

Методологии проектирования БД

Подходы к моделированию данных

Формальный подход

Выявление полного перечня хранимых атрибутов и правил бизнеса.

Определение схемы универсального отношения.

Выявление множества ФЗ, существующих в универсальном отношении.

Нормализация универсального отношения до требуемого уровня.

Результат нормализации

Система проекций универсального отношения, связанных по типу “родитель – потомок”.

Проекции

Модели объектов ПО или связей объектов.

Основной недостаток

Необходимость проведения *детального анализа ПО до начала* проектирования логического макета БД.

Методики, основанные на формальном подходе, непригодны для решения сложных задач.

Семантический подход

Анализ ПО и проектирование логического макета БД выполняются параллельно.

Этап 0

Формирование общих представлений о компонентах бизнеса и их отношениях.

Описание предметной области

Этап 1

Идентификация основных сущностей и связей (ER-уровень модели данных).

Без детальной информации о свойствах объектов ПО и их отношениях.

Этап 2

Детализация модели до уровня идентификаторов экземпляров сущностей и типов связей (KB-уровень модели данных).

Окончательно определяется состав сущностей, ограничения целостности сущности и ссылок. Результат – логический макет БД с точностью до ключей.

Этап 3

Окончательное определение схем отношений (FA-уровень модели данных).

Как правило, все отношения находятся в 3НФ.

На каждом этапе возможно согласование представлений аналитика с представлениями конечных пользователей.

Методологии, основанные на семантическом подходе

Поддерживают параллельное развитие процессов анализа ПО и проектирования структуры БД и ограничений целостности данных.

Имеют простые и строго определённые графические языки описания данных, обеспечивающие однозначную интерпретацию модели.

Обеспечивают возможность согласования представлений аналитика и требований конечных пользователей на всех этапах разработки.

Включают компьютерные средства поддержки проектирования на всех этапах процесса от формулировки замысла до выпуска проектной документации.

Использование семантического подхода

снижает трудозатраты на ранних этапах проектирования, упрощает и облегчает верификацию моделей и обеспечивает создание высококачественных спецификаций СБД.

МЕТОДОЛОГИЯ IDEF1X

Методология моделирования данных.
Одна из методологий семейства **ICAM** (Integrated Computer-Aided Manufacturing).

Методология IDEF1X определяет

- последовательность этапов моделирования данных,
- требования к проектной документации на всех этапах,
- язык описания данных,
- правила взаимодействия участников проекта.

Определение модели данных IDEF1X

Модель данных есть графическое и текстуальное представление, идентифицирующее потребности организации в данных для достижения её целей, выполнения функций и определения стратегий управления.

Модель данных идентифицирует сущности, домены, атрибуты и связи между данными и поддерживает единый концептуальный взгляд на данные и их взаимосвязи.

Компоненты языка описания данных IDEF1X:

графический – средства создания диаграмм, показывающих структуру и взаимосвязи данных;

текстовый – правила создания и ведения текстовых документов, уточняющих и поясняющих графическую модель.

Цель моделирования

На основании локальных представлений пользователей найти обобщенное представление информации, свойственное природе предметной области как целому.

Задачи анализа предметной области

Уяснить цели предприятия, для которого создаётся БД.

Изучить процессы, направленные на достижение целей.

Выявить потоки данных, возникающие в процессах.

Выявить правила бизнеса, принятые на предприятии.

Выделить объекты, вовлеченные в деятельность.

Выявить связи (отношения) объектов.

Результат анализа:

спецификации логического макета БД, содержащие определения

- базовых отношений и
- ограничений целостности данных.

Уровни графической модели данных IDEF1X

Entity-Relationship Level (ER)

Уровень наименее детального представления информации. Используется на начальной стадии моделирования. На ER-уровне сущности и связи представлены на диаграмме только именами. Их свойства ещё не выяснены и не поняты.

Key-Based Level (KB)

Диаграмма показывает имена первичных и внешних ключей сущностей и спецификации связей.

Fully-Attributed Level (FA)

Диаграмма показывает имена всех атрибутов сущностей и связей и полностью определяет структуру и взаимосвязи данных.

Результат моделирования

ФА-диаграмма - графическое представление структуры РБД содержащее полностью определённые схемы отношений.

Словарь данных – набор определений имён сущностей, атрибутов, доменов и других элементов модели.

