

БАЗЫ ДАННЫХ

Автор: Емельянов Н. Е.

Правка: Тригуб Н.А.

Литература

- [А. Д. Хомоненко, В. М. Цыганков, М. Г. Мальцев.](#) Базы данных. Учебник для ВУЗов. – Корона-принт. 2004. - 737 с.
- К. Дж. Дейт. Введение в Системы баз данных, 8-е издание – Вильямс. 2006 – 1315 с.
- Джеймс Р. Грофф, Пол Н. Вайнберг, Эндрю Дж. Оппель. [SQL: полный справочник, 3-е издание.](#) Диалектика-Вильямс. 2010. - 960 с.
- Роберт Виейра - Программирование баз данных MS SQL Server 2008. Базовый курс. Вильямс. 2009. – 816 с.
- Емельянов Н.Е. Введение в СУБД ИНЕС // Наука, М., 1990, 256 с.
- Структуры данных // МИСИС №958

1. Основные понятия

Определение

Реляционная база данных – совокупность изменяющихся во времени нормализованных отношений (relation)

Главный постулат РСУБД – все данные во всех отношениях атомарны

<i>Интерпретация</i>							
Номер рейса	Дни недели	Пункт отправления	Время вылета	Пункт назначения	Время в пути	Тип самолета	Стоимость билета
<i>Данные</i>							
138	2_4_7	Баку	15:20	Москва	3ч40мин	ИЛ-86	115.00
57	3_6	Ереван	0:45	Киев	2ч50мин	ТУ-154	92.00
1234	2_6	Казань	13:20	Баку	1ч55мин	ТУ-134	73.50
242	1 по 7	Киев	23:30	Москва	3ч10мин	ТУ-154	57.00
86	2_3_5	Минск	14:45	Сочи	1ч25мин	ИЛ-86	78.50
137	1_3_6	Москва	18:40	Баку	3ч40мин	ИЛ-86	115.00
241	1 по 7	Москва	2:30	Киев	3ч10мин	ТУ-154	57.00
577	1_3_5	Рига	13:50	Таллин	40мин	АН-24	21.50
78	3_6	Сочи	11:40	Баку	3ч40мин	ТУ-134	44.00
578	2_4_6	Таллин	9:00	Рига	40мин	АН-24	21.50

Рисунок 1. К разделению данных и их интерпретации

По запросу

ВЫБРАТЬ КОЛИЧЕСТВО

(Номер_рейса)

ИЗ ТАБЛИЦЫ Расписание

ГДЕ Пункт_отправления = 'Москва'

И Пункт_назначения = 'Минск';

получим количество рейсов "Москва-Минск".

Основная особенность СУБД – это наличие процедур для ввода и хранения не только самих данных, но и описания их структуры. Файлы, снабженные описанием хранимых в них данных и находящиеся под управлением СУБД, стали называть «**банки данных**» (БнД) и «**базы данных**» (БД).

СУБД должна предоставлять доступ к данным любым пользователям, включая и тех, которые практически не имеют и/или не хотят иметь представления:

- о физическом размещении данных в памяти ЭВМ и их описаний;
- о механизмах поиска запрашиваемых данных;
- о проблемах, возникающих при одновременном запросе одних и тех же данных многими пользователями (прикладными программами);
- о способах обеспечения защиты данных от некорректных обновлений и (или) несанкционированного доступа;
- о поддержании баз данных в актуальном состоянии
- о множестве других функций СУБД.

<i>Компоненты СУБД</i>		<i>обеспечивают</i>
<i>Языки описания данных (ЯОД)</i>	=====	Описание структур файлов БД, записей файлов и их полей данных
<i>Языки работы с БД (ЯрБД)</i>	=====	Получение ответа на санкционированный запрос пользователя или БД-приложения
<i>Общие утилиты для БД</i>	=====	Выполнение общих процедур по поддержанию БД в актуальном состоянии
<i>Генераторы БД-приложений</i>	=====	Создание модульных систем решения задач пользователя по работе с БД
<i>Генераторы отчетов</i>	=====	Вывод результатов работы в виде отчетов

Рисунок 2. Принципиальный состав основных компонент СУБД.

Языковые средства СУБД (две основные группы):

1. языки *описания данных* (**ЯОД**)

2. языки *работы с БД* (**ЯрБД**).

ЯОД предназначены для описания структур данных и отношений между ними, поддерживаемых СУБД.

ЯОД можно классифицировать на:

- языки *описания БД*

/ предназначены для описания состава и логической организации БД */*

- языки *описания внешних данных*.

/ для описания внешней, по отношению к БД информации (входной, выходной, сообщений и т. д.). */*

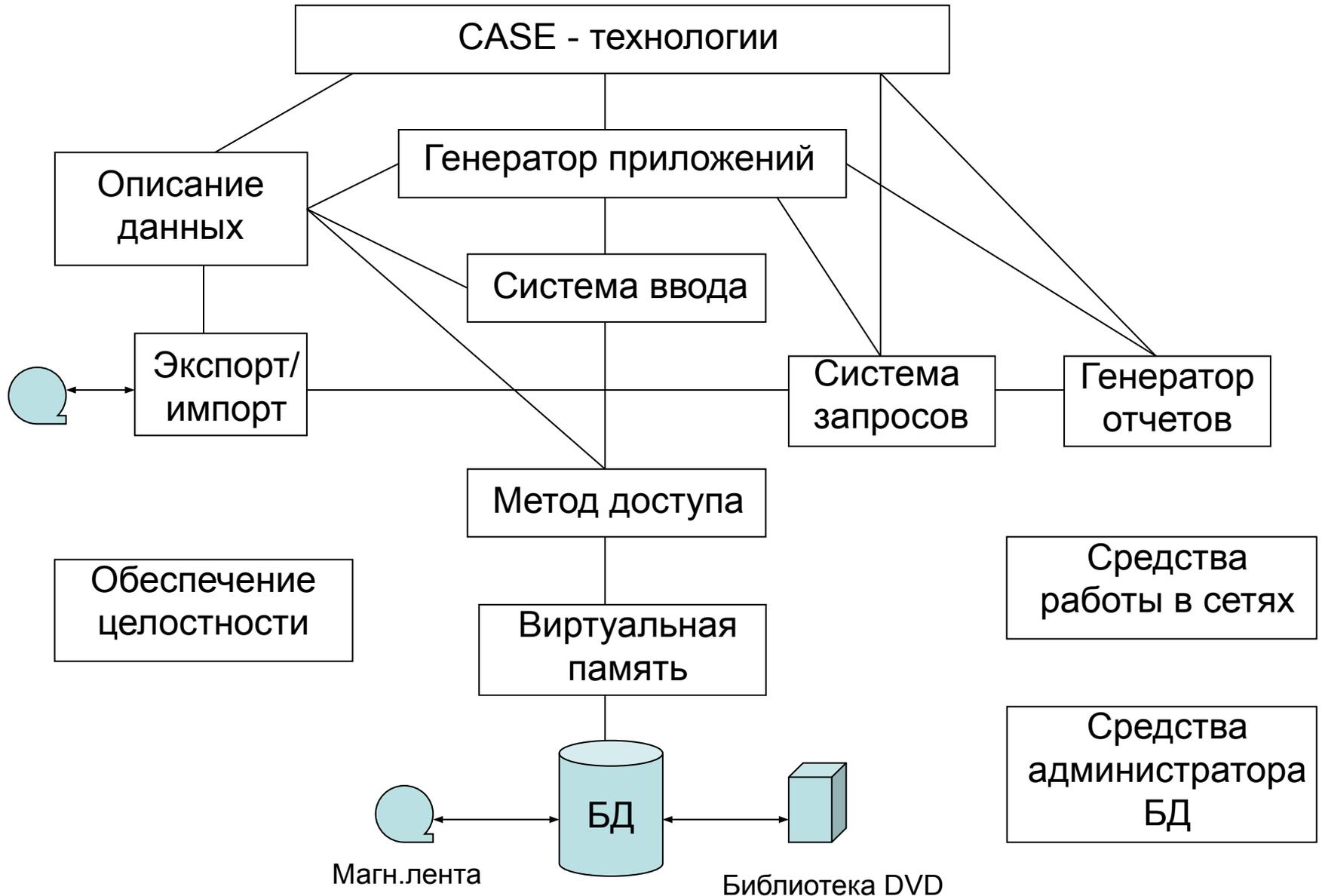
Среди лингвистических средств СУБД **центральное место** занимают **ЯрБД**.

