# Sistemas de Banco de Dados Projeto, implementação e gerenciamento

Capítulo 6

Modelagem de dados avançada

# Objetivos

#### Neste capítulo, você aprenderá:

- Sobre o modelo de entidade-relacionamento estendido (EER)
- Como os grupos de entidades são utilizados para representar múltiplas entidades e relacionamentos
- As características de uma boa chave primária e como selecioná-la
- De que modo utilizar soluções flexíveis para casos especiais de modelagem de dados
- Que problemas verificar ao desenvolver modelos de dados com base em diagramas EER

# O Modelo Entidade-relacionamento Estendido

- Resultado da adição de mais estruturas semânticas ao modelo original de entidaderelacionamento (ER)
- Um diagrama que utilize esse modelo é chamado diagrama EER (EERD, extended entity relationship diagram, ou seja, diagrama entidade-relacionamento estendido).

# Supertipos e Subtipos de Entidades

#### Supertipo de entidade

- É uma entidade genérica que se relaciona com um ou mais subtipos de entidade
- Contém todas as características comuns

#### Subtipos de entidade

Contém características específicas

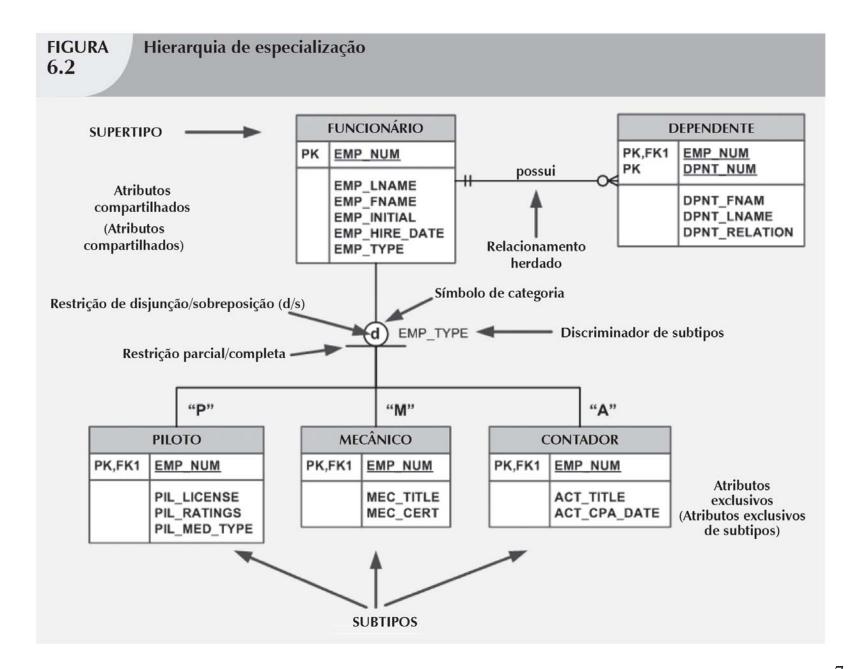
# FIGURA 6.1

#### Nulos criados por atributos exclusivos

EMP_NUM	EMP_LNAME	EMP_FNAME	EMP_INITIAL	EMP_LICENSE	EMP_RATINGS	EMP_MED_TYPE	EMP_HIRE_DATE
100	Kolmycz	Xavier	T				15-Mar-88
101	Lewis	Marcos		ATP	SEL/MEL/Instr/CFII	1	25-Apr-89
102	Vandam	Jean					20-Dec-93
103	Jones	Victoria	R				28-Aug-03
104	Lange	Edith		ATP	SELMEL/Instr	1	20-Oct-97
105	Williams	Gabriel	U	COM	SEL/MEL/Instr/CFI	2	08-Nov-97
106	Duzak	Mario		COM	SELMELAnstr	2	05-Jan-04
107	Diante	Venite	L				02-Jul-97
108	Wiesenbach	Joni					18-Nov-95
109	Travis	Brett	T	COM	SEL/MEL/SES/Instr/CFII	1	14-Apr-01
110	Genkazi	Stan					01-Dec-03

# Hierarquia de Especialização

- Representa a disposição de supertipos de entidades de nível superior e subtipos de entidade de nível inferior
- Os relacionamentos, às vezes são descritos em termos de relacionamentos "é um"
- Em uma hierarquia de especialização pode ocorrer um subtipo apenas no interior do contexto de um supertipo
- Cada subtipo pode ter apenas um supertipo ao qual está diretamente relacionado
- Uma hierarquia de especialização pode apresentar vários níveis de relacionamentos supertipo/subtipo