Нотации и правила языка IDEF1X

Обозначения сущностей и атрибутов

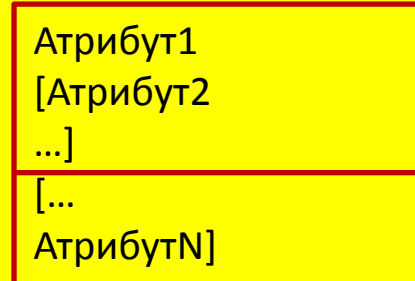
Независимая сущность
<ИМЯ>[/<НОМЕР>]



Зависимая сущность
<ИМЯ>[/<НОМЕР>]



Атрибуты
<ИМЯ>[/<НОМЕР>]



Первичный
ключ

Прочие
атрибуты

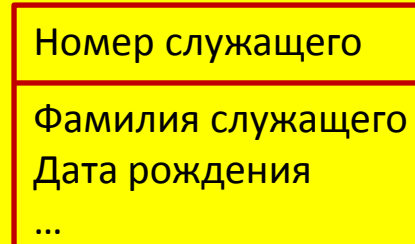
СЛУЖАЩИЙ/25



ПОСТАВКА/39



СЛУЖАЩИЙ/25



На ER-уровне

независимые и зависимые сущности не различаются,
атрибуты не указываются,
имена сущностей вписываются в обозначающие их прямоугольники.

Домены определяются независимо от сущностей и диаграмм.

Правила именования и определения сущностей и атрибутов

Сущности, атрибуты и домены **обязательно** именуются.
Именем может быть **только имя существительное**, возможно с определениями.
В качестве имён допускаются аббревиатуры и акронимы.

Имя должно быть уникальным и осмысленным.
Формальное определение смысла имени **обязательно включается в глоссарий**.

Имя сущности, атрибута или домена **должно иметь единственный смысл** и этот смысл **всегда** должен выражаться этим именем.
Тот же смысл не может вкладываться в другое имя, если оно не является псевдонимом или синонимом основного.

Сущности и атрибуты **всегда** именуются в **единственном числе**.
Имя должно относиться к **одному экземпляру** сущности или **одному** значению атрибута.

Правила для атрибутов (основные)

На диаграммах КВ- и FA-уровней каждая сущность имеет не менее одного атрибута.

Каждый атрибут может быть **собственным** атрибутом сущности или **присоединённым**, полученным от другой сущности через связь (внешний ключ).

Каждый атрибут является **собственным** атрибутом **точно одной** сущности.

Присоединённый атрибут должен быть частью первичного ключа передавшей его родительской (или родовой) сущности.

Присоединённый атрибут помечается символом **"(FK)"**, следующим за именем атрибута.

Каждый экземпляр сущности должен иметь определённое (не NULL) значение каждого атрибута, являющегося частью первичного ключа.

Не может быть экземпляра сущности, имеющего более чем одно значение какого-либо из атрибутов.

На КВ-диаграммах показываются только атрибуты, входящие в состав первичных, альтернативных и внешних ключей.

Атрибуты альтернативных ключей обязательно помечаются символом **"(AKn)"**, где **n** - номер альтернативного ключа.

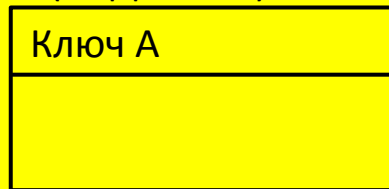
На FA-диаграмме показываются **все** атрибуты **каждой** сущности.

Соединение

Ассоциация двух сущностей или двух экземпляров одной и той же сущности

Один из двух видов связей, используемых в IDEF1X.

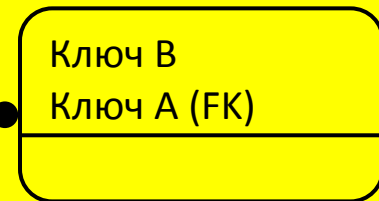
А (РОДИТЕЛЬ)



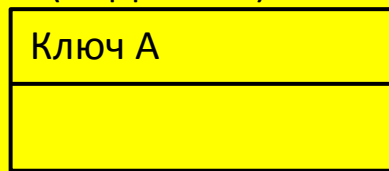
а) Идентифицирующее соединение

<имя от родителя>/
<имя от потомка>

В (ПОТОМОК)



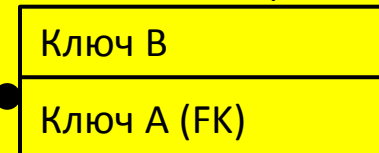
А (РОДИТЕЛЬ)



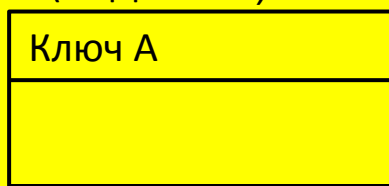
**б) Обязательное
неидентифицирующее соединение**

<имя от родителя>/
<имя от потомка>

В (ПОТОМОК)