Языки данной группы позволяют не только организовывать запросы из БД нужной информации, но и программировать БД-приложения.

ЯрБД по их функциональному назначению можно классифицировать по трем группам:

1. *управления данными (ЯУД),*
2. *управления запросами (ЯУЗ),*
3. *организации диалоговых режимов.*

Состав функционально полной СУБД



Функционально-полную СУБД можно определять следующим составом функциональных характеристик:

1. Поддерживаемая системой *даталогическая* модель;
2. Средства *администратора* БД;
3. Средства разработки *БД-приложений*;
4. Интерфейсы с пользователями и другими *БД-приложениями*;
5. Интерфейсы с другими СУБД;
6. Средства обеспечения *сетевой* и *распределенной* обработки информации.

2. Модели данных. Классификация СУБД

Модель данных (Data Model) =
структура данных + методы доступа

Классические модели

- Иерархическая м. (1965 – 1970)
- Сетевая м. (1970 – 1975)
- Реляционная м. (1975 – 1980)

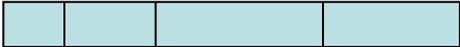
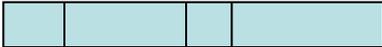
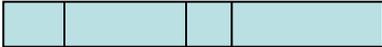
Новые модели

- Семантическая м. (1980 – 1985)
- Типово-полная м. (1985 – 1990)
- Объектно-ориент. (1990 – 1995)

Новейшие модели

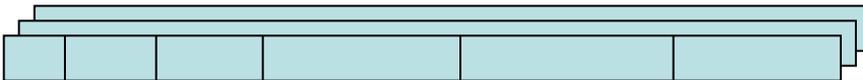
- Многомерная м.
- XML-DB (с ~ 2000 г.)

2.1. Иерархическая модель

- СУБД IMS (фирмы IBM)
- ОКА (для ЕС ЭВМ)
- Структура данных: файлы с разнотипными записями
- Обл. 1 
- Район 1.1 
- Город 1.1.1 
- Город 1.1.2. 
-
- Район 1.2 

2.2. Сетевая модель

- СУБД IDMS (комитета CODASYL)
- Рабочие группы по БД (РГБД - DBTG)
- СЕТЬ (для ЕС ЭВМ)
- Структура данных: 1) файлы с однотипными записями
- Обл. 1, Обл. 2, ..., Обл. N
- 
- Район 1.1, Район 1.2, ..., Район N.Mi
- 
- Город 1.1.1, Город 1.1.2, ..., Город N.Mi.Kj



- Структура данных: 2) наборы (SET) содержащие ссылки на подчиненные записи хозяин (OWNER) → член набора (MEMBER)



2.3. Реляционная модель

- СУБД ORACLE (фирмы ORACLE), MS SQL (фирмы MicroSoft)
- Автор - математик Код (Codd)
- Пальма (для ЕС ЭВМ)

- Структура данных: файлы с однотипными записями

Понятия

- Отношение (relation)
- Кортеж $\langle d_1, d_2, \dots, d_k \rangle$
- Домен $d_i \in D$

Понятия

- Отношение (relation)
- Таблица (двумерный массив)
- Кортеж $\langle d_1, d_2, \dots, d_k \rangle$
- Запись
- Домен $d_i \in D$
- Множество допустимых значений

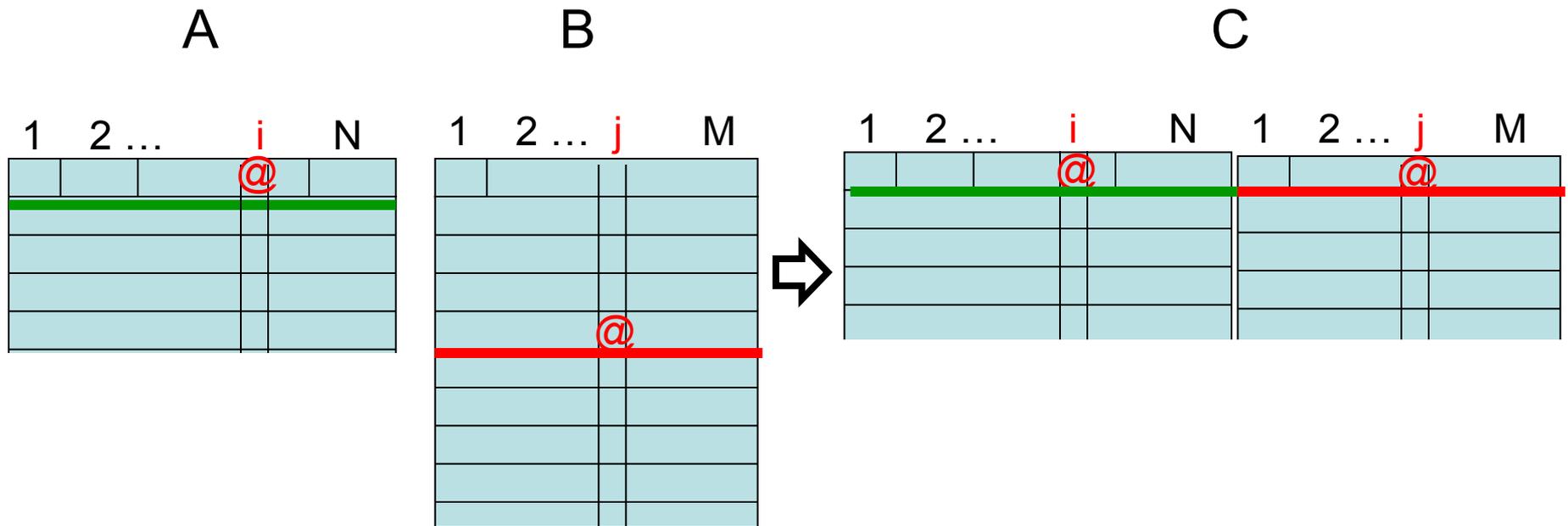
Определение

- Реляционная модель – совокупность изменяющихся во времени нормализованных отношений
- Нормализованное отношение – все данные атомарны

		*		

d_{ij} — атомарное
данное

- Пример двуместной операции (JOIN):



- Достоинства

- 1) Простая математическая модель

- Недостатки

- 1) Утрачена семантика данных

2.4. Семантическая модель

Возникла из описания семантики языков

Требования:

- a) Все объекты в процессе работы могут быть связаны
- b) Логическая близость приводит к близкому расположению
- c) В процессе работы можно менять схему БД