### Herança

- Permite que um subtipo de entidade herde os atributos e relacionamentos do supertipo
- Todos os subtipos de entidades herdam seu atributo de chave primária do respectivo supertipo
- No nível da implementação, o supertipo e seu(s) subtipo(s) representados na hierarquia de especialização mantêm um relacionamento 1:1
- Os subtipos de entidade herdam todos os relacionamentos de que o supertipo participa
- Em hierarquias de especialização com vários níveis de supertipos/subtipos, um subtipo de nível mais baixo herda todos os atributos e relacionamentos de todos os seus supertipos de nível superior

### FIGURA 6.3

#### Relacionamento supertipo/subtipo entre FUNCIONÁRIO e PILOTO

#### Nome da tabela: FUNCIONÁRIO

EMP_NUM	EMP_LNAME	EMP_FNAME	EMP_INITIAL	EMP_HIRE_DATE	EMP_TYPE
100	Kolmycz	Xavier	T	15-Mar-88	
101	Lewis	Marcos		25-Apr-89	P
102	Vandam	Jean		20-Dec-93	А
103	Jones	Victoria	R	28-Aug-03	
104	Lange	Edith		20-Oct-97	P
105	Williams	Gabriel	U	08-Nov-97	P
106	Duzak	Mario		05-Jan-04	P
107	Diante	Venite	L	02-Jul-97	М
108	Wiesenbach	Joni		18-Nov-95	M
109	Travis	Brett	T	14-Apr-01	Р
110	Genkazi	Stan		01-Dec-03	А

#### Nome da tabela: PILOTO

EMP_NUM	PIL_LICENSE	PIL_RATINGS	PIL_MED_TYPE
101	ATP	SEL/MEL/Instr/CFII	1
104	ATP	SELMELInstr	1
105	COM	SEL/MEL/Instr/CFI	2
106	COM	SEL/MEL/Instr	2
109	COM	SEL/MEL/SES/Instr/CFII	1

## Discriminador de Subtipos

#### Discriminador de subtipo

- O atributo de uma entidade supertipo que determina a qual subtipo a ocorrência de supertipo está relacionada
- A condição-padrão de comparação para o atributo discriminador de subtipos é a comparação de igualdade
- Pode haver situações em que o discriminador não se baseia necessariamente nesse tipo de comparação

### Restrições de Disjunção e Sobreposição

#### Subtipos disjuntos

- Também conhecidos como subtipos de não sobrepostos
- Contêm um subconjunto exclusivo do conjunto do supertipo

#### Subtipos sobrepostos

Contêm subconjuntos não exclusivos do conjunto da entidade supertipo

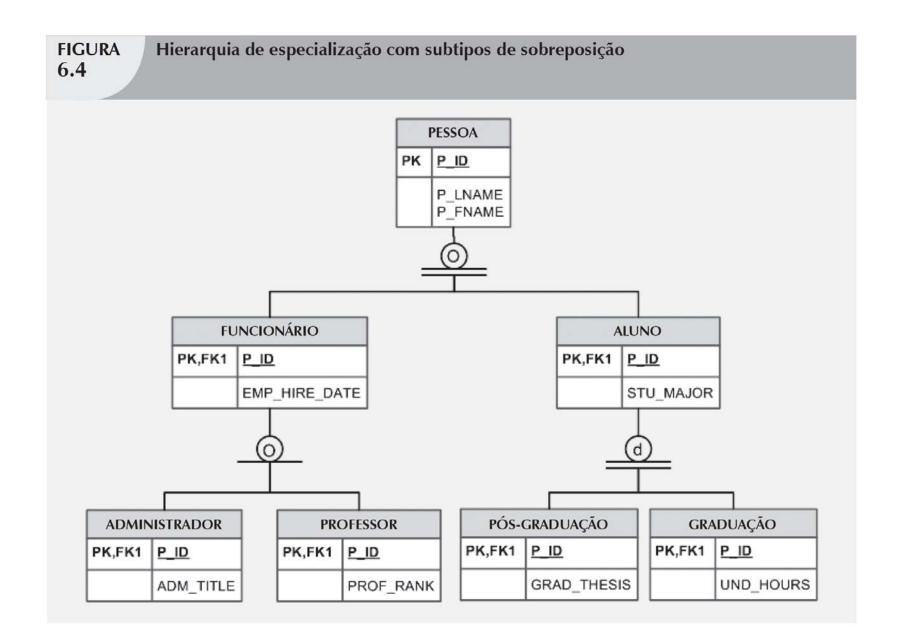


TABELA 6.1 Atributos discriminadores com subtipos sobrepostos

ATRIBUTOS	DISCRIMINADORES	COMENTÁRIO	
PROFESSOR ADMINISTRADOR		COMENTARIO	
"S"	"N"	O Funcionário é membro do subtipo Professor.	
"N"	"S"	O Funcionário é membro do subtipo Administrador.	
"S"	"S"	O Funcionário é tanto Professor como Administrador.	