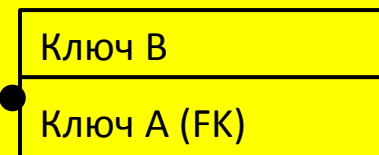
А (РОДИТЕЛЬ)



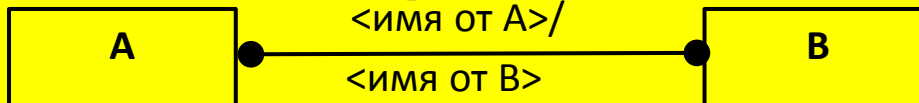
**в) НЕобязательное
неидентифицирующее соединение**

<имя от родителя>/
<имя от потомка>

В (ПОТОМОК)



Неспецифическое соединение



**Допускается
только на ER-уровне**

Правила именования соединений (основные)

Соединению присваивается имя, выражаемое глагольным оборотом.
Имя зрительно привязывается к дуге, изображающей соединение.
Имена соединений могут быть неуникальными в пределах диаграммы.

Имя каждого соединения одной и той же пары сущностей
должно быть уникальным во множестве имён связей этих сущностей.

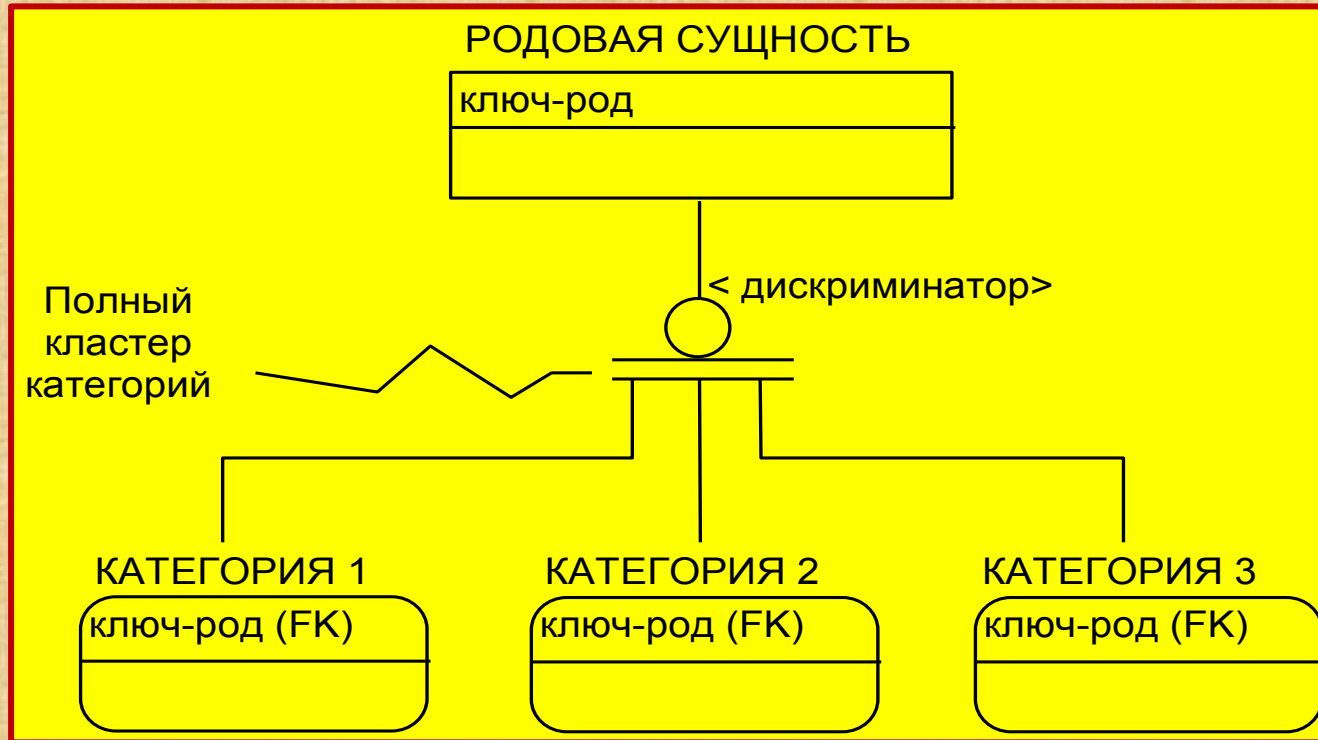
Имена специфических соединений выбираются так,
чтобы можно было построить осмысленную фразу, составленную из
имени родительской сущности,
имени связи,
выражения кардинальности и
имени потомка.

Связь может быть поименована “от родителя” и от “потомка”.
Имя “от родителя” обязательно!

