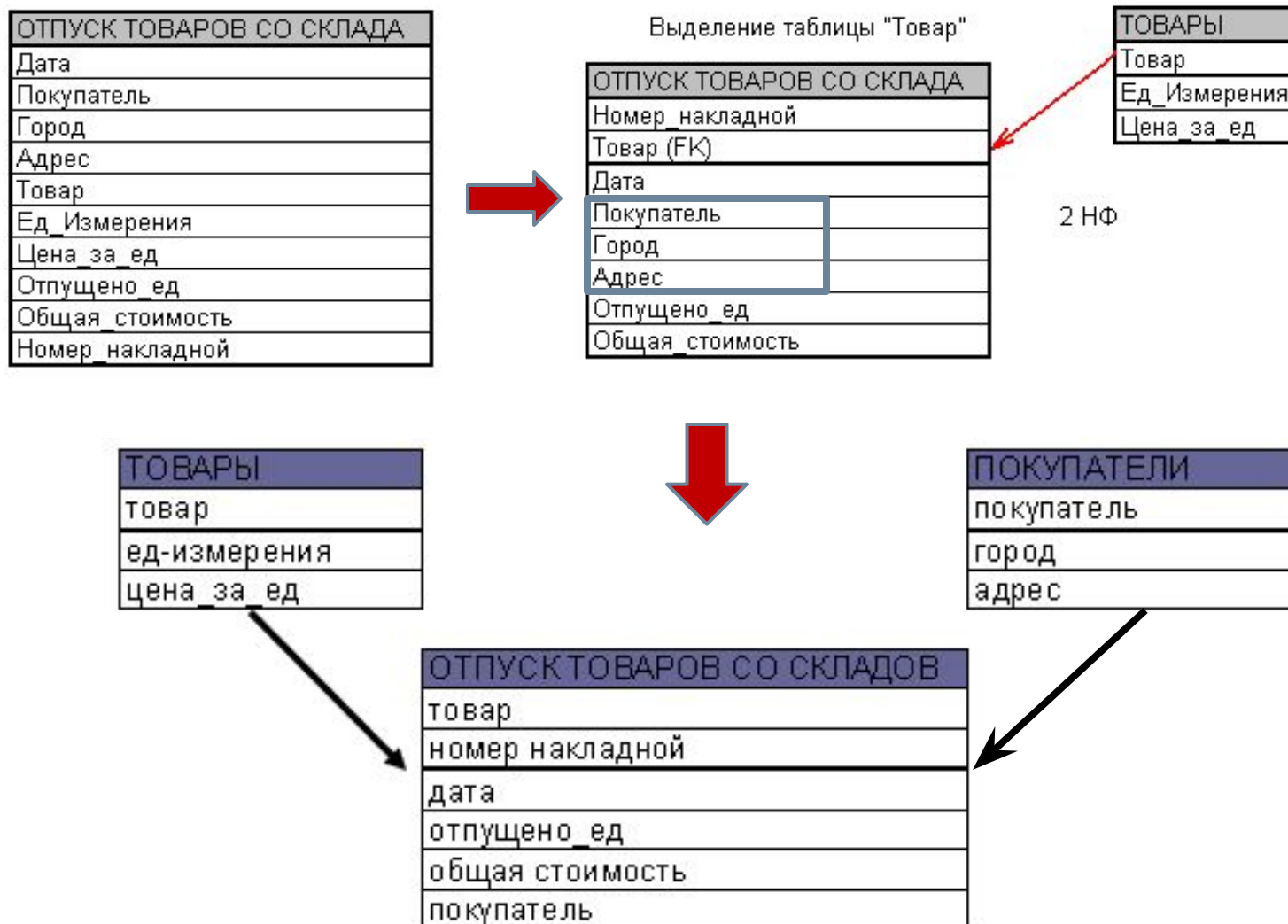
2 НФ

ТОВАРЫ
Товар
Ед_Измерения
Цена за ед



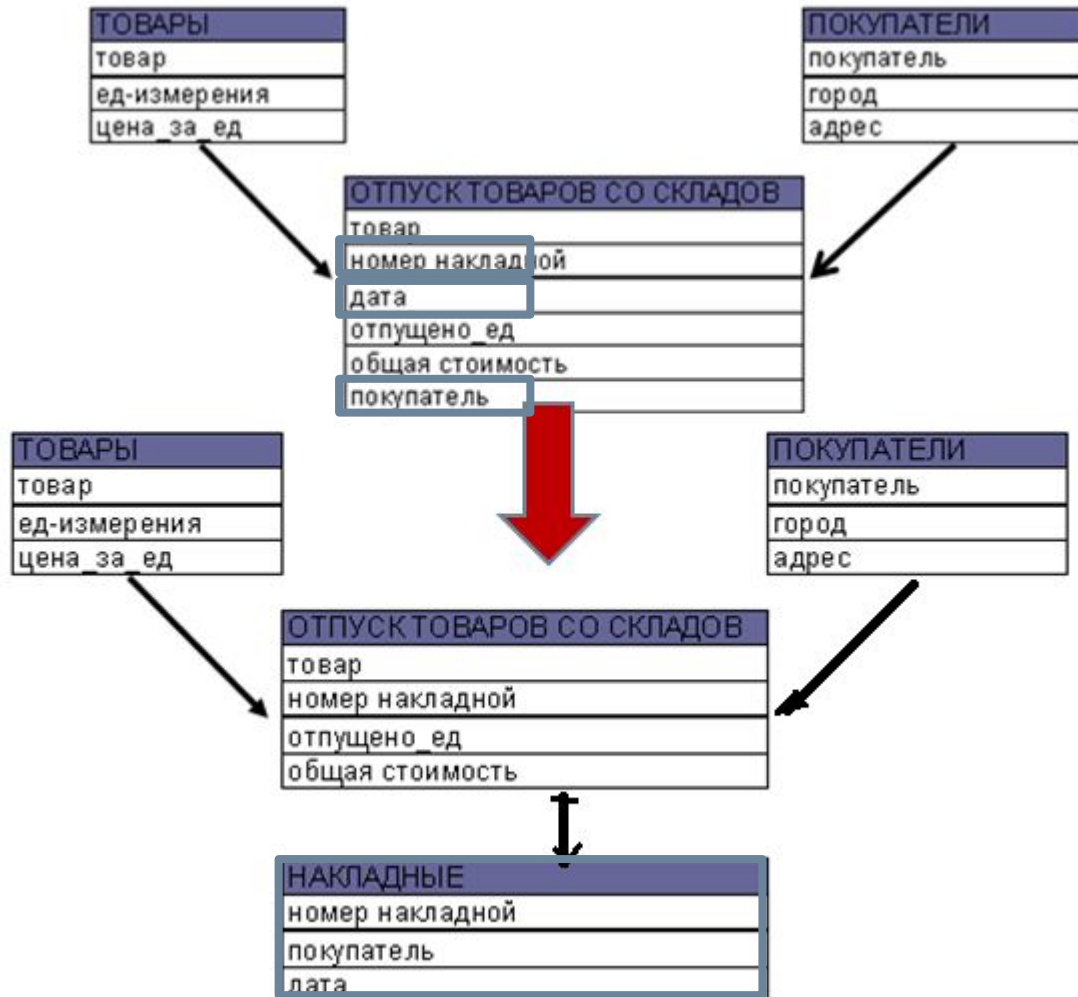
ПРИВЕДЕНИЕ К 3 НФ

40



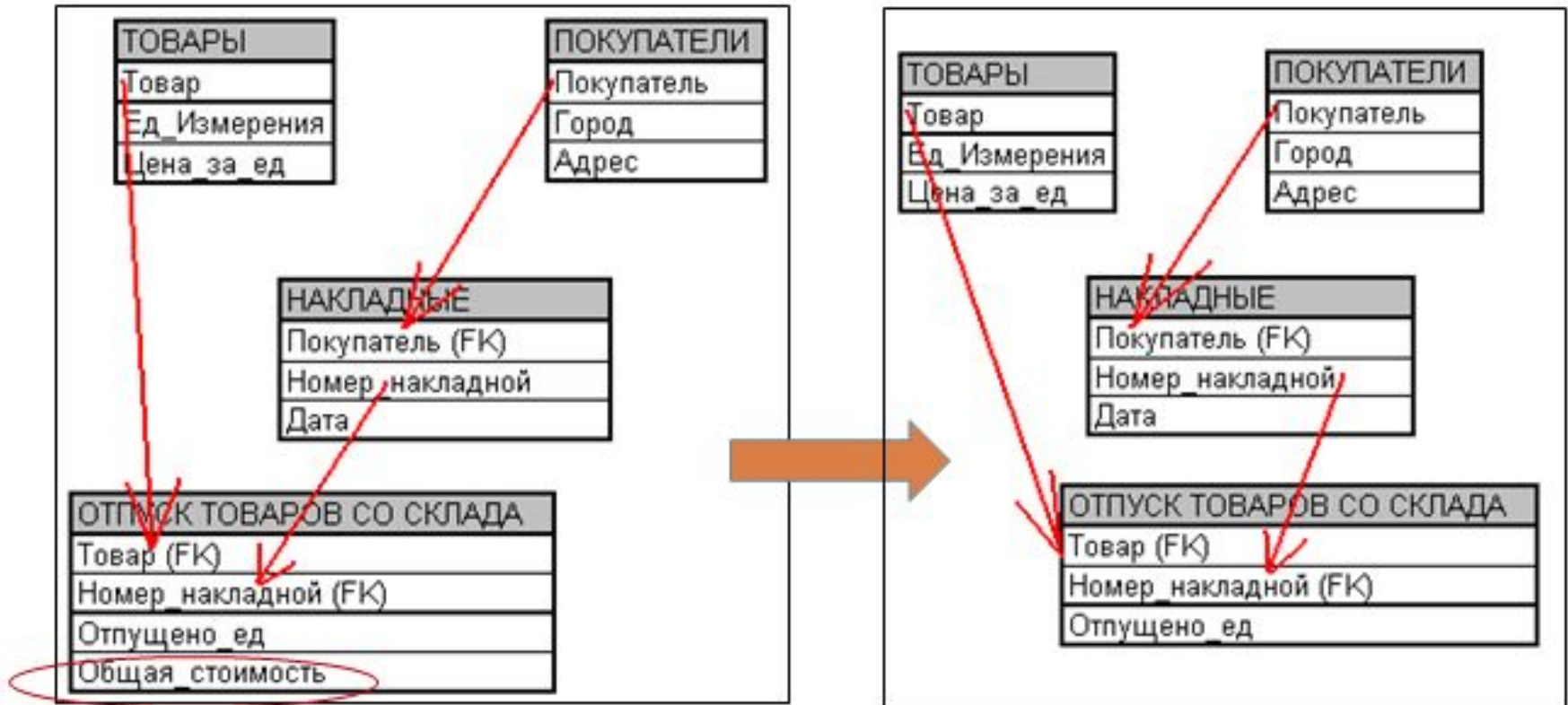
ПРИВЕДЕНИЕ К 2 НФ

41



ПРИВЕДЕНИЕ К 3 НФ

42



НЕДОСТАТКИ НОРМАЛИЗАЦИИ

43

- Большое количество сущностей БД. Сопровождение и поддержка такой БД сложна
- Трудности построения запросов к таким БД, так как необходимо связывать несколько таблиц.
- Оперативность **выборки данных низкая для высоко нормализованных БД (3 НФ).**

OLTP И OLAP-СИСТЕМЫ

- Сильно нормализованные модели данных хорошо подходят для OLTP-приложений (**On-Line Transaction Processing (OLTP)**-оперативная обработка транзакций)
- OLAP-приложения (**On-Line Analytical Processing (OLAP)** - оперативная аналитическая обработка данных) используют слабо нормализованные модели данных

OLTP-ПРИЛОЖЕНИЯ

45

- Поддерживает большое число пользователей, работающих параллельно
- Большое значение имеет время ответа на запрос
- OLTP-системы сопряжены с интенсивными процессами чтения-записи