2.5. Типово-полная модель

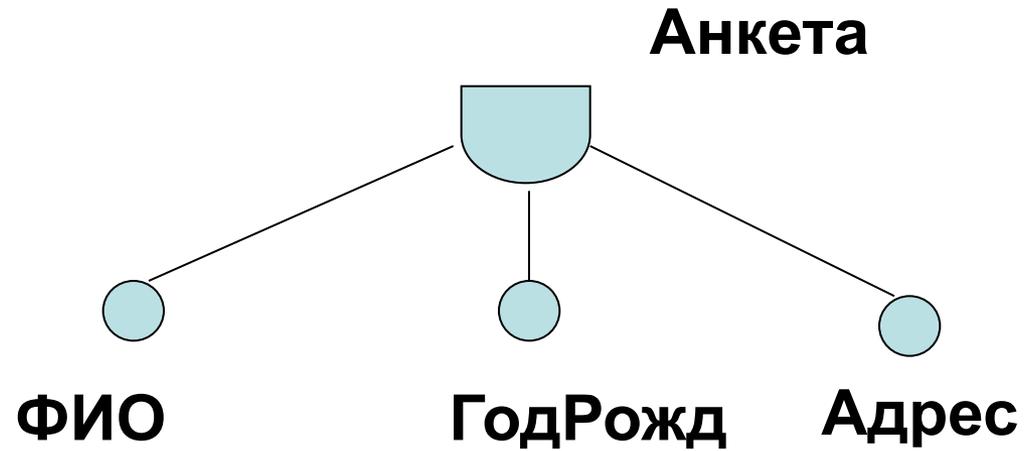
Теоретический подход к описанию
сложных структур

Основные понятия:

- a) Простые типы данных (TEXT, INT, REAL, DATA...)
- b) Конструкторы новых типов из простых и ранее описанных типов
- c) Возможна любая суперпозиция типов

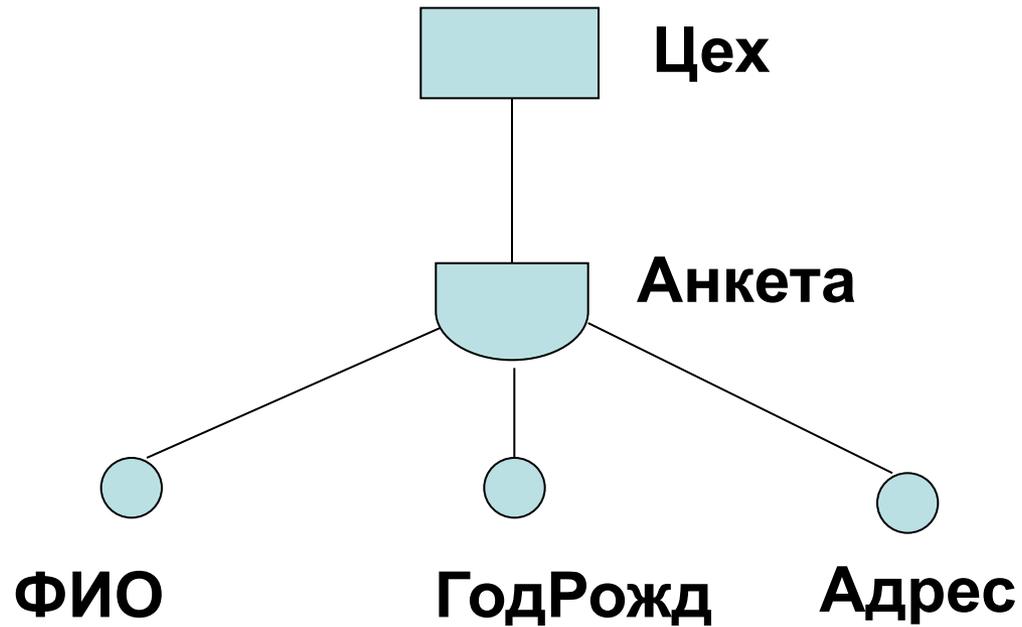
Конструктор агрегации

(Структура)

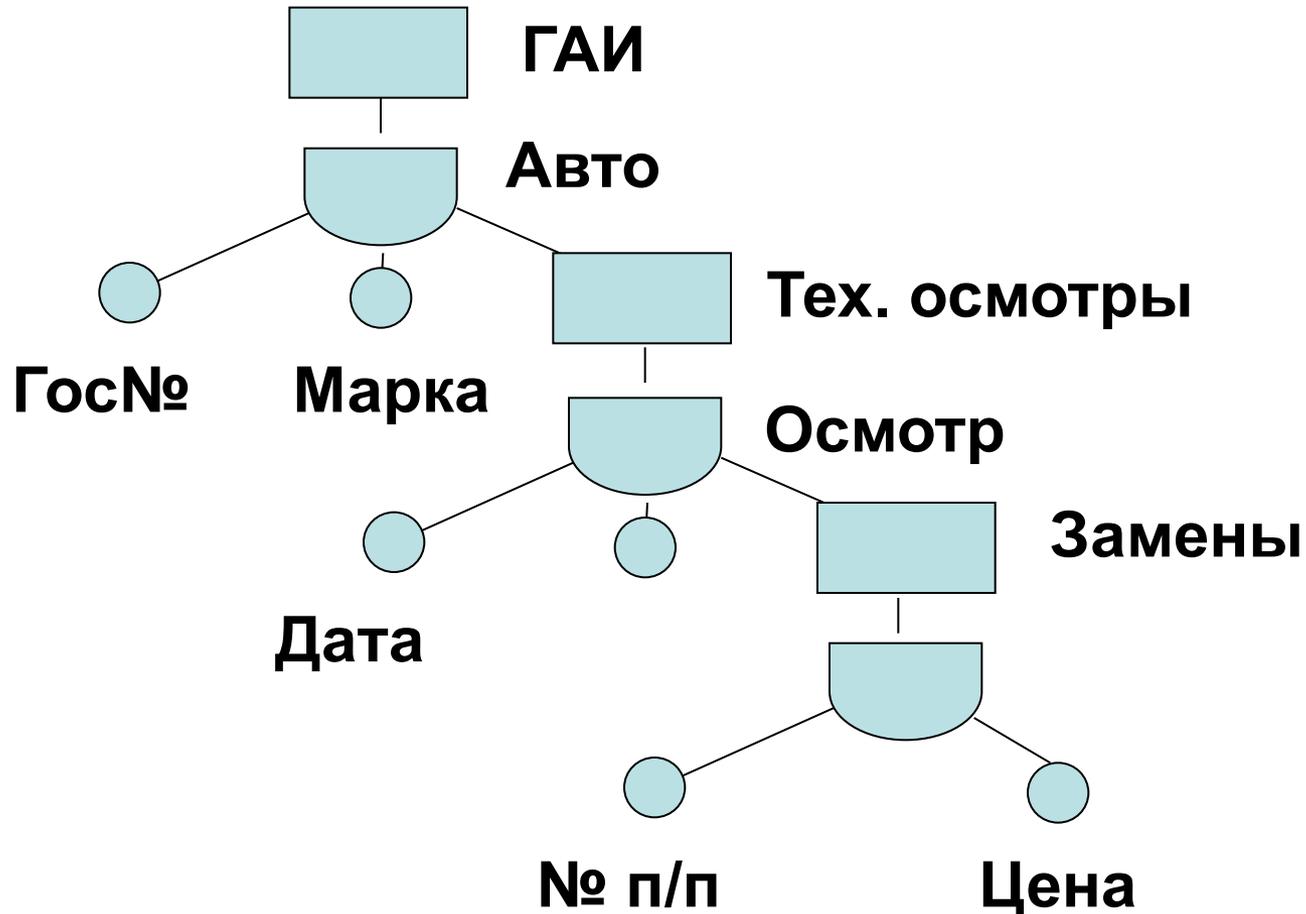


Конструктор ассоциации

(Массив)



Пример суперпозиции типов



Нет ограничений на ширину и глубину

2.6. Объектно-ориентированная модель

На основе объектного программирования

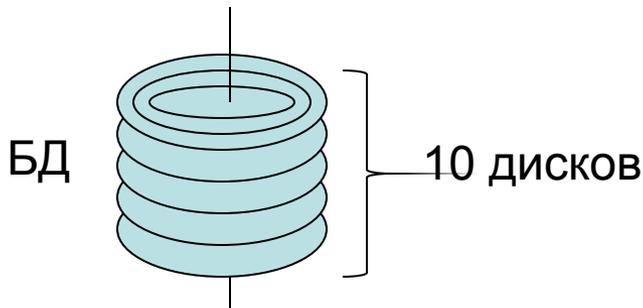
Основные понятия:

- a) Многократное использование типов
- b) Иерархия типов
- c) Инкапсуляция – смесь описания данных (типов) и методов работы с ними

2.7. Машины БД

- СУБД относятся к техническим наукам, так как зависят от существующей техники
- Всегда стояла задача разработки специального оборудования для работы с БД.
- Например. Можно ли за один оборот диска найти все записи, содержащие слово «сталь»?