## Restrição de Integralidade

- Especifica se cada ocorrência de supertipo de entidade também deve ser membro de, pelo menos, um subtipo
- Pode ser parcial ou total
  - Representada por um círculo sobre uma única linha
  - Podem haver algumas ocorrências de supertipos que não sejam membros de nenhum subtipo
- Integralidade total
  - Representada por um círculo sobre linha dupla
  - Toda ocorrência de supertipo deve ser membro de, pelo menos, um subtipo

TABELA 6.2 Cenários de restrição de hierarquia de especialização

TIPO	RESTRIÇÃO DE DISJUNÇÃO	RESTRIÇÃO DE SOBREPOSIÇÃO
Parcial	O supertipo possui subtipos opcionais.	O supertipo possui subtipos opcionais.
	O discriminador de subtipos pode ser nulo.	Os discriminadores de subtipos podem ser nulos.
	Os conjuntos de subtipos são exclusivos.	Os conjuntos de subtipos não são exclusivos.
Total	Todas as ocorrências de supertipos são membros de (pelo menos) um subtipo.	Todas as ocorrências de supertipos são membros de (pelo menos) um subtipo.
	O discriminador de subtipos não pode ser nulo.  Os conjuntos de subtipos são exclusivos.	Os discriminadores de subtipos não podem ser nulos.
		Os conjuntos de subtipos não são exclusivos.

# Especialização e Generalização

#### Epecialização

- Processo de identificação, de cima para baixo, dos subtipos de entidade de nível inferior, a partir de um supertipo de nível superior
- Baseia-se no agrupamento de características e relacionamentos exclusivos dos subtipos

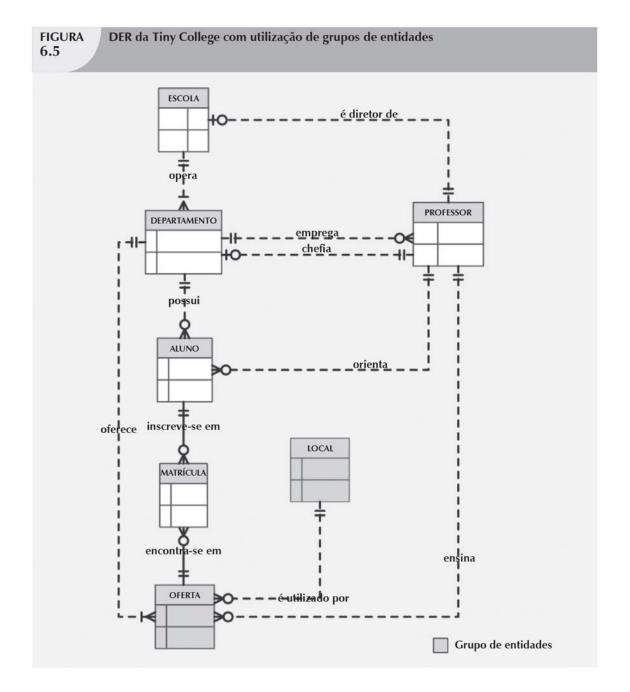
### Especialização e Generalização (cont.)

#### Generalização

- Processo de identificação, de baixo para cima, de um supertipo de entidade mais genérico de nível superior, a partir de subtipos de nível inferior
- Baseia-se no agrupamento de características e relacionamentos comuns dos subtipos

## Agrupamento de Entidades

- Tipo de entidade "virtual" utilizado para representar várias entidades e relacionamentos no DER
- É considerado "virtual" ou "abstrato" no sentido de que não é realmente uma entidade no DER final
- Trata-se de uma entidade temporária utilizada para representar várias entidades e relacionamentos
- Elimina consequências indesejadas
  - Evitar a apresentação de atributos quando se utilizam grupos de entidades



# Integridade de Entidades: Seleção de Chaves Primárias

- A característica mais importante de uma entidade é sua chave primária
  - Um único atributo ou uma combinação de atributos
- A função da chave primária é garantir a integridade de entidade
- As chaves primárias e estrangeiras atuam em conjunto para implementar relacionamentos no modelo relacional
- A importância de seleção adequada de chaves primárias possui um efeito direto sobre a eficiência e a eficácia da implementação do banco de dados

### Chaves Naturais e Chaves Primárias

- A chave natural ou identificador natural é um identificador real utilizado para distinguir objetos reais
  - É familiar aos usuários finais e faz parte de seu vocabulário comercial cotidiano
- Normalmente, se uma entidade tiver um identificador natural, o modelador de dados deve utilizá-lo como chave primária da entidade que estiver sendo modelada