- Примеры OLTP-приложений - системы складского учета, системы заказов билетов, банковские системы, выполняющие операции по переводу

OLAP-ПРИЛОЖЕНИЯ

46

- оперируют с большими массивами данных
- добавление в систему новых данных происходит относительно редко крупными блоками (например, раз в квартал загружаются данные по итогам квартальных продаж из OLTP-приложения)
- добавленные данные в систему обычно никогда не удаляются
- перед загрузкой данные проходят различные процедуры "очистки", связанные с тем, что в одну систему могут поступать данные из многих источников, имеющих различные форматы

OLAP-ПРИЛОЖЕНИЯ

47

- запросы к системе являются нерегламентированными, достаточно сложными.
- скорость выполнения запросов важна, но не критична.

ФИЗИЧЕСКОЕ ПРОЕКТИРОВАНИЕ БД

48

- Проектирование базовых отношений в среде целевой СУБД, отношений, содержащих производные данные.
- Реализация ограничений предметной области.
- Проектирование физического представления БД
- Анализ транзакций.
- Выбор файловой структуры.
- Определение индексов.

ФИЗИЧЕСКОЕ ПРОЕКТИРОВАНИЕ БД

- Определение требований к дисковой памяти.
- Разработка пользовательских представлений.
- Анализ необходимости введения контролируемой избыточности.
- Организация мониторинга и настройка функционирования ОС.
- Разработка средств и механизмов защиты.

ФИЗИЧЕСКОЕ ПРОЕКТИРОВАНИЕ БД

50

- выбор типа носителя, способа организации данных, методов доступа (**определение пользователей базы данных, их уровней доступа, разработка и внедрение правил безопасности доступа**),
- определение размеров физического блока, управление размещением данных на внешнем носителе,
- управление свободной памятью, определение целесообразности сжатия данных и используемых методов сжатия,

ФИЗИЧЕСКОЕ ПРОЕКТИРОВАНИЕ БД

- оценка размеров объектов базы (определение размеров табличных пространств и особенностей их размещения на носителях информации,
- определение спецификации носителей информации для промышленной системы (например, тип raid-массивов, их количество),
- разработка топологии базы данных в случае распределенной базы данных, определение механизмов доступа к удаленным данным.

СПАСИБО ЗА ВНИМАНИЕ!