Оборот за $1/50$ сек.

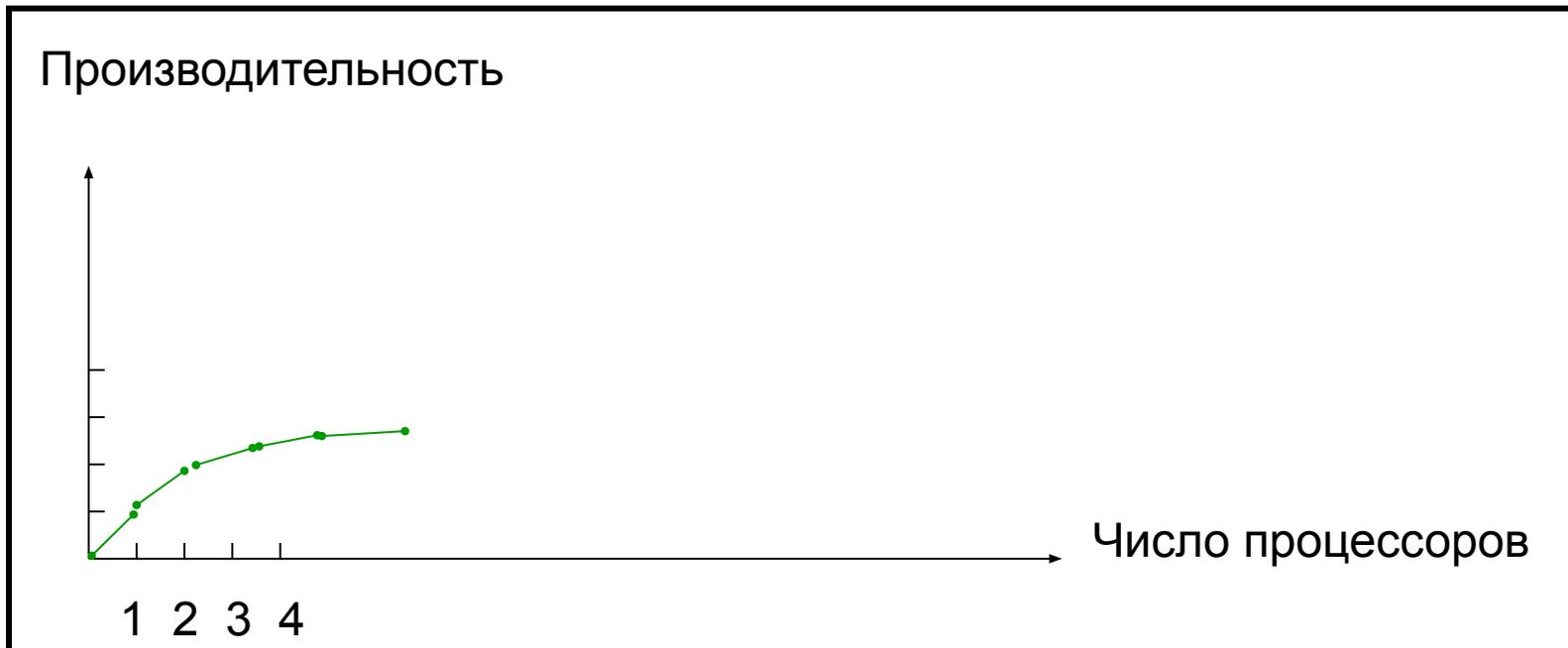


ДА. Можно!

Нужно поставить головки с процессорами на все дорожки всех дисков.

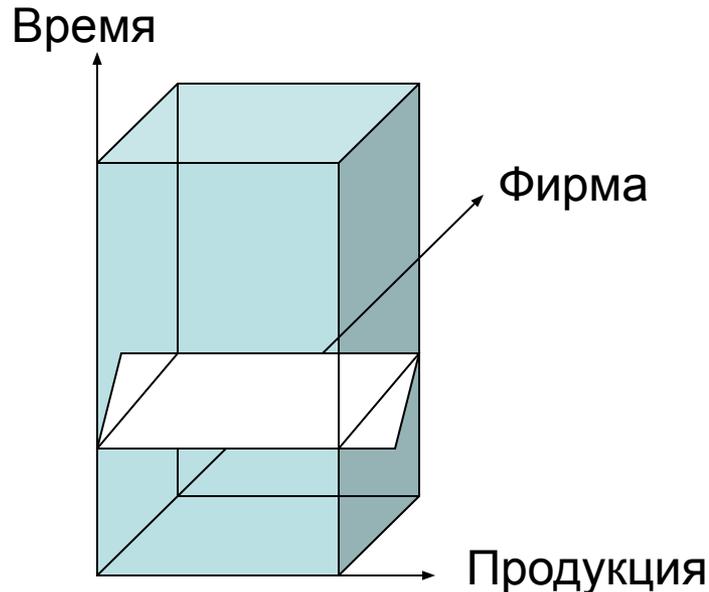
СУБД - Teradata (Фирма NCD)

- Ориентирована на многопроцессорные комплексы
- Основная проблема:



2.8. Многомерный куб

- Для систем оперативного анализа - OLAP (On Line Analytical Processing)
- База данных – многомерный куб

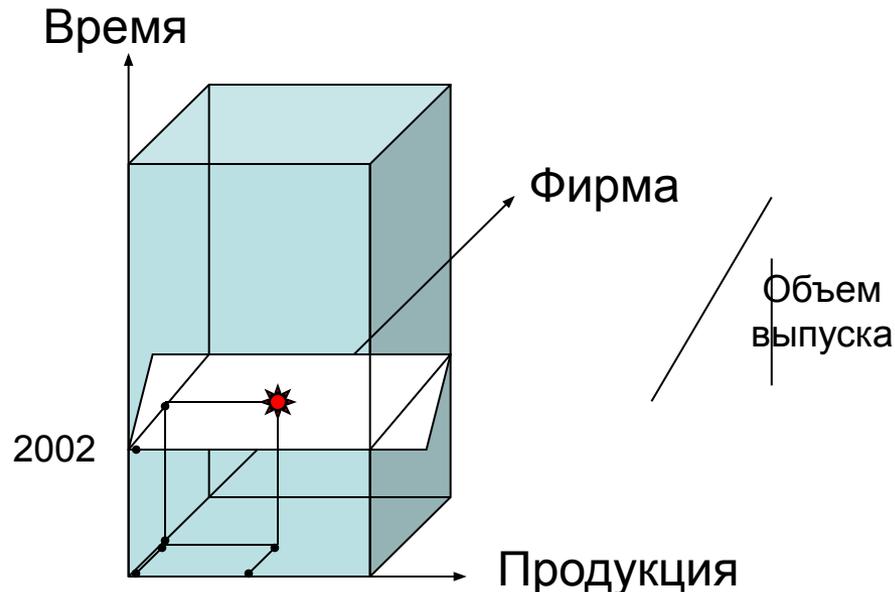


Требования

- **Возможность агрегации** (Например: Область, район, город)
- **Историчность** (есть временная ось)
- **Статичность** (данные в многомерном кубе не изменяются во время проведения анализа)