### Diretrizes de Chaves Primárias

- É uma designação ou uma combinação de atributos que identifica exclusivamente as instâncias de um conjunto de entidades
- Sua principal função é identificar exclusivamente uma instância ou linha de entidade em uma tabela
- A função da chave primária é garantir a integridade de entidade, não "descrever" a entidade
- As chaves primárias e as chaves estrangeiras são utilizadas para implementar relacionamentos entre entidades
  - É feita, principalmente, nos "bastidores", oculta dos usuários finais

TABELA 6.3 Características desejáveis de chaves principais

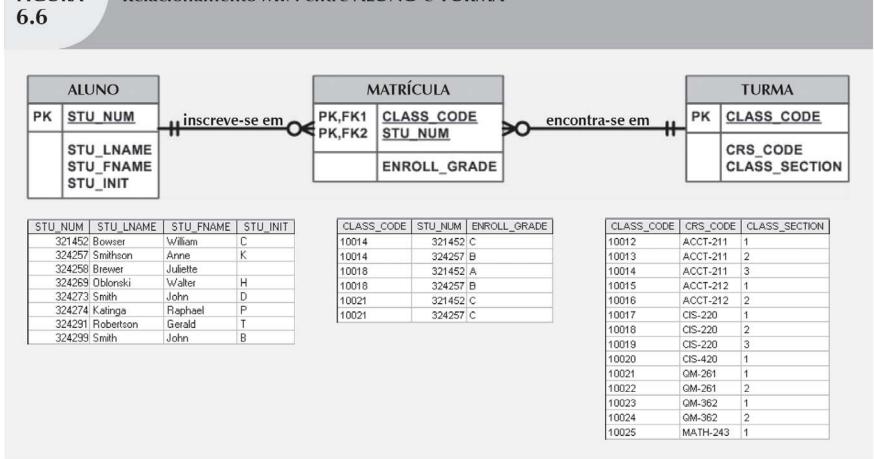
CARACTERÍSTICAS DE PK	MOTIVO
Valores exclusivos	A PK deve identificar de modo exclusivo cada instância de entidade. Uma chave primária deve ser capaz de garantir valores exclusivos. Ela não pode conter nulos.
Não intuitiva	A PK não deve apresentar conteúdo semântico incorporado (não deve se relacionar a um fato). Um atributo com conteúdo semântico provavelmente é mais bem utilizado como característica descritiva da entidade e não como identificador. Em outras palavras, seria preferível a ID de aluno 650973 como chave primária do que Smith, Martha L. Em resumo, a PK não deve ter conteúdo factual.
Sem alteração com o passar do tempo	Se um atributo possui conteúdo semântico, pode estar sujeito a atualizações. Por isso, nomes não constituem boas chaves primárias. Tendo-se Vickie Smith como chave primária, o que ocorre quando ela se casar? Se a chave primária estiver sujeita a mudança, os valores de chave estrangeira devem ser atualizados, contribuindo, assim, para a carga de trabalho do banco de dados. Além disso, a alteração de um valor de chave primária significa que se está basicamente alterando a identidade de uma entidade. Em resumo, a PK deve ser permanente e imutável.
Preferencialmente de um único atributo	Uma chave primária deve ter o número mínimo de atributos possível (irredutível).  As chaves primárias com um único atributo são desejáveis, mas não necessárias. Elas simplificam a implementação de chaves estrangeiras. Ter chaves primárias com vários atributos pode fazer com que chaves primárias de entidades relacionadas cresçam em razão da possível adição de vários atributos, contribuindo, assim, para a carga de trabalho do banco de dados e tornando mais trabalhosa a codificação (de aplicações).
Preferencialmente numérica	Valores exclusivos podem ser mais bem gerenciados quando são numéricos, pois o banco de dados pode utilizar rotinas internas para implementar um atributo com contador que incremente automaticamente os valores quando da adição de cada nova coluna. Na verdade, a maioria dos sistemas de bancos de dados inclui a possibilidade de utilizar estruturas especiais, como o Autonumber do Microsoft Access, para dar suporte a atributos de chave primária autoincrementados.
Segurança adequada	A chave primária selecionada não deve ser composta de atributos que possam ser considerados um risco ou violação de segurança. Por exemplo, utilizar o número da Previdência Social como PK de uma tabela FUNCIONÁRIO não é uma boa ideia, pois, nos Estados Unidos, ele pode ser usado em fraudes.