Если связь не именуется со стороны потомка, то имя “от родителя”
должно выбираться так, чтобы связь легко читалась и со стороны потомка.

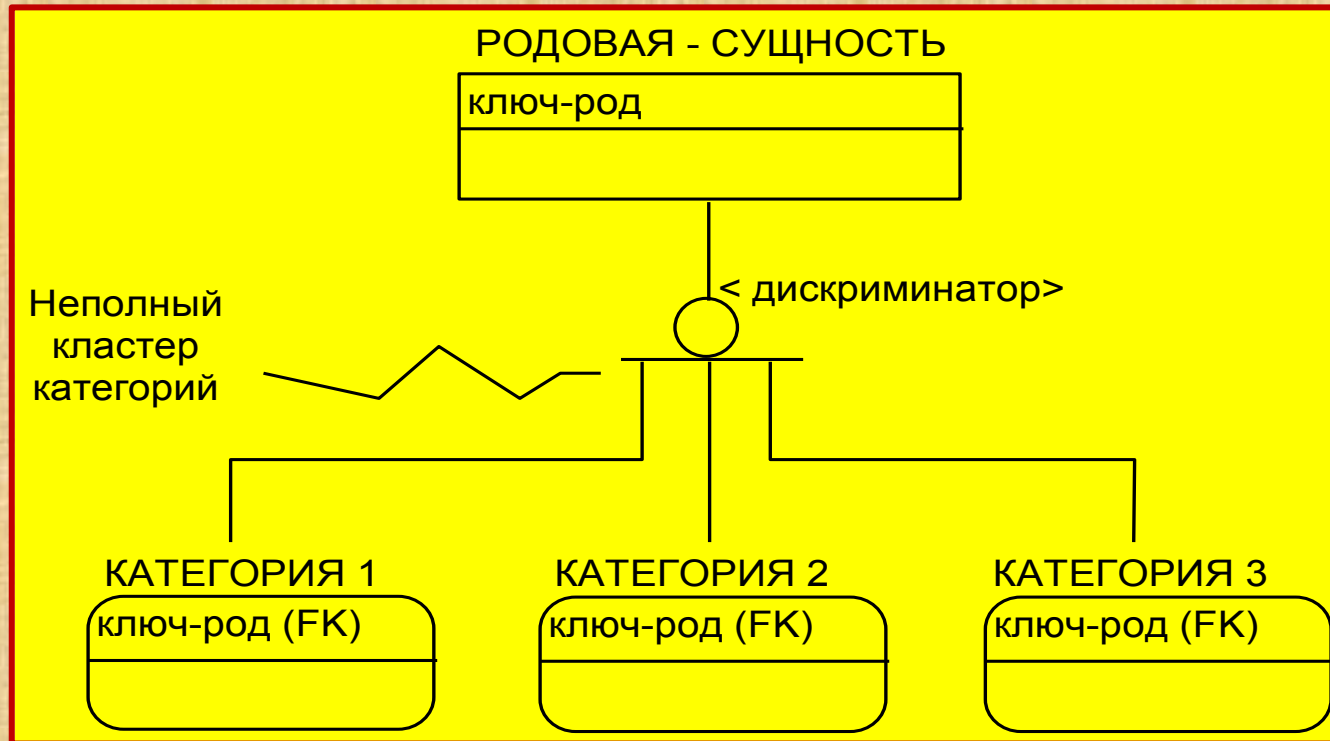
Неспецифические соединения обязательно именуются с обеих сторон.

Связи категоризации



Полный кластер категорий

Связи категоризации



Неполный кластер категорий

Правила для связей категоризации

Категория может иметь только одну родовую сущность.
Она может принадлежать только одному кластеру категорий.

Категория в одной категоризационной связи может быть родовой сущностью в другой категоризационной связи.

Сущность может быть родовой для любого числа кластеров категорий.

Все экземпляры одной категории имеют одно и то же значение дискриминатора, и все экземпляры различных категорий должны иметь различные значения дискриминатора.

Не существует двух кластеров категорий одной и той же родовой сущности, имеющих один и тот же дискриминатор.

Дискриминатором полного кластера категорий не может быть атрибут, значения которого могут быть неопределёнными.

Первичный ключ любой категории должен совпадать с первичным ключом родовой сущности. Однако в качестве имён ключевых атрибутов категорий могут использоваться имена ролей.

[https://intuit.ru/studies/courses/93/93/lecture/28079#:~:text=Методология%20IDEF1X%20-%20один%20из,структуру%20информации%20об%20объектах%20системы\)](https://intuit.ru/studies/courses/93/93/lecture/28079#:~:text=Методология%20IDEF1X%20-%20один%20из,структуру%20информации%20об%20объектах%20системы))