Основные понятия

- Измерение (dimension)
- Ячейка (cell)
- Срез (slice) подмножество - куб меньшей размерности



Пример

1. Фирмы

- BMW
 - Филиалы по странам
 -
- Ford
 - Филиалы по странам
 -

2. Продукция

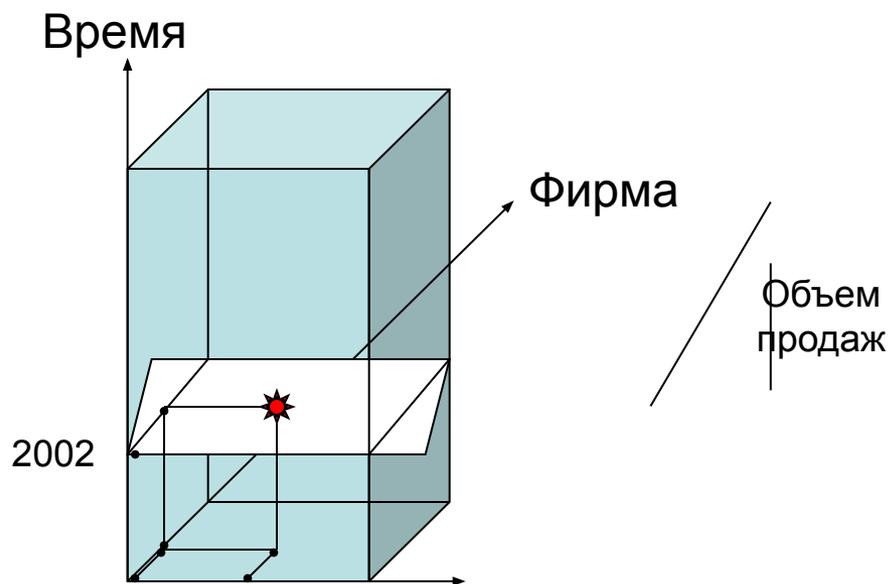
- Грузовики
 - Легкие
 - Тяжелые
 -
- Автобусы
 - Микроавтобусы
 -

3. Время

- Год
 - Январь
 - Февраль
 - Март
 - Апрель
 -

Пример

Ячейки (cell) – объем продаж в денежном выражении



2.9. XML - DB

- Для описания сложных структур данных на языке XML (Extensible Markup Language)
- База данных содержит XML-документы:

```
<имя объекта1>  
  <имя подоб1.1>  
    <имя данного 1> данное 1 </имя данного 1>  
    <имя данного 2> данное 2 </имя данного 2>  
  
    .....  
  </имя подоб1.1>  
  
  .....  
</имя объекта1>
```

2.10. XML data type.

Революция в реляционных СУБД
(2005 – 2006 г.)

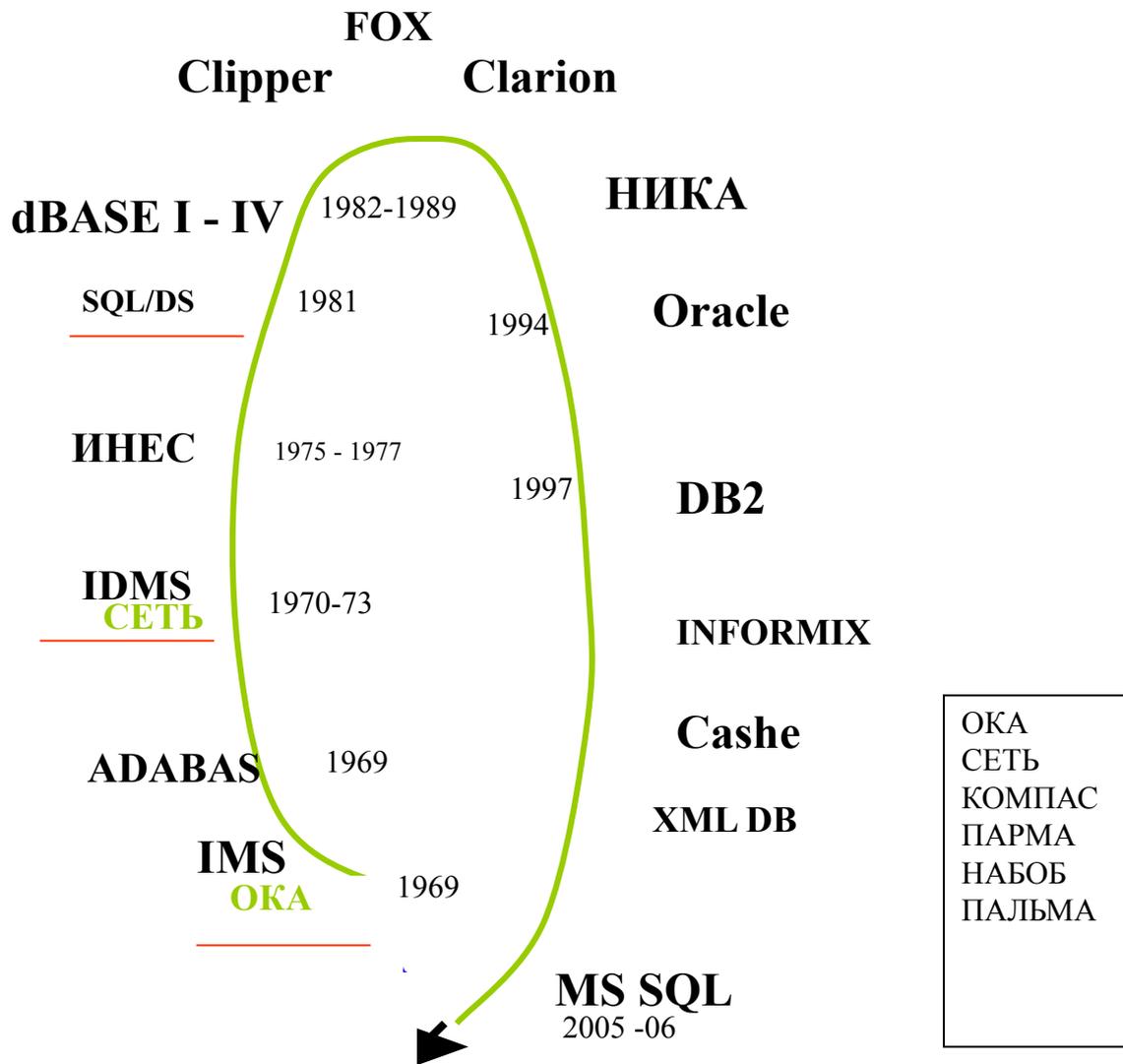
Что произошло?