### Quando Utilizar Chaves Surrogates como Chaves Primárias

- Esse tipo de chave é especialmente útil em dois casos:
  - Com identificadores de entidades compostas, em que cada combinação de chave primária é permitida somente uma vez no relacionamento M:N
  - Como identificadores de entidades fracas, em que essa entidade possui uma relação de identificação forte com a entidade pai
- Proporciona automaticamente o benefício de garantir que não haja valores duplicados

#### **FIGURA** 6.6

#### Relacionamento M:N entre ALUNO e TURMA



# Quando Utilizar Chaves Surrogates como Chaves Primárias (cont.)

- Uma entidade fraca em uma relação de identificação forte com uma entidade pai normalmente é utilizada para representar uma dessas duas situações:
  - Um objeto real dependente da existência de outro objeto real
  - Um objeto real representado no modelo de dados como duas entidades distintas em um relacionamento de identificação forte
- Aseleção de uma chave primária composta para tipos de entidades compostas ou fracas proporciona benefícios que aprimoram a integridade e a consistência do modelo

# Quando Utilizar Chaves Surrogates como Chaves Primárias (cont.)

- Especialmente útil quando:
  - Não há chave natural
  - A chave candidata selecionada possui conteúdo semântico incorporado ou é muito longa ou trabalhosa

# Quando Utilizar Chaves Surrogates como Chaves Primárias (cont.)

- Quando se utiliza uma chave surrogate, deve-se:
  - Garantir que a chave candidata da entidade em questão funcione adequadamente
    - Por meio da utilização de restrições de "índice exclusivo"
    - Pela "ausência de nulos"

TABELA 6.4 Dados utilizados para o rastreamento de eventos

DATA	HORA_INÍCIO	HORA_FINAL	SALÃO	EVENTO_NOME	FESTA_ CONVIDADOS
17/6/08	11h00	14h00	Allure	Casamento dos Burton	60
17/6/08	11h00	14h00	Bonanza	Escritório Adams	12
17/6/08	15h00	17h30	Allure	Família Smith	15
17/6/08	15h30	17h30	Bonanza	Escritório Adams	12
18/6/08	13h00	15h00	Bonanza	Escoteiros	33
18/6/08	11h00	14h00	Allure	Instituição de Caridade March of Dimes	25
18/6/08	11h00	12h30	Bonanza	Família Smith	12

# Casos de Projetos: Banco de Dados Flexível

- A modelagem e o projeto de dados exigem habilidades adquiridas com a experiência
- A experiência, por sua vez, é adquirida por meio da prática
- Quatro casos especiais enfatizados:
  - A importância de projetos flexíveis
  - A identificação adequada de chaves primárias
  - A disposição das chaves estrangeiras

# Caso de Projeto 1: Implementação de Relacionamentos 1:1

- As chaves estrangeiras funcionam em conjunto com as chaves primárias para identificar adequadamente os relacionamentos no modelo relacional
- Coloque a chave primária do lado "um" como chave estrangeira do lado "muitos"
  - Chave primária: a entidade pai
  - Chave estrangeira: a entidade dependente

# Caso de Projeto 1: Implementação de Relacionamentos 1:1 (cont.)

- Há duas opções para a seleção e o posicionamento da chave estrangeira:
  - Colocar a chave estrangeira em ambas as entidades (não é recomendada)
  - Colocar a chave estrangeira em uma das entidade
    - A chave primária de uma entidade aparece como chave estrangeira da outra

TABELA 6.5 Seleção de chave estrangeira em um relacionamento 1:1

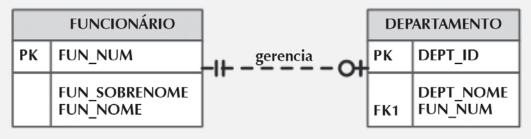
CASO	RESTRIÇÕES DO RELACIONAMENTO ER	AÇÃO
1	Um lado é obrigatório e outro é opcional.	Coloque a PK da entidade do lado obrigatório como FK da entidade do lado opcional e torne a PK obrigatória.
11	Ambos os lados são opcionais.	Selecione a FK que resulta no menor número de nulos ou coloque a FK na entidade em que o papel (do relacionamento) é executado.
III	Ambos os lados são obrigatórios.	Veja o Caso II ou considere a revisão do modelo para garantir que as duas entidades não pertençam a uma única entidade.