Введен новый тип данных
XML data type

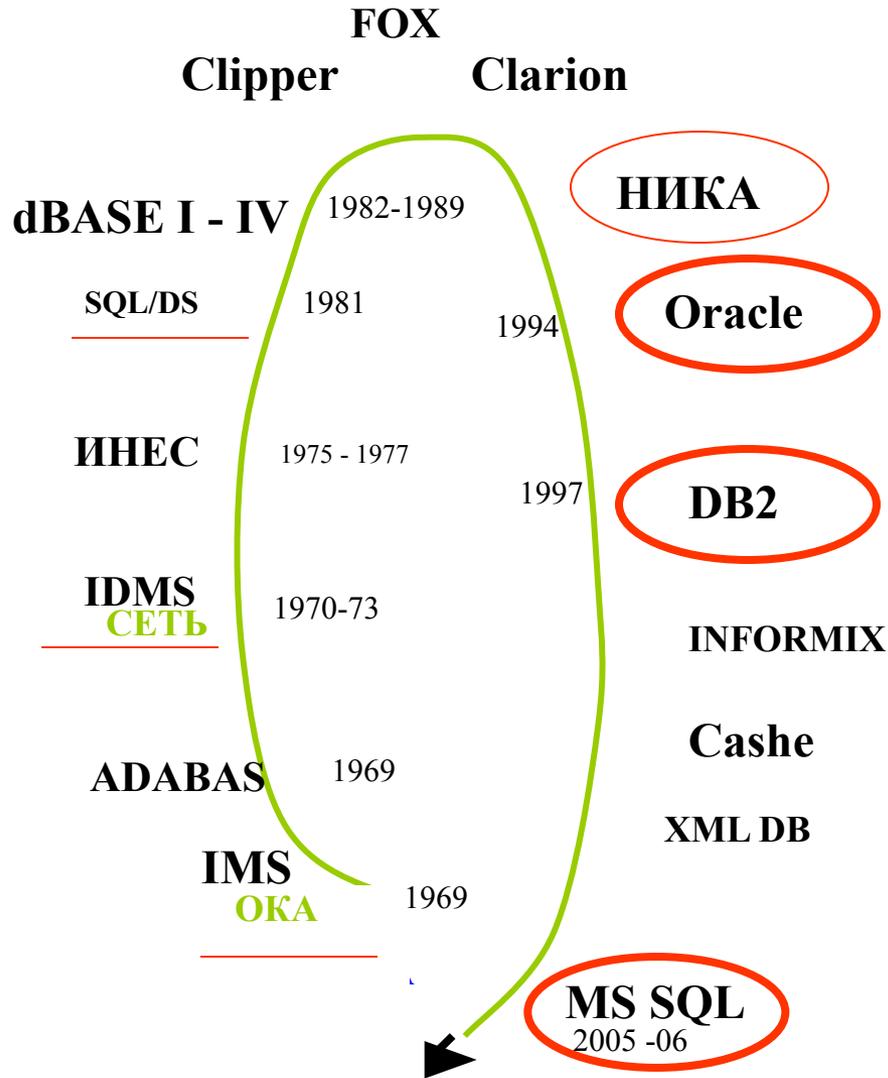
Теперь один реквизит может содержать
сколь угодно сложно структурированное
данное - вплоть до всей БД!

В один день устарели миллионы учебников по
СУБД, написанные на всех языках мира

Почему это случилось? Немного истории



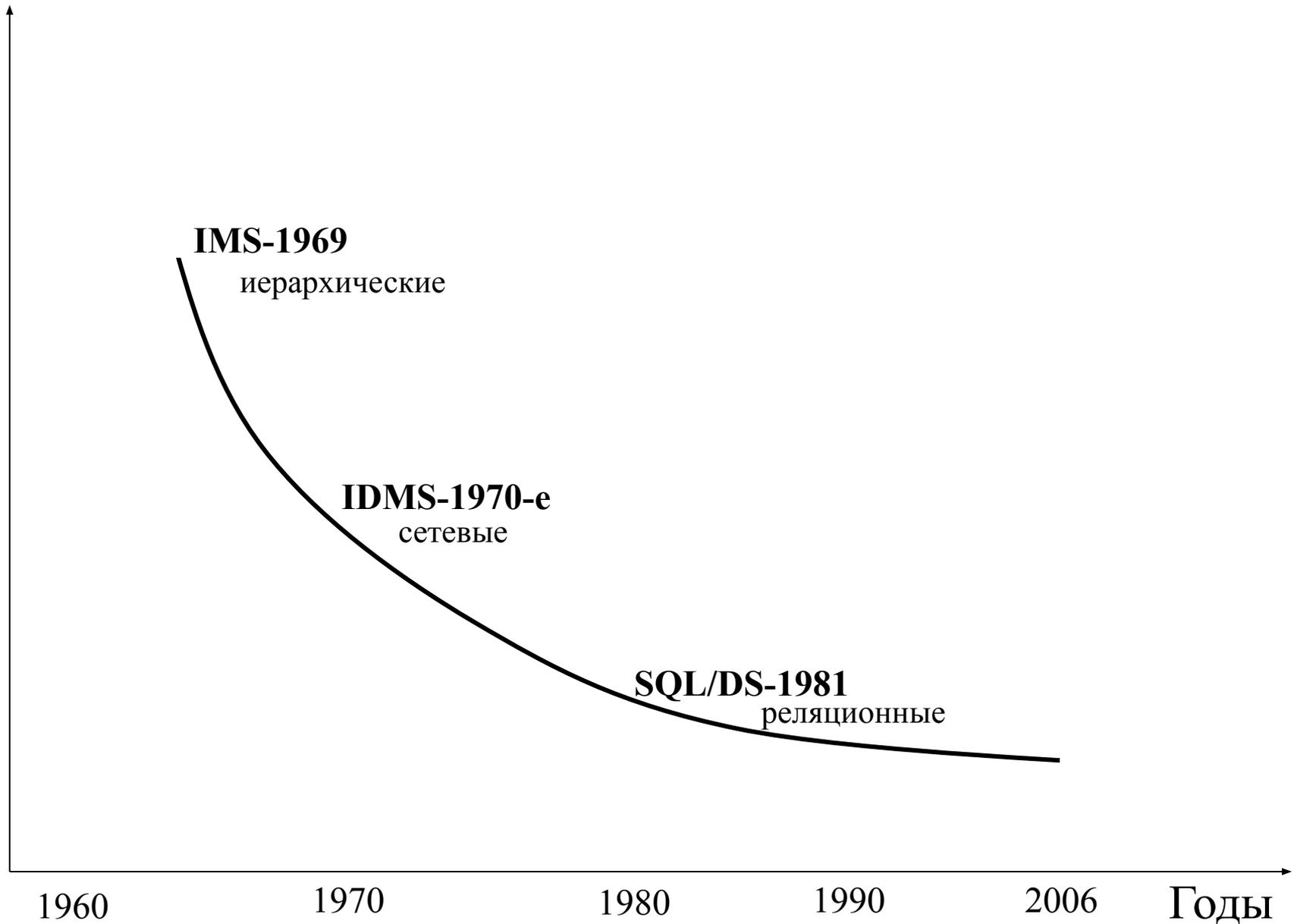
Более 1500 СУБД — в нашем справочнике для РФФИ



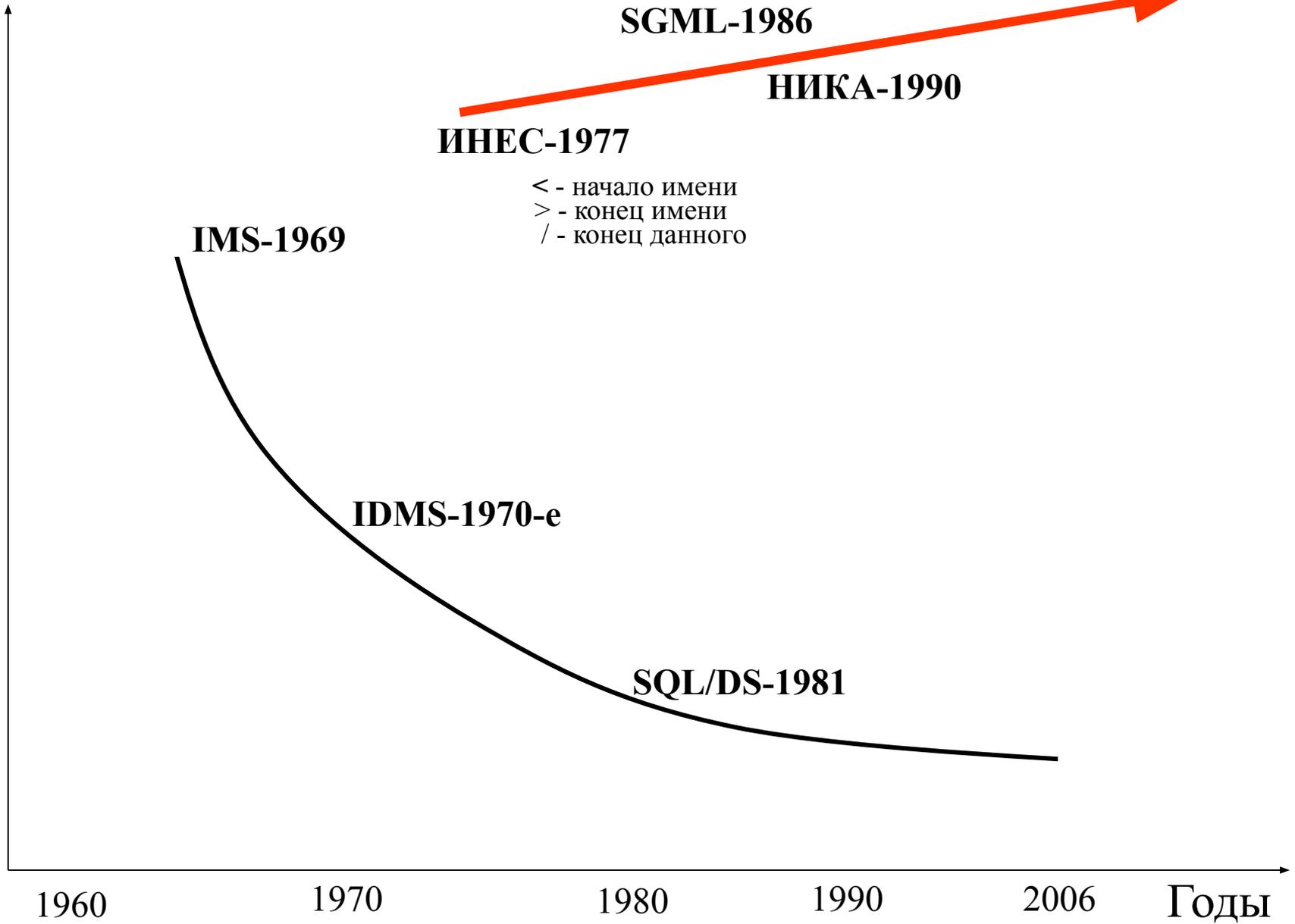
- | |
|-------------------|
| НИКА |
| ОКА |
| СЕТЬ |
| КОМПАС |
| ПАРМА |
| НАБОБ |
| ПАЛЪМА |

Сложность
данных

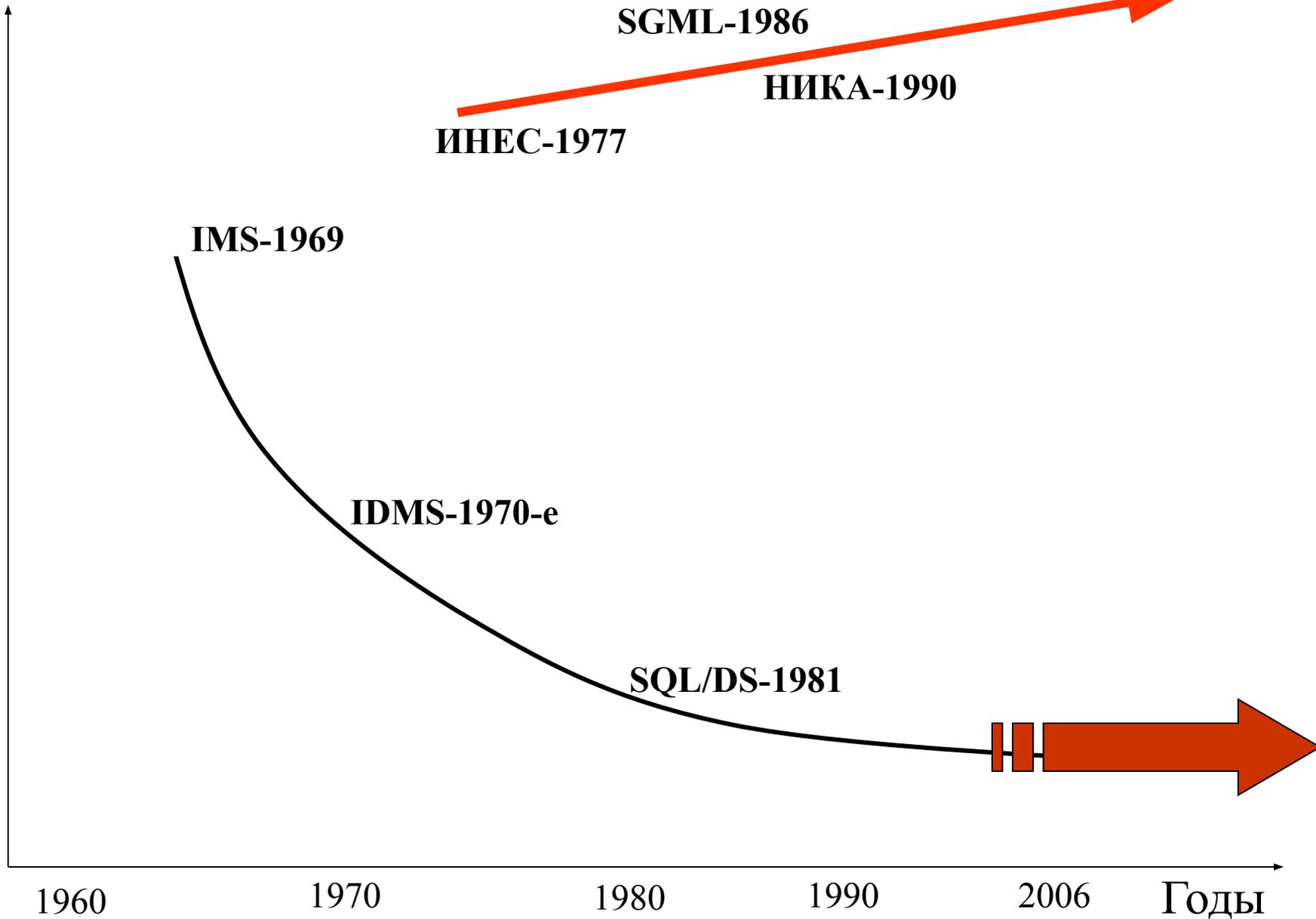
«Развитие» СУБД в последние 30 лет



Сложность
данных



Сложность
данных



Что не нравилось
пользователям и
разработчикам?