### FIGURA 6.7

#### Relacionamento 1:1 entre DEPARTAMENTO e FUNCIONÁRIO

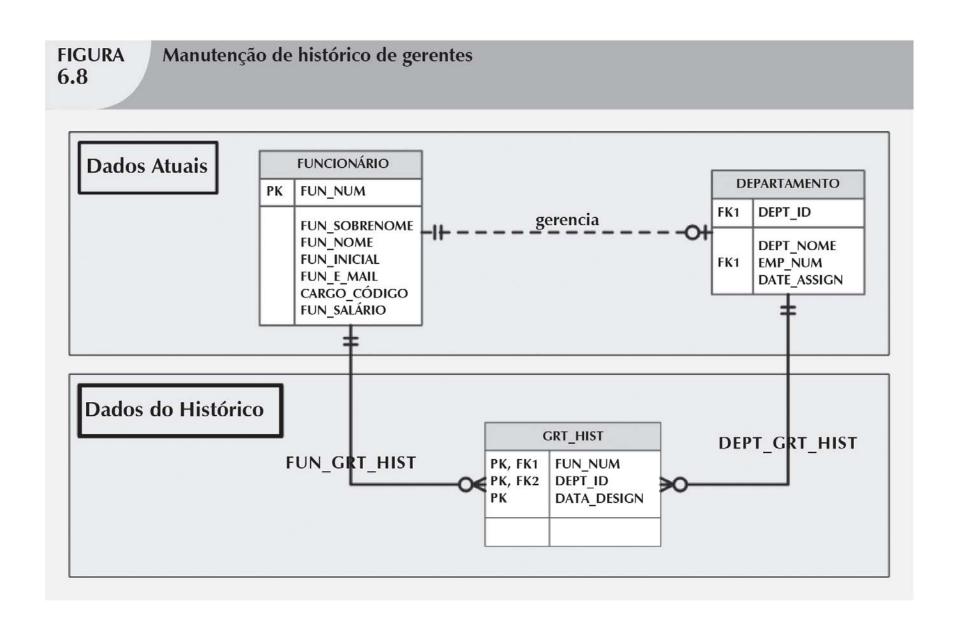
#### Relacionamento um a um (1:1)

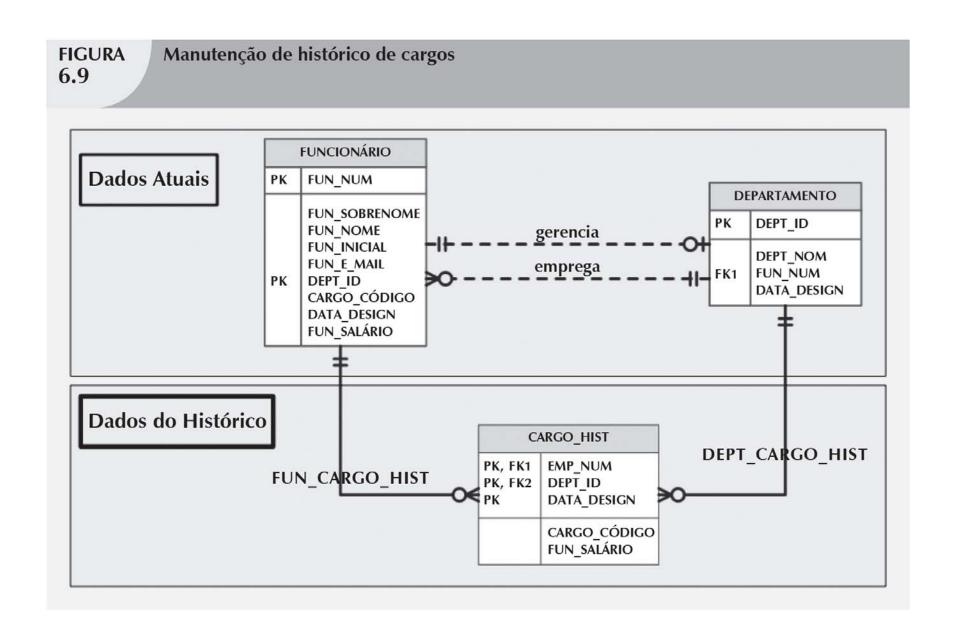
Um FUNCIONÁRIO gerencia zero ou um DEPARTAMENTO; cada DEPARTAMENTO é gerenciado por um FUNCIONÁRIO.



# Caso de Pojeto 2: Manutenção de Histórico de Dados que Variam no Tempo

- Normalmente, as mudanças de dados são gerenciadas substituindo-se o valor existente do atributo pelo novo valor, sem se preocupar com o anterior
- Dados variáveis no tempo
  - Os valores se alteram com o passar do tempo
  - É necessário manter um histórico de alterações
- Manter o histórico de dados variáveis no tempo é equivalente a ter um atributo com diversos valores em sua entidade
- Para modelar esse tipo de dado, deve-se criar uma nova entidade em um relacionamento 1:M com a entidade original
- Essa nova entidade conterá o novo valor e a data da alteração





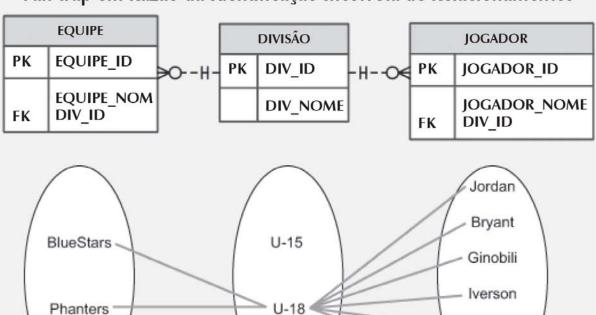
# Caso de Projeto 3: Fan Traps

- Uma armadilha de projeto ocorre quando um relacionamento é identificado de maneira inadequada ou incompleta
  - Representada de um modo inconsistente em relação à realidade
- A mais comum é a fan trap
- Uma fan trap ocorre quando uma entidade está em dois relacionamentos 1:M com outras entidades
  - Produz, assim, uma associação entre as outras entidades que não é expressa no modelo

# FIGURA 6.10

### DER incorreto com problema de fan trap

#### Fan trap em Razão da Identificação Incorreta de Relacionamentos



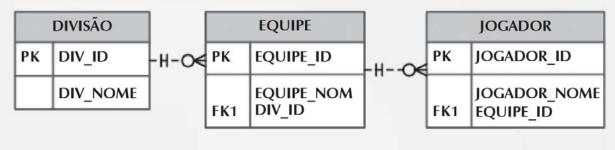
Kidd

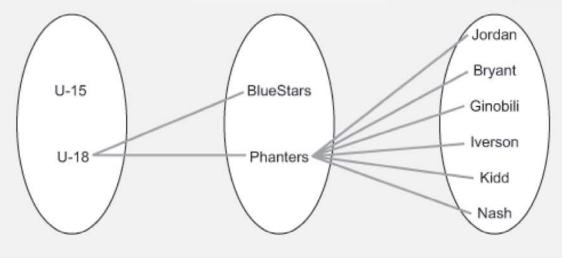
Nash

### FIGURA 6.11

### DER corrigido após remoção da fan trap

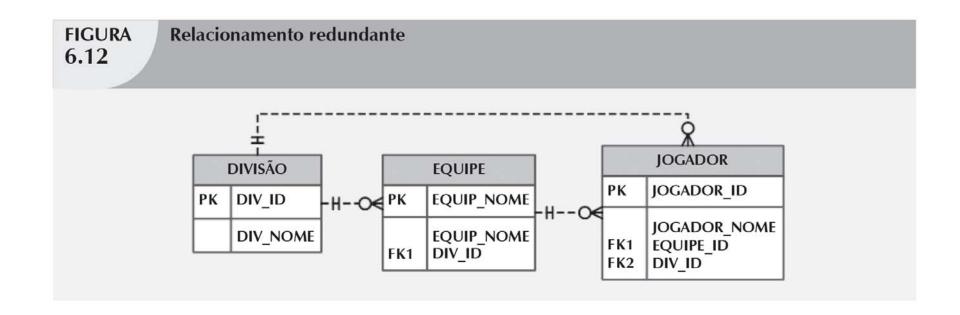
### Fan trap Eliminada pela Identificação Adequada dos Relacionamentos





# Caso de Projeto 4: Relacionamentos Redundantes

- A redundância raramente é boa no ambiente de banco de dados
- Ocorrem quando há vários caminhos de relacionamento entre as entidades relacionadas
- A principal preocupação quanto aos relacionamentos redundantes é que permaneçam consistentes por todo o modelo
- Alguns projetos utilizam esse tipo de relacionamento como um modo de simplificação



# Lista de Verificação de Modelagem de Dados

- A modelagem de dados traduz um ambiente real específico em um modelo de dados
  - Representa dados, usuários, processos e interações reais
- O EERM permite que o projetista adicione maior conteúdo semântico ao modelo
- A lista de modelagem de dados apresentada ajudará a garantir a execução adequada das tarefas de modelagem de dados
- Essa lista se baseia nos conceitos e ferramentas apresentados no início do Capítulo 3

#### TABELA 6.6 Lista de verificação de modelagem de dados

#### **REGRAS DE NEGÓCIO**

- Documente e verifique adequadamente todas as regras de negócio com os usuários finais.
- Garanta que todas as regras sejam escritas de modo preciso, claro e simples, pois devem ajudar a identificar as entidades, atributos, relacionamentos e restrições.
- Identifique a fonte de todas as regras de negócio e garanta que sejam acompanhadas do motivo de sua existência e da data e pessoa(s) responsável(is) por sua verificação e aprovação.

#### **MODELAGEM DE DADOS**

Convenções de nomenclatura: Todos os nomes devem ter comprimento limitado (tamanho dependente do banco de dados).