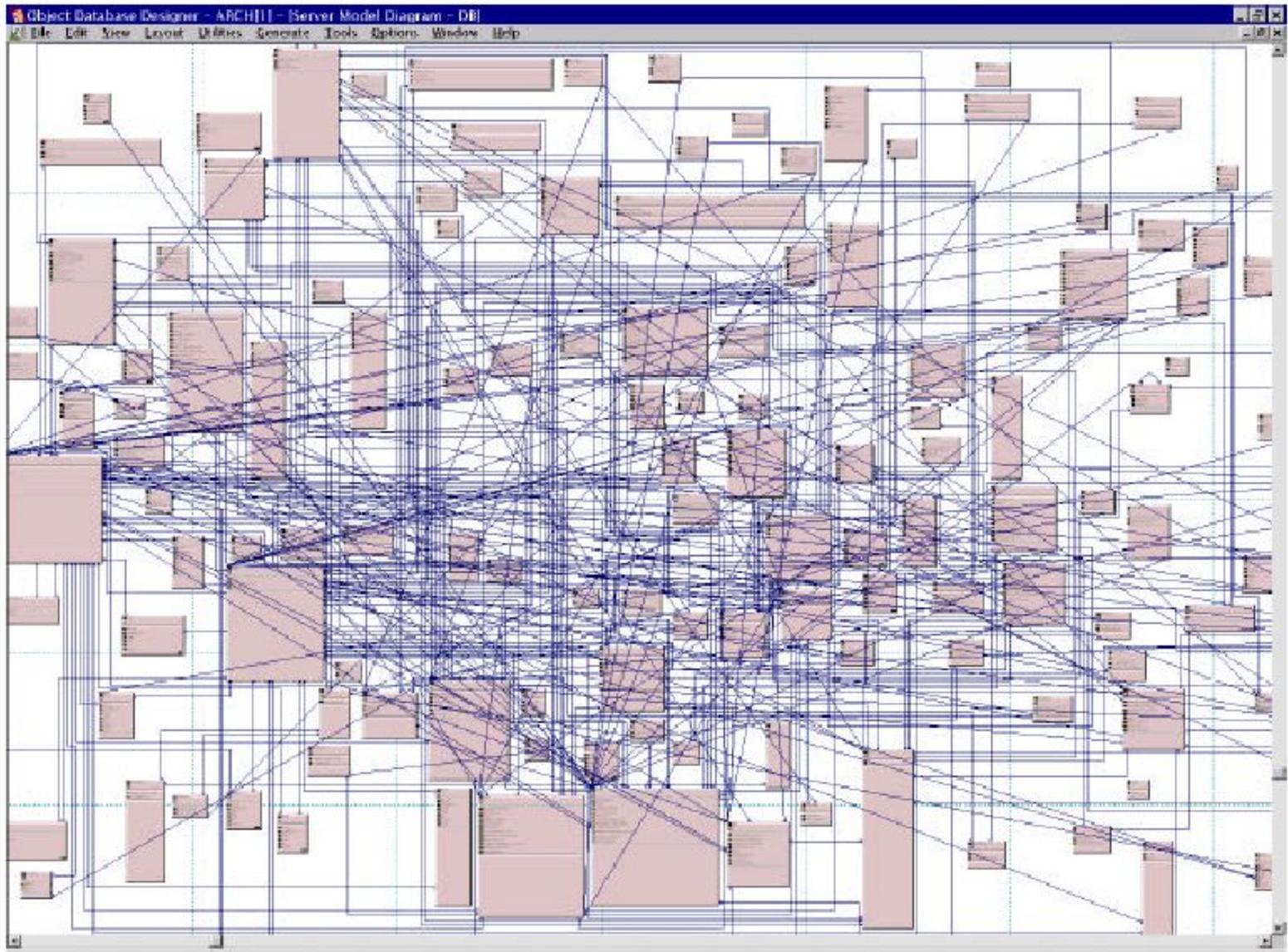


Fig.7 Object-relational Database generated after the CIDOC model

While the object-relational databases give more straightforward implementation of the object-oriented data model, they are also less

time it gives much wider range of opportunities for software development.

Революция (контрреволюция) в реляционных СУБД

(2005 – 2006 гг.)

XML хорош, если :

- Данные разрежены
- Структура данных может существенно изменяться
- Возможна рекурсия в описании данных

Если нет ни одного из этих условий,
ТО МОЖНО (но не обязательно) ИСПОЛЬЗОВАТЬ РБД

XML Best Practices (<http://msdn2.microsoft.com>)

Oracle XML Milestones

10gR2

10g

9iR2

9i

8i

10^g

XDK
XSU

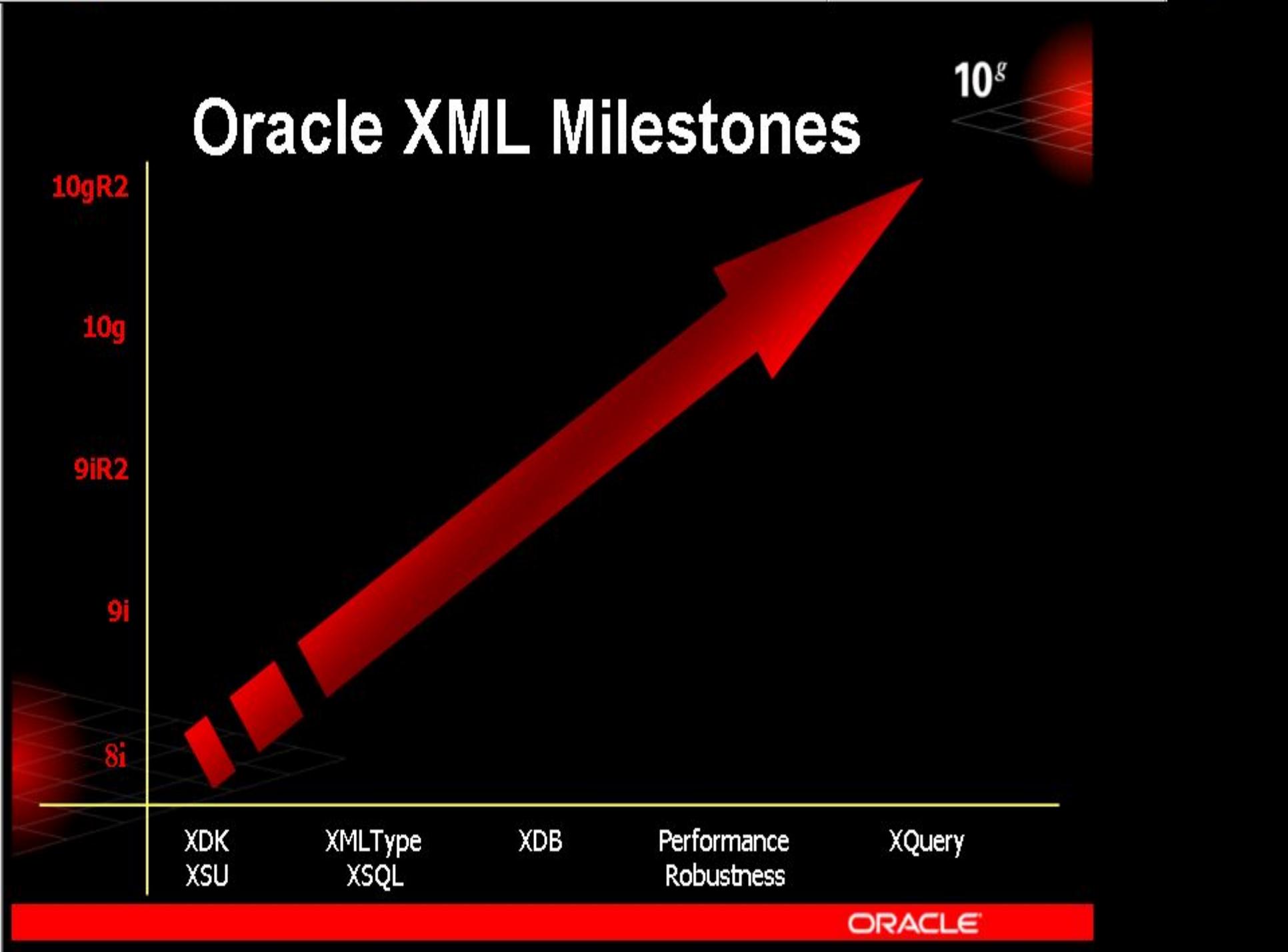
XMLType
XSQL

XDB

Performance
Robustness

XQuery

ORACLE



Oracle XML Milestones

10gR2

10g

9iR2

9i

8i

10^g

XDK
XSU

XMLType
XSQL

XDB

Performance
Robustness

XQuery

Верстовые столбы ORACLE.
Возврат к семантике!

ORACLE