#### Nomes de Entidades:

- Devem ser substantivos curtos, significativos e familiares aos negócios.
- Devem incluir abreviaturas, sinônimos e alias para cada entidade.
- Devem ser exclusivos no interior do modelo.
- Para entidades compostas, pode incluir uma combinação de nomes abreviados das entidades ligadas por meio dessa entidade.

#### Nomes de Atributos:

- Devem ser exclusivos no interior da entidade.
- Devem utilizar a abreviatura ou prefixo da entidade.
- Devem ser descritivos da característica.
- Devem utilizar sufixos como \_ID, \_NUM ou \_CÓDIGO para atributo de FK.
- Não devem ser uma palavra reservada.
- Não devem conter espaços ou caracteres especiais como @, ! ou &.

#### Nomes de Relacionamentos:

- Devem ser verbos na voz ativa ou passiva que indiquem claramente a natureza do relacionamento.

(1 de 2)

#### **Entidades:**

- Cada entidade deve representar um único assunto.
- Cada entidade deve representar um conjunto de instâncias de entidades distinguíveis.
- Todas as entidades devem estar em 3NF ou superior.
- A granularidade da instância de entidade é claramente definida.
- A PK é claramente definida e dá suporte à granularidade de dados selecionada.

#### **Atributos:**

- Devem ser simples e monovalorados (dados atômicos).
- Devem incluir valores-padrão, restrições, sinônimos e alias.
- Atributos derivados devem ser claramente identificados e incluir fonte(s).
- Não devem ser redundantes, a menos que sejam necessários para precisão de transações, manutenção de histórico ou chave estrangeira.
- Atributos que não são chave devem ser totalmente dependentes do atributo de PK.

#### Relacionamentos:

- Devem identificar claramente seus participantes.
- Devem definir claramente as regras de participação e de cardinalidade.

#### Modelo ER:

- Deve ser validado em relação aos processos esperados: inserções, atualizações e exclusões.
- Devem avaliar onde, quando e como manter um histórico.
- Não devem conter relacionamentos redundantes, exceto quando necessário (veja atributos).
- Devem minimizar a redundância de dados para garantir atualizações em um único local.
- Devem se adequar à regra dos dados mínimos: tudo o que é necessário está à disposição e tudo o que está à disposição é necessário.

### (2 de 2)

### Resumo

- O modelo de entidade-relacionamento estendido (EER) adiciona semântica ao modelo de ER
  - Por meio de supertipos, subtipos e grupos de entidades
  - Um supertipo é um tipo genérico de entidade relacionado a um ou mais subtipos
- Hierarquia de especialização
  - Representa a organização e os relacionamentos entre supertipos e subtipos de entidades
- A herança significa que um subtipo de entidade herda os atributos e relacionamentos do supertipo

- Utiliza-se um discriminador de subtipos para determinar a qual subtipo de entidade a ocorrência de supertipo está relacionada:
  - Integralidade parcial ou total
  - Especialização e generalização
- O grupo de entidades é um tipo de entidade "virtual"
  - Representa várias entidades e relacionamentos no DER
  - É formado pela combinação de várias entidades interrelacionadas e relacionamentos em um único objeto

- As chaves naturais são identificadores existentes na realidade
  - Às vezes, constituem boas chaves primárias
- As chaves primárias devem apresentar as seguintes características:
  - Ter valores exclusivos
  - Não serem intuitivas
  - Não mudar com o passar do tempo
  - Serem preferencialmente numéricas e compostas de um único atributo

- As chaves compostas são úteis para representar
  - Relacionamentos M:N
  - Entidades fracas (de identificação forte)
- As chaves primárias surrogates são úteis quando não há chave natural que possa ser uma chave primária adequada
- Em um relacionamento 1:1, coloque a PK da entidade obrigatória como chave estrangeira da entidade opcional como FK na entidade que provoque o menor número de nulos ou como PK no local em que o papel é executado

- Os dados variáveis no tempo
  - São aqueles cujos valores se alteram com o passar do tempo
  - Cujas necessidades impõem a manutenção de um histórico de alterações de dados
- Para manter o histórico desses dados:
  - Deve-se criar uma entidade que contenha o novo valor, a data de alteração e quaisquer outros dados relevantes sobre tempo
  - Essa entidade mantém um relacionamento 1:M com a entidade para a qual o histórico é sustentado

# Summary (continued)

### Fan trap:

- Quando uma entidade está em dois relacionamentos 1:M com outras entidades
- Há uma associação entre as outras entidades que não é expressa no modelo
- Os relacionamentos redundantes ocorrem quando há vários caminhos de relacionamento entre as entidades relacionadas
  - A principal preocupação quanto aos relacionamentos redundantes é que permaneçam consistentes por todo o modelo
- A lista de verificação de modelagem de dados propicia um modo de o projetista verificar se o DER atende a um conjunto de exigências mínimas